ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mathématique

Luran Bengono Mintogo, Emmanuel de Dieu Nkou and Guy Martial Nkiet

Rates of strong uniform consistency for the k-nearest neighbors kernel estimators of density and regression function

Volume 363 (2025), p. 1035-1046

Online since: 5 September 2025

https://doi.org/10.5802/crmath.784

CENTRE
MERSENNE

The Comptes Rendus. Mathématique *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1778-3569

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mathématique
2025, Vol. 363, p. 1035-1046
https://doi.org/10.5802/crmath.784

Research article / *Article de recherche*
Statistics / *Statistiques*

Rates of strong uniform consistency for the k-nearest neighbors kernel estimators of density and regression function

Vitesses de convergence uniforme presque sûre pour les estimateurs de la densité et la fonction de régression par la méthode des k plus proches voisins à noyau

Luran Bengono Mintogo [a], **Emmanuel de Dieu Nkou** [a] **and Guy Martial Nkiet** [a]

[a] Laboratoire de Probabilités, Statistique et Informatique, Unité de Recherche en Mathématiques et Informatique, Université des Sciences et Techniques de Masuku, BP 813 Franceville, Gabon

E-mails: luranbengono@gmail.com, emmanueldedieunkou@gmail.com, guymartial.nkiet@univ-masuku.org

Abstract. We address the problem of consistency of the k-nearest neighbors kernel estimators of the density and the regression function in the multivariate case. We get the rates of strong uniform consistency on the whole space $\mathbb{R}^p$ for these estimators under specified assumptions.

Résumé. Cet article aborde le problème de la convergence des estimateurs de la densité et de la fonction de régression par la méthode des k plus proches voisins, dans le cas multivarié. Nous déterminons les vitesses de convergence uniforme presque sûre sur tout l'espace $\mathbb{R}^p$, sous des conditions spécifiées.

Keywords. k-nearest neighbors, kernel estimators, density, regression function, strong uniform consistency, rates of convergence.

Mots-clés. k plus proches voisins, estimateurs à noyau, densité, fonction de régression, convergence uniforme presque sûre, vitesse de convergence.

2020 Mathematics Subject Classification. 62G05, 62G07, 62G08.

Manuscript received 15 August 2024, revised 1 April 2025, accepted 15 July 2025.

1. Introduction

Estimation of the density and the regression function are important and classical issues in nonparametric statistics which have been intensively addressed since many years, so leading to an abundant literature. Without a doubt, the most popular estimators that have been tackled in this context are the kernel estimators, namely the Parzen–Rosenblatt estimator of the density and the Nadaraya–Watson estimator of the regression function. However, the practical choice of the bandwidth on which these estimators rely is not straightforward and still a challenging issue. This is why alternative estimators, which do not require to make such a choice, have been proposed. Among them, the k-nearest neighbors (k-NN) kernel estimators have attracted particular attention. They have the same form than the kernel estimators, but with bandwidth replaced by the Euclidean distance between the point to which the estimator is calculated and the k-th nearest neighbor of this point among the observations. Earlier works on this topic go back to [15] for density estimation and to [8] for the case of regression function. These papers established the strong uniform consistency and the strong pointwise consistency, respectively, of the tackled estimators. After these works, some others studying various aspects related to the aforementioned estimators were introduced in the literature. For example, [14] derived expressions describing the asymptotic behavior of the bias and variance of the k-NN density estimates, [4] introduced an adaptative optimal choice of k in multivariate k-NN density and regression estimation and [13] proved strong pointwise consistency for the k-NN estimators of the density and the regression function in the context of α-mixing stationary sequences. The most recent works on the k-NN kernel estimators concern the case of functional data (e.g., [3,6,11,12]), and that of spatial data (see [2]). In fact, these estimators belong to the family of variable bandwidth kernel estimators, where the bandwidth depends on the estimation point and the sample. The use of such adaptive bandwidths has been extensively explored in the literature, with numerous theoretical results, particularly concerning bias reduction (see, e.g., [5]). In this vein, [1] considers the estimation of the density by proposing a kernel estimator of this function at $x = 0$, where the bandwidth depends on the observations through a scalar function on $\mathbb{R}^p$ which is not explicitly assigned, and using the mean squared error (MSE) at 0 as an optimality criterion. [19] proposed a density estimation methodology based on kernels with variable bandwidth, emphasizing the importance of local adaptation to improve estimator accuracy. [10] distinguished two types of variable-bandwidth kernel estimators: those where the bandwidth varies for each data point and those where the bandwidth depends only on the estimation point. The estimators that we tackle combine several key ideas of previous works: local adaptation of the bandwidth as in [1], the distinction between different forms of bandwidth variation as in [10].

Curiously, there is almost no work devoted to determining the convergence rates of the aforementioned k-NN kernel estimators in both the univariate and the multivariate cases. However, [20] derived rates of strong uniform convergence, on any compact subset of $\mathbb{R}$, of the k-NN kernel density estimator, but only in the univariate case. To the best of our knowledge, there is no work dealing with derivation of such rates for the k-NN kernel estimator of the regression function either in the univariate case or in the multivariate case.

In this paper, we address the problem of determining rates of strong uniform consistency for the k-NN kernel estimators of multivariate density and regression function. In Section 2, we define the estimators that will be tackled. For the density, it is the usual k-NN kernel estimator but for the regression function, we slightly modify the classical one as it was done in [22]. Section 3 presents the used assumptions and gives the main results. The proofs of the theorems are postponed in Section 4.

2. The k-NN kernel estimators

For $n \in \mathbb{N}^*$, let $\{(X_i, Y_i)\}_{1\leqslant i\leqslant n}$ be an i.i.d. sample of a pair (X, Y) of random variables valued into $\mathbb{R}^p \times \mathbb{R}$, with $p \geqslant 1$. We denote by f the density of X and, assuming that $\mathbb{E}(|Y|) < +\infty$, we consider the regression function r defined as

$$r(x) = \mathbb{E}\big(Y \,\big|\, X = x\big), \quad x \in \mathbb{R}^p.$$

For estimating f, the k-NN kernel density estimator $\widehat{f}_n$ was introduced (see, e.g., [15]); it is defined as

$$\widehat{f}_n(x) = \frac{1}{n\big(R_n(x)\big)^p}\sum_{i=1}^{n} K\left(\frac{X_i - x}{R_n(x)}\right), \quad x \in \mathbb{R}^p,$$

where $K\colon \mathbb{R}^p \to \mathbb{R}$ is a multivariate kernel, and

$$R_n(x) = \min\left\{h \in \mathbb{R}_+^* \Big/ \sum_{i=1}^{n} \mathbb{1}_{\mathscr{B}(x,h)}(X_i) = k_n\right\}$$

with $\mathscr{B}(x,h) = \{t \in \mathbb{R}^p / \|t - x\| < h\}$, $\|\cdot\|$ denoting the Euclidean norm of $\mathbb{R}^p$. In what precedes, $(k_n)_{n\in\mathbb{N}^*}$ is a sequence of integers such that $k_n \to +\infty$ as $n \to +\infty$. Note that the main difference between this estimator and the usual Parzen–Rosenblatt kernel density estimator is that the bandwidth $R_n(x)$ is random and depends on the X_i's. This estimator is also used for determining an estimator of the regression function r. Indeed, assuming that $f(x) > 0$, one can easily see that

$$r(x) = \frac{g(x)}{f(x)}, \tag{1}$$

where

$$g(x) = \int_{\mathbb{R}} y f_{(X,Y)}(x,y)\,\mathrm{d}y,$$

the function $f_{(X,Y)}$ being the density of (X,Y). Considering the k-NN kernel estimator $\widehat{g}_n$ of g defined as

$$\widehat{g}_n(x) = \frac{1}{n\big(R_n(x)\big)^p}\sum_{i=1}^{n} Y_i K\left(\frac{X_i - x}{R_n(x)}\right), \quad x \in \mathbb{R}^p, \tag{2}$$

it is seen that by replacing in (1) g and f by $\widehat{g}_n$ and $\widehat{f}_n$ respectively, we obtain an estimator of r which just is the one introduced in [8]. We will modify this estimator as done in [22], and repeated in [17]. Specifically, considering a sequence $(b_n)_{n\in\mathbb{N}^*}$ of positive real numbers converging to 0 as $n \to +\infty$, we define

$$\widehat{f}_{b_n}(x) = \max\big(\widehat{f}_n(x), b_n\big),$$

and consider the estimator $\widehat{r}_n$ of r given by:

$$\widehat{r}_n(x) = \frac{\widehat{g}_n(x)}{\widehat{f}_{b_n}(x)}.$$

3. Rates of uniform consistency

In this section, we present our assumptions, then we give the main results that establish rates of strong uniform consistency for the estimators of the density and the regression function.

3.1. *Assumptions*

Assumption 1. The density f of X is bounded and bounded from below: there exists $c_0 > 0$ such that $\inf_{x\in\mathbb{R}^p} f(x) \geqslant c_0$.

Assumption 2. For a given $r \in \mathbb{N}^*$, the density f belongs to the class $\mathscr{C}(c,r)$ of functions $\phi\colon \mathbb{R}^p \to \mathbb{R}$ that are r times differentiable and have r-th derivatives $\frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}$, with $(i_1,\ldots,i_r) \in \{1,\ldots,p\}^r$, satisfying the following Lipschitz condition:

$$\left|\frac{\partial^r \phi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x) - \frac{\partial^r \phi}{\partial x_{i_1}\cdots\partial x_{i_r}}(y)\right| \leqslant c\|x - y\|,$$

where $\|\cdot\|$ denotes the Euclidean norm of $\mathbb{R}^p$.

Assumption 3. The functions g_1 and g_2 defined as $g_1(x) = \mathbb{E}\big(Y\mathbb{1}_{\{Y\geqslant 0\}} \,\big|\, X = x\big)f(x)$ and $g_2(x) = \mathbb{E}\big(-Y\mathbb{1}_{\{Y<0\}} \,\big|\, X = x\big)f(x)$ are bounded and belong to the class $\mathscr{C}(c,r)$ previously defined.

Assumption 4. The kernel $K\colon \mathbb{R}^p \to \mathbb{R}$ satisfies the following properties.

(i) K is bounded, that is $G = \sup_{x\in\mathbb{R}^p} \left|K(x)\right| < +\infty$.

(ii) K is symmetric with respect to 0, that is $K(x) = K(-x)$, $\forall\, x \in \mathbb{R}^p$.

(iii) $\int_{\mathbb{R}^p} K(x)\,\mathrm{d}x_1\cdots\mathrm{d}x_p = 1$.

(iv) K is of order r, that is

$$\int_{\mathbb{R}^p} x_{i_1}\cdots x_{i_\ell} K(x)\,\mathrm{d}x_1\cdots\mathrm{d}x_p = 0$$

for any $\ell \in \{1,\dots,r\}$ and $(i_1,\dots,i_\ell) \in \{1,\dots,p\}^\ell$.

(v)

$$\int_{\mathbb{R}^p} \|x\|^{r+1}\left|K(x)\right|\mathrm{d}x_1\cdots\mathrm{d}x_p < +\infty.$$

(vi) $\forall\, x \in \mathbb{R}^p$, $\forall\, a \in [0,1]$, $K(ax) \geqslant K(x)$.

Assumption 5. The number k_n of neighbors is a sequence of positive integers such that: $k_n \sim \lfloor n^{c_1}\rfloor$, where $c_1 \in \left]\frac{1}{2},1\right[$ and $\lfloor a\rfloor$ denotes the integer part of a.

Assumption 6. The sequence $(b_n)_{n\in\mathbb{N}^*}$ satisfies $b_n \sim n^{-c_2}$ with $c_2 \in \left]0,\frac{1}{10}\right[$.

Assumption 7. There exists a sequence M_n of strictly positive numbers such that $M_n \sim \sqrt{\log(n)}$ and $\max_{1\leqslant i\leqslant n}|Y_i| \leq M_n$.

Assumption 1 has been made several times in the nonparametric statistics literature. For example, it was made in [21]. Assumptions 2 and 3 are classical ones; one can find them in [17,21,22] for the univariate case. Assumption 4(iv) just is the translation to the multivariate case of the fact that the kernel K is of order $r \in \mathbb{N}^*$. It is useful in Taylor's expansion used for deriving the consistency rates. Assumption 4(vi) was made in several works on k-NN kernel estimators (e.g., [2,8]); it is satisfied, for instance, by the Gaussian kernel. Assumption 7 is weaker than boundness assumption; for instance, it has been considered in [16].

3.2. *Main results*

Now, we give the main results of the paper, that is rates of strong uniform consistency of the estimators introduced in Section 2. First, for the k-NN kernel density estimator, we have the following.

Theorem 8. *Under Assumptions 1, 2, 4 and 5, we have:*

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_n(x) - f(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)}{k_n^2}}\right).$$

Remark 9. Rates of strong uniform consistency for this estimator were already obtained in [20], but it was in the univariate case and on any compact subset of $\mathbb{R}$, which is a more restricted framework than ours. In addition, the strong assumption that the kernel has bounded variation on $\mathbb{R}$ was required. The obtained rate involves the term $\log\big(\log(n)\big)$, whereas the rate we obtain just involves $\log(n)$. This difference is due to the different approaches used for obtaining these rates and, therefore, it is due to distinct assumptions. Lemma 1 of [20], which gives the aforementioned rate, relies on a result from [7], which in turn is derived from Kolmogorov's methods. In the opposite, our Theorem 8 relies on an exponential inequality of Talagrand type, which highlights $\log(n)$ instead of $\log\big(\log(n)\big)$.

In order to get the rate for the k-NN kernel estimator of the regression function, we first need to deal with the case of the estimator $\widehat{g}_n$ given in (2). We have the following.

Theorem 10. *Under Assumptions 1, 3, 4, 5 and 7, we have:*

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{g}_n(x)-g(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}}+\sqrt{\frac{n\log(n)M_n^2}{k_n^2}}\right).$$

From this result, we obtain as a consequence the following theorem which gives the rate of strong uniform consistency of the k-NN kernel estimator of the regression function.

Theorem 11. *Under Assumptions 1 to 7, we have:*

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{r}_n(x)-r(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}}+\sqrt{\frac{n\log(n)M_n^2}{k_n^2}}+b_n\right).$$

Remark 12. The term b_n that appears in the above rate is not always negligible compared to the other terms when considering the sequence $(k_n)_n$ that optimizes the rate. Indeed, if $k_n^\star$ denotes the optimal choice of k_n, that is the value k_n that minimizes the rate, it follows from straightforward calculations that

$$k_n^\star = \left(\left(\frac{p}{r+1}\right)n^{\frac{r+1}{p}+\frac{1}{2}}(\log n)^{\frac{1}{2}}M_n\right)^{\frac{1}{\frac{r+1}{p}+1}}.$$

The term b_n is negligible compared to other terms if $b_n = o(T_{1,n})$ and $b_n = o(T_{2,n})$, where

$$T_{1,n} = n^{-\frac{r+1}{p}}(k_n^\star)^{\frac{r+1}{p}} = \left(\frac{p}{r+1}\right)^{\frac{r+1}{r+1+p}} n^{-\frac{r+1}{2(r+1+p)}}\left((\log n)^{\frac{1}{2}}M_n\right)^{\frac{r+1}{r+1+p}}$$

and

$$T_{2,n} = n^{\frac{1}{2}}(k_n^\star)^{-1}(\log n)^{\frac{1}{2}}M_n = \left(\frac{p}{r+1}\right)^{-\frac{p}{r+1+p}} n^{-\frac{r+1}{2(r+1+p)}}(\log n)^{\frac{r+1}{2(r+1+p)}}M_n^{\frac{r+1}{r+1+p}}.$$

Since $b_n \sim n^{-c_2}$, it follows that we must have $n^{-c_2} = o(T_{1,n})$ and $n^{-c_2} = o(T_{2,n})$, that is

$$n^{-c_2} = o\left(n^{-\frac{r+1}{2(r+p+1)}}(\log n^{\frac{p}{r+1}})^{\frac{r+1}{r+1+p}}\right) \tag{3}$$

and

$$n^{-c_2} = o\left(\left(\frac{p}{r+1}\right)^{-\frac{p}{r+1+p}} n^{-\frac{r+1}{2(r+1+p)}}(\log n)^{\frac{r+1}{r+1+p}}\right). \tag{4}$$

Conditions (3) and (4) are satisfied if, and only if, $4r < p-4$. For instance, if the dimension p is in $\{1,\dots,7\}$, then the latter condition cannot be satisfied and, therefore, b_n is not negligible compared to the other terms.

4. Proofs of the theorems

In this section, we give the proofs of the main results of the paper. First, a result useful for proving the main theorems is established in Lemma 13. Then, the proofs of Theorems 8, 10 and 11 are given.

4.1. *A preliminary result*

Let $h\colon \mathbb{R}\to\mathbb{R}$ be a measurable function for which there exists a sequence $(\eta_n)_{n\in\mathbb{N}^*}$ such that $|y|\leqslant M_n \Rightarrow \left|h(y)\right|\leqslant \eta_n$, and $1\leqslant\eta_n\leqslant M_n$ for n large enough. We put

$$\widehat{\varphi}_n(x) = \frac{1}{n\mathscr{D}_{1,n}^p}\sum_{i=1}^n h(Y_i)K\left(\frac{X_i-x}{\mathscr{D}_{2,n}}\right),$$

where $(\mathscr{D}_{1,n})_{n\in\mathbb{N}^*}$ and $(\mathscr{D}_{2,n})_{n\in\mathbb{N}^*}$ are sequences satisfying

$$\mathscr{D}_{1,n}^p \geqslant C_1 n^{-1}k_n, \qquad \mathscr{D}_{2,n}^p \leqslant C_2 n^{-1}k_n, \tag{5}$$

for n large enough and some $C_1 > 0$ and $C_2 > 0$, and

$$\left|\frac{\mathscr{D}_{2,n}^p}{\mathscr{D}_{1,n}^p} - 1\right| \sim n^{-\frac{r+1}{p}}. \tag{6}$$

Considering

$$\varphi(x) = \int_{\mathbb{R}} h(y) f_{(X,Y)}(x,y)\,\mathrm{d}y,$$

where $f_{(X,Y)}$ is the density of the pair (X,Y), we have the following.

Lemma 13. *Under the conditions* (i) *to* (v) *of Assumption 4, if φ is bounded and belongs to the class $\mathscr{C}(c,r)$ defined in Assumption 1, we have:*

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{\varphi}_n(x) - \varphi(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)\eta_n^2}{k_n^2}}\right).$$

Proof. It suffices to prove the two following properties:

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{\varphi}_n(x) - \mathbb{E}\big(\widehat{\varphi}_n(x)\big)\right| = O_{a.s.}\left(\sqrt{\frac{n\log(n)\eta_n^2}{k_n^2}}\right) \tag{7}$$

and

$$\sup_{x\in\mathbb{R}^p}\left|\mathbb{E}\big(\widehat{\varphi}_n(x)\big) - \varphi(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}}\right). \tag{8}$$

Proof of (7)**.** From the class of functions

$$\mathscr{G}_n = \left\{\psi_x\colon (t,y)\in\mathbb{R}^p\times[-M_n,M_n]\mapsto \psi_x(t,y) = \frac{h(y)}{n\mathscr{D}_{1,n}^p}K\left(\frac{t-x}{\mathscr{D}_{2,n}}\right),\ x\in\mathbb{R}^p\right\},$$

we use a similar reasoning than in the proof of [17, Theorem 3.1, p. 1299]. Since for any $\psi_x\in\mathscr{G}_n$, and for n large enough,

$$\left|\psi_x(t,y)\right| \leqslant \frac{G\left|h(y)\right|}{n\mathscr{D}_{1,n}^p} \leqslant \frac{G\eta_n}{C_1k_n},$$

it follows

$$\mathbb{E}\Big[\left|\psi_x(X_j,Y_j)\right|\Big] \leqslant \frac{G\eta_n}{C_1k_n} =: U_n, \qquad \mathbb{E}\big[\psi_x^2(X_j,Y_j)\big] \leqslant \frac{G^2\eta_n^2}{C_1^2k_n^2} =: \sigma_n^2.$$

We can apply Talagrand's inequality (see [18] and [9, Proposition 2.2]): there exist $A>0$, $K_1>0$ and $K_2>0$, such that for all t satisfying

$$t \geqslant K_1\left[U_n\log\frac{AU_n}{\sigma_n} + \sqrt{n}\sigma_n\sqrt{\log\frac{AU_n}{\sigma_n}}\right] = K_1U_n\Big[\log(A) + \sqrt{n}\sqrt{\log(A)}\Big],$$

one has

$$P\left\{\sup_{\psi_x\in\mathscr{G}_n}\left|\sum_{i=1}^n\Big\{\psi_x(X_i,Y_i) - \mathbb{E}\big(\psi_x(X,Y)\big)\Big\}\right| > t\right\}$$
$$\leqslant K_2\exp\left\{-\frac{1}{K_2}\frac{t}{U_n}\log\left(1 + \frac{tU_n}{K_2\left(\sqrt{n}\sigma_n + U_n\sqrt{\log\frac{AU_n}{\sigma_n}}\right)^2}\right)\right\},$$

that is

$$\begin{aligned} P\left\{\sup_{x\in\mathbb{R}^p}\left|\widehat{\varphi}_n(x) - \mathbb{E}\big(\widehat{\varphi}_n(x)\big)\right| > t\right\} &\leqslant K_2\exp\left\{-\frac{1}{K_2}\frac{t}{U_n}\log\left(1 + \frac{tU_n}{K_2\left(\sqrt{n}\sigma_n + U_n\sqrt{\log\frac{AU_n}{\sigma_n}}\right)^2}\right)\right\} \\ &= K_2\exp\left\{-\frac{1}{K_2}\frac{C_1tk_n}{G\eta_n}\log\left(1 + \frac{C_1tk_n}{nK_2G\eta_n\big(1+\sqrt{\log(A)}\big)^2}\right)\right\}. \end{aligned} \tag{9}$$

Let us put $t_n = C_3 n^{1/2} k_n^{-1} \log^{1/2}(n)\,\eta_n$ and $L = \frac{K_1 G}{C_1}\sqrt{\log(A)}$, where

$$C_3 > \frac{K_2 G\big(1+\sqrt{\log(A)}\big)}{C_1}. \tag{10}$$

We have, for n large enough,

$$C_3 \log^{1/2}(n) \geqslant 2L \geqslant L(1+1) \geqslant L\left(1+\sqrt{\frac{\log(A)}{n}}\right) = L\left(\frac{\sqrt{n}+\sqrt{\log(A)}}{\sqrt{n}}\right),$$

which implies

$$C_3 \frac{n^{1/2}\log^{1/2}(n)\eta_n}{k_n} \geqslant \frac{K_1 G n^{1/2}\eta_n}{C_1 k_n}\sqrt{\log(A)}\left[\frac{\sqrt{n}+\sqrt{\log(A)}}{\sqrt{n}}\right],$$

that is

$$t_n \geqslant K_1 U_n\Big[\log(A)+\sqrt{n}\sqrt{\log(A)}\Big] = K_1\left[U_n \log\frac{AU_n}{\sigma_n} + \sqrt{n}\sigma_n\sqrt{\log\frac{AU_n}{\sigma_n}}\right].$$

Then, (9) can be applied to t_n so as to yield $P\{\sup_{x\in\mathbb{R}^p}\big|\widehat{\varphi}_n(x)-\mathbb{E}\big(\widehat{\varphi}_n(x)\big)\big| > t_n\} \leqslant u_n$, where

$$u_n = K_2 \exp\left\{-\frac{1}{K_2}\frac{C_1 t_n k_n}{G\eta_n}\log\left(1+\frac{C_1 t_n k_n}{nK_2 G\eta_n\big(1+\sqrt{\log(A)}\big)^2}\right)\right\}.$$

Since $t_n k_n/(n\eta_n) = C_3 n^{-1/2}\log^{1/2}(n) \to 0$ as $n\to+\infty$, it follows that $u_n \sim v_n$, where

$$\begin{aligned}
v_n &= K_2 \exp\left\{-\frac{C_1 t_n k_n}{K_2 G\eta_n} \times \frac{C_1 t_n k_n}{nK_2 G\eta_n\big(1+\sqrt{\log(A)}\big)^2}\right\}\\
&= K_2 \exp\left\{-\left(\frac{C_1 t_n k_n n^{-1/2}}{K_2 G\eta_n\big(1+\sqrt{\log(A)}\big)}\right)^2\right\}\\
&= K_2 \exp\left\{-\left(\frac{C_1 C_3}{K_2 G\big(1+\sqrt{\log(A)}\big)}\right)^2 \log(n)\right\}\\
&= \frac{K_2}{n^{\alpha}},
\end{aligned}$$

with $\alpha = \left(\frac{C_1 C_3}{K_2 G\left(1+\sqrt{\log(A)}\right)}\right)^2$. From (10) we have $\alpha > 1$, thus $\sum_{n=0}^{+\infty} v_n < +\infty$ and $\sum_{n=0}^{+\infty} u_n < +\infty$. Consequently,

$$\sum_{n\geqslant 0} P\left\{\sup_{x\in\mathbb{R}^p}\Big|\widehat{\varphi}_n(x)-\mathbb{E}\big(\widehat{\varphi}_n(x)\big)\Big| > C_3\frac{n^{1/2}\log^{1/2}(n)\,\eta_n}{k_n}\right\} < +\infty,$$

and by Borel–Cantelli lemma we deduce (7).

Proof of (8)**.**

$$\begin{aligned}
\mathbb{E}\big(\widehat{\varphi}_n(x)\big) &= \frac{1}{\mathscr{D}_{1,n}^p}\mathbb{E}\left(h(Y_1)K\left(\frac{X_1-x}{\mathscr{D}_{2,n}}\right)\right)\\
&= \frac{1}{\mathscr{D}_{1,n}^p}\int_{\mathbb{R}^{p+1}} h(y)K\left(\frac{t-x}{\mathscr{D}_{2,n}}\right) f_{(X,Y)}(t,y)\,\mathrm{d}t_1\cdots\mathrm{d}t_p\,\mathrm{d}y\\
&= \frac{1}{\mathscr{D}_{1,n}^p}\int_{\mathbb{R}^p} K\left(\frac{t-x}{\mathscr{D}_{2,n}}\right)\left(\int_{\mathbb{R}} h(y) f_{(X,Y)}(t,y)\,\mathrm{d}y\right)\mathrm{d}t_1\cdots\mathrm{d}t_p\\
&= \frac{1}{\mathscr{D}_{1,n}^p}\int_{\mathbb{R}^p} K\left(\frac{t-x}{\mathscr{D}_{2,n}}\right)\varphi(t)\,\mathrm{d}t_1\cdots\mathrm{d}t_p\\
&= \gamma_n \int_{\mathbb{R}^p} K(u)\varphi(x+\mathscr{D}_{2,n}u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p,
\end{aligned}$$

where $\gamma_n = \frac{\mathscr{D}_{2,n}^p}{\mathscr{D}_{1,n}^p}$. From Taylor's theorem, there exists $\theta \in]0,1[$ such that

$$\varphi(x+\mathscr{D}_{2,n}u) = \varphi(x) + \sum_{k=1}^{r-1} \frac{1}{k!} \sum_{1\leqslant i_1,\dots,i_k\leqslant p} \frac{\partial^k \varphi}{\partial x_{i_1}\cdots\partial x_{i_k}}(x)\, \mathscr{D}_{2,n}^k u_{i_1}\cdots u_{i_k} + \frac{1}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x+\theta\mathscr{D}_{2,n}u)\, \mathscr{D}_{2,n}^r u_{i_1}\cdots u_{i_r}.$$

Hence

$$\begin{aligned}
&\mathbb{E}\big(\widehat{\varphi}_n(x)\big)\\
&= \gamma_n \varphi(x) \int_{\mathbb{R}^p} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&\quad + \gamma_n \sum_{k=1}^{r-1} \frac{1}{k!} \sum_{1\leqslant i_1,\dots,i_k\leqslant p} \frac{\partial^k \varphi}{\partial x_{i_1}\cdots\partial x_{i_k}}(x)\, \mathscr{D}_{2,n}^k \int_{\mathbb{R}^p} u_{i_1}\cdots u_{i_k} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&\quad + \gamma_n \frac{1}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^r \int_{\mathbb{R}^p} \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x+\theta\mathscr{D}_{2,n}u)\, u_{i_1}\cdots u_{i_r} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&= \gamma_n \varphi(x) + \gamma_n \frac{1}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^r \int_{\mathbb{R}^p} \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x+\theta\mathscr{D}_{2,n}u)\, u_{i_1}\cdots u_{i_r} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&= \gamma_n \varphi(x) + \gamma_n \frac{1}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^r \int_{\mathbb{R}^p} \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x+\theta\mathscr{D}_{2,n}u)\, u_{i_1}\cdots u_{i_r} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&\quad - \gamma_n \frac{1}{r!} \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x) \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^r \int_{\mathbb{R}^p} u_{i_1}\cdots u_{i_r} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&= \gamma_n \varphi(x)\\
&\quad + \frac{\gamma_n}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^r \int_{\mathbb{R}^p} \left(\frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x+\theta\mathscr{D}_{2,n}u) - \frac{\partial^r \varphi}{\partial x_{i_1}\cdots\partial x_{i_r}}(x) \right) u_{i_1}\cdots u_{i_r} K(u)\,\mathrm{d}u_1\cdots\mathrm{d}u_p.
\end{aligned}$$

Since φ belongs to $\mathscr{C}(c,r)$, it follows

$$\begin{aligned}
\left|\mathbb{E}\big(\widehat{\varphi}_n(x)\big) - \varphi(x)\right| &\leqslant |\gamma_n - 1|\,\left|\varphi(x)\right| + c\frac{\gamma_n}{r!} \sum_{1\leqslant i_1,\dots,i_r\leqslant p} \mathscr{D}_{2,n}^{r+1}\theta \int_{\mathbb{R}^p} \|u\|\,|u_{i_1}|\cdots|u_{i_r}|\,\left|K(u)\right|\mathrm{d}u_1\cdots\mathrm{d}u_p\\
&\leqslant |\gamma_n - 1|\,\|\varphi\|_\infty + c\frac{\gamma_n}{r!} p^r \mathscr{D}_{2,n}^{r+1} \int_{\mathbb{R}^p} \|u\|^{r+1} \left|K(u)\right|\mathrm{d}u_1\cdots\mathrm{d}u_p.
\end{aligned}$$

Since $\gamma_n \to 1$ as $n \to +\infty$, we have for n large enough $\gamma_n \leqslant 3/2$ and, therefore, $\gamma_n \mathscr{D}_{2,n}^{r+1} \leqslant \frac{3}{2} C_2 n^{-\frac{r+1}{p}} k_n^{\frac{r+1}{p}}$. Thus,

$$\left|\mathbb{E}\big(\widehat{\varphi}_n(x)\big) - \varphi(x)\right| \leqslant |\gamma_n - 1|\,\|\varphi\|_\infty + C_4 \left(\frac{k_n}{n}\right)^{\frac{r+1}{p}}$$

for some $C_4 > 0$. Since $|\gamma_n - 1| \sim n^{-\frac{r+1}{p}}$ it follows that, for n large enough, $\left|\mathbb{E}\big(\widehat{\varphi}_n(x)\big) - \varphi(x)\right| \leqslant C_5 \left(\frac{k_n}{n}\right)^{\frac{r+1}{p}}$ for some $C_5 > 0$, which implies (8). □

4.2. *Proof of Theorem 8*

Considering a sequence $(\beta_n)_{n\in\mathbb{N}^*}$ in $]0,1[$ such that $1 - \beta_n \sim n^{-\frac{r+1}{p}}$, we put

$$D_n^-(x) = \left[\frac{k_n}{nf(x)}\right]^{1/p} \beta_n^{1/2p}, \qquad D_n^+(x) = \left[\frac{k_n}{nf(x)}\right]^{1/p} \beta_n^{-1/2p}.$$

Then, for n large enough we have almost surely: $D_n^-(x) \leqslant R_n(x) \leqslant D_n^+(x)$ (see, e.g., [8]). According to Assumption 4(v) we have

$$K\left(\frac{X_i - x}{R_n(x)}\right) = K\left(\frac{D_n^-(x)}{R_n(x)}\frac{X_i - x}{D_n^-(x)}\right) \geqslant K\left(\frac{X_i - x}{D_n^-(x)}\right)$$
$$\text{and} \qquad K\left(\frac{X_i - x}{D_n^+(x)}\right) = K\left(\frac{R_n(x)}{D_n^+(x)}\frac{X_i - x}{R_n(x)}\right) \geqslant K\left(\frac{X_i - x}{R_n(x)}\right).$$

Thus

$$K\left(\frac{X_i - x}{D_n^-(x)}\right) \leqslant K\left(\frac{X_i - x}{R_n(x)}\right) \leqslant K\left(\frac{X_i - x}{D_n^+(x)}\right) \tag{11}$$

and, therefore, $\widehat{f}_{1,n}(x) \leqslant \widehat{f}_n(x) \leqslant \widehat{f}_{2,n}(x)$, where

$$\widehat{f}_{1,n}(x) = \frac{1}{n\left(D_n^+(x)\right)^p}\sum_{i=1}^n K\left(\frac{X_i - x}{D_n^-(x)}\right) \qquad \text{and} \qquad \widehat{f}_{2,n}(x) = \frac{1}{n\left(D_n^-(x)\right)^p}\sum_{i=1}^n K\left(\frac{X_i - x}{D_n^+(x)}\right).$$

Hence

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_n(x) - f(x)\right| \leqslant \max\left\{\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_{1,n}(x) - f(x)\right|, \sup_{x\in\mathbb{R}^p}\left|\widehat{f}_{2,n}(x) - f(x)\right|\right\},$$

and it remains to prove that

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_{1,n}(x) - f(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)}{k_n^2}}\right) \tag{12}$$

and

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_{2,n}(x) - f(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)}{k_n^2}}\right). \tag{13}$$

For proving (12) we apply Lemma 13 with $h \equiv 1$, $\eta_n \equiv 1$, $\mathscr{D}_{1,n} = D_n^+(x)$ and $\mathscr{D}_{2,n} = D_n^-(x)$. In this case, the properties (5) and (6) are satisfied. Indeed, since $\beta_n \to 1$ as $n \to +\infty$, we have for n large enough $1/2 \leqslant \beta_n \leqslant 3/2$ and, therefore,

$$\left(D_n^+(x)\right)^p \geqslant \sqrt{\frac{2}{3}}\frac{1}{\|f\|_\infty}n^{-1}k_n, \qquad \left(D_n^-(x)\right)^p \leqslant \sqrt{\frac{3}{2}}\frac{1}{c_0}n^{-1}k_n, \tag{14}$$

and also

$$\left|\frac{\left(D_n^-(x)\right)^p}{\left(D_n^+(x)\right)^p} - 1\right| = 1 - \frac{\left(D_n^-(x)\right)^p}{\left(D_n^+(x)\right)^p} = 1 - \beta_n \sim n^{-\frac{r+1}{p}}. \tag{15}$$

On the other hand, $\widehat{\varphi}_n(x) = \widehat{f}_{1,n}(x)$ and

$$\varphi(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)\,\mathrm{d}y = f(x).$$

Then, applying Lemma 13 yields (12). Similarly, applying Lemma 13 to the case where $h \equiv 1$, $\eta_n \equiv 1$, $\mathscr{D}_{1,n} = D_n^-(x)$ and $\mathscr{D}_{2,n} = D_n^+(x)$ leads to (13) since

$$\left(D_n^-(x)\right)^p \geqslant \frac{1}{\sqrt{2}\|f\|_\infty}n^{-1}k_n, \qquad \left(D_n^+(x)\right)^p \leqslant \frac{\sqrt{2}}{c_0}n^{-1}k_n, \tag{16}$$

and

$$\left|\frac{\left(D_n^+(x)\right)^p}{\left(D_n^-(x)\right)^p} - 1\right| = \frac{\left(D_n^+(x)\right)^p}{\left(D_n^-(x)\right)^p} - 1 = \frac{1-\beta_n}{\beta_n} \sim n^{-\frac{r+1}{p}}. \tag{17}$$

4.3. *Proof of Theorem 10*

Clearly, $\widehat{g}_n(x) = \widehat{g}_{1,n}(x) - \widehat{g}_{2,n}(x)$, where

$$\widehat{g}_{1,n}(x) = \frac{1}{n\big(R_n(x)\big)^p}\sum_{i=1}^n Y_i \mathbb{1}_{\{Y_i \geqslant 0\}} K\left(\frac{X_i - x}{R_n(x)}\right)$$

and

$$\widehat{g}_{2,n}(x) = \frac{1}{n\big(R_n(x)\big)^p}\sum_{i=1}^n (-Y_i) \mathbb{1}_{\{Y_i < 0\}} K\left(\frac{X_i - x}{R_n(x)}\right).$$

From (11) we get $\widehat{g}^-_{1,n}(x) \leqslant \widehat{g}_{1,n}(x) \leqslant \widehat{g}^+_{1,n}(x)$ and $\widehat{g}^+_{2,n}(x) \leqslant \widehat{g}_{2,n}(x) \leqslant \widehat{g}^-_{2,n}(x)$, where

$$\widehat{g}^-_{1,n}(x) = \frac{1}{n\big(D^+_n(x)\big)^p}\sum_{i=1}^n Y_i \mathbb{1}_{\{Y_i \geqslant 0\}} K\left(\frac{X_i - x}{D^-_n(x)}\right),$$

$$\widehat{g}^+_{1,n}(x) = \frac{1}{n\big(D^-_n(x)\big)^p}\sum_{i=1}^n Y_i \mathbb{1}_{\{Y_i \geqslant 0\}} K\left(\frac{X_i - x}{D^+_n(x)}\right),$$

$$\widehat{g}^-_{2,n}(x) = \frac{1}{n\big(D^-_n(x)\big)^p}\sum_{i=1}^n (-Y_i) \mathbb{1}_{\{Y_i < 0\}} K\left(\frac{X_i - x}{D^+_n(x)}\right)$$

and

$$\widehat{g}^+_{2,n}(x) = \frac{1}{n\big(D^+_n(x)\big)^p}\sum_{i=1}^n (-Y_i) \mathbb{1}_{\{Y_i < 0\}} K\left(\frac{X_i - x}{D^-_n(x)}\right).$$

Then, since $\widehat{g}_n(x) - g(x) = \big(\widehat{g}_{1,n}(x) - g_1(x)\big) - \big(\widehat{g}_{2,n}(x) - g_2(x)\big)$, it follows

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{g}_n(x) - g(x)\right| \leqslant 2\max\left\{\sup_{x\in\mathbb{R}^p}\left|\widehat{g}^+_{1,n}(x) - g_1(x)\right| + \sup_{x\in\mathbb{R}^p}\left|\widehat{g}^-_{2,n}(x) - g_2(x)\right|,\right.$$
$$\left.\sup_{x\in\mathbb{R}^p}\left|\widehat{g}^-_{1,n}(x) - g_1(x)\right| + \sup_{x\in\mathbb{R}^p}\left|\widehat{g}^+_{2,n}(x) - g_2(x)\right|\right\},$$

and it suffices to show that

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{g}^+_{\ell,n}(x) - g_\ell(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)M_n^2}{k_n^2}}\right) \tag{18}$$

and

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{g}^-_{\ell,n}(x) - g_\ell(x)\right| = O_{a.s.}\left(\left(\frac{k_n}{n}\right)^{\frac{r+1}{p}} + \sqrt{\frac{n\log(n)M_n^2}{k_n^2}}\right) \tag{19}$$

for $\ell \in \{1,2\}$. For proving (18) with $\ell = 1$, we apply Lemma 13 with $h(y) = y\mathbb{1}_{\mathbb{R}_+}(y)$, $\eta_n = M_n$, $\mathscr{D}_{1,n} = D^-_n(x)$ and $\mathscr{D}_{2,n} = D^+_n(x)$. In this case, the properties (5) and (6) are satisfied in (16) and (17) respectively, and we have $\widehat{\varphi}_n(x) = \widehat{g}^+_{1,n}(x)$ and

$$\varphi(x) = \int_{\mathbb{R}} y\mathbb{1}_{\mathbb{R}_+}(y) f_{(X,Y)}(x,y)\,\mathrm{d}y = f(x)\int_{\mathbb{R}} y\mathbb{1}_{\mathbb{R}_+}(y) f_{Y\,|\,X=x}(y)\,\mathrm{d}y = g_1(x).$$

Similarly, applying Lemma 13 to the case where $h(y) = -y\mathbb{1}_{]-\infty,0[}(y)$, $\eta_n = M_n$, $\mathscr{D}_{1,n} = D^+_n(x)$ and $\mathscr{D}_{2,n} = D^-_n(x)$ leads to (18) with $\ell = 2$ since the properties (5) and (6) are satisfied in (14) and (15) respectively, and we have $\widehat{\varphi}_n(x) = \widehat{g}^+_{2,n}(x)$ and

$$\varphi(x) = -\int_{\mathbb{R}} y\mathbb{1}_{]-\infty,0[}(y) f_{(X,Y)}(x,y)\,\mathrm{d}y = -f(x)\int_{\mathbb{R}} y\mathbb{1}_{]-\infty,0[}(y) f_{Y\,|\,X=x}(y)\,\mathrm{d}y = g_2(x).$$

Equation (19) is obtained from a similar reasoning.

4.4. *Proof of Theorem 11*

Clearly,

$$\begin{aligned}\left|\widehat{r}_n(x)-r(x)\right| &= \left|\frac{\widehat{g}_n(x)}{\widehat{f}_{b_n}(x)}-\frac{g(x)}{f(x)}\right| \\ &= \frac{\left|\widehat{g}_n(x)f(x)-\widehat{f}_{b_n}(x)g(x)\right|}{\widehat{f}_{b_n}(x)f(x)} \\ &\leqslant c_0^{-1}\frac{\left|\left(\widehat{g}_n(x)-g(x)\right)f(x)+g(x)\left(f(x)-\widehat{f}_{b_n}(x)\right)\right|}{\widehat{f}_{b_n}(x)} \\ &\leqslant c_0^{-1}\frac{\|f\|_\infty\left|\widehat{g}_n(x)-g(x)\right|+\|g\|_\infty\left|\widehat{f}_{b_n}(x)-f(x)\right|}{\widehat{f}_{b_n}(x)}.\end{aligned}$$

Since $\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_n(x)-f(x)\right|\to 0$, a.s., as $n\to+\infty$, we have for n large enough, $\left|\widehat{f}_n(x)-f(x)\right|\leqslant c_0/2$ and, therefore,

$$c_0\leqslant f(x)\leqslant\left|\widehat{f}_n(x)-f(x)\right|+\widehat{f}_n(x)\leqslant\frac{c_0}{2}+\widehat{f}_n(x).$$

Hence, $\widehat{f}_n(x)\geqslant c_0/2$ and, since $\widehat{f}_{b_n}(x)\geqslant\widehat{f}_n(x)$ it follows that $\widehat{f}_{b_n}(x)\geqslant c_0/2$. Thus

$$\left|\widehat{r}_n(x)-r(x)\right|\leqslant 2c_0^{-2}\left(\|f\|_\infty\left|\widehat{g}_n(x)-g(x)\right|+\|g\|_\infty\left|\widehat{f}_{b_n}(x)-f(x)\right|\right).$$

On the other hand, since $\widehat{f}_n(x)\leqslant\widehat{f}_{b_n}(x)\leqslant\widehat{f}_n(x)+b_n$, it follows that $\left|\widehat{f}_{b_n}(x)-\widehat{f}_n(x)\right|\leqslant b_n$ and, therefore,

$$\left|\widehat{f}_{b_n}(x)-f(x)\right|\leqslant\left|\widehat{f}_{b_n}(x)-\widehat{f}_n(x)\right|+\left|\widehat{f}_n(x)-f(x)\right|\leqslant b_n+\left|\widehat{f}_n(x)-f(x)\right|.$$

Consequently,

$$\sup_{x\in\mathbb{R}^p}\left|\widehat{r}_n(x)-r(x)\right|\leqslant 2c_0^{-2}\left(\|f\|_\infty\sup_{x\in\mathbb{R}^p}\left|\widehat{g}_n(x)-g(x)\right|+\|g\|_\infty b_n+\|g\|_\infty\sup_{x\in\mathbb{R}^p}\left|\widehat{f}_n(x)-f(x)\right|\right),$$

and the proof is completed by using Theorems 8 and 10.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] I. S. Abramson, "On bandwidth variation in kernel estimates. A square root law", *Ann. Stat.* **10** (1982), pp. 1217–1223.

[2] M.-S. Ahmed, M. N'diaye, M. K. Attouch and S. Dabo-Niang, "k-nearest neighbors prediction and classifcation for spatial data", *J. Spatial Econom.* **4** (2023), article no. 12.

[3] M. K. Attouch, A. Laksaci and R. Rafaa, "Estimation locale linéaire de la régression non paramétrique fonctionnelle par la méthode des k plus proches voisins", *C. R. Math.* **355** (2017), pp. 824–829.

[4] P. K. Bhattacharya and Y. P. Mack, "Multivariate data-driven k-NN function estimation", *J. Multivariate Anal.* **35** (1990), no. 1, pp. 1–11.

[5] G. Biau and L. Devroye, *Lectures on the nearest neighbor method*, Springer Series in the Data Sciences, Springer, 2015, ix+290 pages.

[6] F. Burba, F. Ferraty and P. Vieu, "Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique", *C. R. Math.* **346** (2008), pp. 339–342.

[7] K. L. Chung, "An estimate concerning the Kolmogoroff limit distribution", *Trans. Am. Math. Soc.* **67** (1949), pp. 36–50.

[8] G. Collomb, "Estimation de la régression par la méthode des k points les plus proches avec noyau: quelques propriétés de convergence ponctuelle", in *Nonparametric asymptotic statistics (Proc. Conf., Rouen, 1979) (French)*, Lecture Notes in Mathematics, Springer, 1980, pp. 159–175.

[9] E. Giné and A. Guillou, "On consistency of kernel density estimators for randomly censored data: rates holding uniformly over adaptive intervals", *Ann. Inst. Henri Poincaré* **37** (2001), pp. 503–522.

[10] M. C. Jones, "Variable kernel density estimates and variable kernel density estimates", *Aust. J. Stat.* **32** (1990), no. 3, pp. 361–371.

[11] N. L. Kudraszow and P. Vieu, "Uniform consistency of k-NN regressors for functional variables", *Stat. Probab. Lett.* **83** (2013), pp. 1863–1870.

[12] H. Lian, "Convergence of functional k-nearest neighbor regression estimate with functional responses", *Electron. J. Stat.* **5** (2011), pp. 31–40.

[13] Z. Lu and P. Cheng, "Strong consistency of nearest neighbor kernel regression estimation for stationary dependent samples", *Sci. China, Ser. A* **41** (1998), pp. 918–926.

[14] Y. P. Mack and M. Rosenblatt, "Multivariate k-nearest neighbor density estimates", *J. Multivariate Anal.* **9** (1979), pp. 1–15.

[15] D. S. Moore and J. W. Yackel, "Consistency properties of nearest neighbor density function estimators", *Ann. Stat.* **5** (1977), pp. 143–154.

[16] E. d. D. Nkou, "Strong consistency of kernel method for sliced average variance estimation", *Commun. Stat., Theory Methods* **52** (2023), pp. 7586–7600.

[17] E. d. D. Nkou and G. M. Nkiet, "Strong consistency of kernel estimator in a semiparametric regression model", *Statistics* **53** (2019), pp. 1289–1305.

[18] M. Talagrand, "Sharper bounds for Gaussian and empirical processes", *Ann. Probab.* **22** (1994), pp. 28–76.

[19] G. R. Terrel and D. W. Scott, "Variable kernel density estimation", *Ann. Stat.* **20** (1992), pp. 1236–1265.

[20] D.-X. Zhang, "Rates of strong uniform convergence of nearest neighbor density estimates on any compact set", *Appl. Math. Mech., Engl. Ed.* **11** (1990), no. 4, pp. 385–393.

[21] L. Zhu and L. X. Zhu, "On kernel method for sliced average variance estimation", *J. Multivariate Anal.* **98** (2007), pp. 970–991.

[22] L. X. Zhu and K. T. Fang, "Asymptotics for kernel estimate of sliced inverse regression", *Ann. Stat.* **24** (1996), pp. 1053–1068.