Transactions of A. Razmadze
Mathematical Institute
Vol. **173** (2019), issue 3, 93–102

CONSTRUCTION OF A KERNEL DENSITY ESTIMATOR OF ROSENBLATT-PARZEN TYPE BY CONDITIONALLY INDEPENDENT OBSERVATIONS

ZURAB KVATADZE AND BEKNU PARJIANI

Abstract. On the probabilistic space (Ω, F, P), we consider the conditionally independent sequence $\{X_i\}_{i\geq 1}$ controlled by the sequence $\{\xi_i\}_{i\geq 1}$. The members of $\{\xi_i\}_{i\geq 1}$ are independent, identically distributed random variables $\xi_i = b_1 I_{(\xi_1=b_1)} + b_2 I_{(\xi_1=b_2)} + \cdots + b_r I_{(\xi_1=b_r)}$. The elements of the sequence $\{X_i\}_{i\geq 1}$ are the observations of some random variable X. Conditional distributions $\mathcal{P}_{X_i|\xi_i=b_i}$, $i=\overline{1,r}$, have unknown densities $f_i(x)$, $i=\overline{1,r}$, respectively. Using observations $\{X_i\}_{i\geq 1}$, a kernel density estimator $\bar{f}(x) = \sum\limits_{i=1}^{r} p_i f_i(x)$ of Rosenblatt-Parzen type is constructed, where $p_i = P(\xi_1 = b_i)$. The accuracy of approximation of the constructed estimator to the unknown function $\bar{f}(x)$ is established.

Distribution density estimators are intensively studied by many authors. In this paper, a nonparametric density estimator is constructed by dependent observations. The class of conditionally independent observations is considered. The nonparametric density estimators which have so far been considered are constructed by independent samples.

Below we present some definitions and auxiliary facts for nonparametric estimates of a distribution density which were constructed by independent observations.

Let the values X_i, $x_i \in R$, i=1,2,..., be independent observations of some random value X_i with unknown density $g(x)$. Various methods are available for obtaining estimators of $g(x)$. In the works of M. Rosenblatt and E. Parzen (see [8,9]) the estimators of$g_n^*(x)$ obtained by the kernel $k(x)$

$$g_n^*(x, h_n) = \frac{1}{nh_n}\sum_{i=1}^{n} k\left(\frac{(x-X_i)}{h_n}\right)$$

were considered, where $\{h_n\}_{n\geq 1}$ is the sequence of positive numbers such that

$$\lim_{n\to\infty} h_n = 0; \quad \lim_{n\to\infty} nh_n = \infty,$$

and the kernel $k(x)$ is some Lebesgue-integrable Borel function. In [5, 7, 10], the results of [9] were generalized by modifying the conditions on $k(x)$ and using the observations of vectors $X_i \in R^m$ $(m > 1)$.

Along with estimators of Rosenblatt-Parzen type, projection type estimators were also considered (see [2,7]) using the spectral decomposition of the kernel $k(x)$ with respect to the orthonormal basis of functions. Applying smoothing functions, L. Devroye and L. Györfi (see [3]) constructed the adaptive kernel estimators for densities with a finite number of discontinuity points. As a divergence measure of the constructed estimators of $g(x)$ some authors considered various characteristics in terms of metrics L_1, ([3,6]); L_2 ([7,9]) and so on.

In [7], E. Nadaraya obtained the sufficient conditions for the uniform convergence of the estimator

$$\hat{g}_n(x, a_n) = \frac{a_n}{n}\sum_{i=1}^{n} k(a_n(x-X_i))$$

2010 *Mathematics Subject Classification.* 62G05, 62G07.

Key words and phrases. Conditionally independent sequence; Kernel estimate; Conditional density.

to $g(x)$ with probability 1. The divergence measure between $g(x)$ and $\hat{g}_n(x, a_n)$ is the value

$$E\int\limits_{-\infty}^{\infty}\left[\hat{g}_n(x, a_n) - g(x)\right]^2 dx,$$

where $\{a_n\}_{n\geq 1}$ is the sequence of positive numbers such that

$$\lim_{n\to\infty} a_n = \infty, \quad a_n = o(n). \tag{1}$$

Definition 1. Denote by H_s ($s \geq 2$; s is an even number) a set of functions $k(x)$ satisfying the conditions

$$k(-x) = k(x), \quad \int\limits_{-\infty}^{\infty} k(x)\,dx = 1, \quad \sup|k(x)| \leq A < \infty,$$

$$\int\limits_{-\infty}^{\infty} x^i k(x)\,dx = 0, \quad i = 1, 2, \dots, s-1; \quad \int\limits_{-\infty}^{\infty} x^s k(x)\,dx \neq 0, \quad \int\limits_{-\infty}^{\infty} x^s |k(x)|\,dx < \infty. \tag{2}$$

Definition 2. Denote by W_s a set of functions $\varphi(x)$ having derivatives up to the s-th order ($s \geq 2$) inclusive, and note that $\varphi^{(s)}(x)$ is a continuous bounded function from the class $L_2(-\infty, \infty)$.

Lemma (see [7])**.** *If the variables X_i, $x_i \in R$, $i = 1, 2, \dots$, are independent observations of some random variable X with unknown density $g(x)$, $g(x) \in W_s \cap L_2(-\infty, \infty)$, $k(x) \in H_s \cap L_2(-\infty, \infty)$ and*

$$\hat{g}_n(x, a_n) = \frac{a_n}{n}\sum_{i=1}^{n} k(a_n(x - X_i)),$$

then for $n \to \infty$, the equalities

$$\int\limits_{-\infty}^{\infty} D\hat{g}_n(x, a_n)\,dx = \frac{a_n}{n}\int\limits_{-\infty}^{\infty} k^2(x)\,dx + O\left(\frac{a_n}{n}\right), \tag{3}$$

$$\int\limits_{-\infty}^{\infty}\left[E\hat{g}_n(x, a_n) - g(x)\right]^2 dx = a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int\limits_{-\infty}^{\infty}\left[g^{(s)}(x)\right]^2 dx + O\left(a_n^{-2s}\right) \tag{4}$$

hold, where $\{a_n\}_{n\geq 1}$ is the sequence (1)*, and*

$$\alpha = \int\limits_{-\infty}^{\infty} x^s k(x)\,dx.$$

Let us present our result.

In practice we encounter the situation when at random moments of time the distribution of the observed variable X changes depending on the conditions (of the controlling sequence $\{\xi_i\}_{i\geq 1}$). This brings about changes of the densities of observations $\{X_i\}_{i\geq 1}$. For example, in stock-exchange transactions the price of some commodities changes depending on a season, though this price is fixed at the auction. As a result, the flow of revenues due to such transactions also changes, and so on.

In such situations, to estimate the density X it is appropriate to consider dependent observations. Here we consider the class of conditionally independent observations.

On the probability space (Ω, F, P), let us consider the two-component stationary (in the narrow sense) sequence of random variables

$$\{\xi_i, X_i\}_{i\geq 1}, \tag{5}$$

where $\xi_i : \Omega \to \Xi$, $X_i : \Omega \to R^m$ and Ξ is some space.

Definition 3. The sequence $\{X_i\}_{i\geq 1}$ from (4) is called a conditionally independent sequence (see [1]) controlled by the sequence $\{\xi_i\}_{i\geq 1}$ if for any natural n and the fixed trajectory $\overline{\xi}_{1n} = (\xi_1, \xi_2, \dots, \xi_n)$, the values $X_1, X_2, \dots, X_n$ become independent and for all natural numbers i, k, n, $j_1, j_2, \dots, j_k$, $(2 \leq k \leq n;\ i \leq n;\ 1 \leq j_1 < j_2 <, \cdots, < j_k \leq n)$ the equalities

$$\mathcal{P}_{(X_{j_1}, X_{j_2}, \dots, X_{j_k})|\xi_{1n}} = \mathcal{P}_{X_{j_1}|\xi_{j_1}} \times \mathcal{P}_{X_{j_2}|\xi_{j_2}} \times \dots \times \mathcal{P}_{X_{j_k}|\xi_{j_k}},$$
$$\mathcal{P}_{X_i|\xi_{1n}} = \mathcal{P}_{X_i|\xi_i},$$

are fulfilled, where $\mathcal{P}_{X|Y}$ is the conditional distribution of the variable X under the condition Y.

Consider the sequence (5). Let ξ_i, $i = 1, 2, \dots,$ be independent, identically distributed random variables and let

$$\Xi = \{b_1, b_2, \dots, b_r\}; \quad P(\xi_1 = b_i) = p_i, \quad i = \overline{1, r}, \quad p_1 + p_2 + \dots + p_r = 1,$$

$\{X_i\}_{i\geq 1}$ is the conditionally independent sequence whose elements are observations of the variable X. It is assumed that the conditional distributions $\mathcal{P}_{X_1|\xi_1 = b_i}$, $i = \overline{1, r}$, have unknown densities $f_i(x)$, $i = \overline{1, r}$, respectively. The sum $\hat{f}_n(x, a_n) = \frac{a_n}{n} \sum\limits_{j=1}^{n} k(a_n(x - X_j))$ constructed by conditionally independent observations is considered as the density estimator $\bar{f}(x) = \sum\limits_{i=1}^{r} p_i f_i(x)$, while the estimator accuracy is established by the expression $u(a_n) = E \int\limits_{-\infty}^{\infty} \left[\hat{f}_n(x, a_n) - \bar{f}(x)\right]^2 dx$.

On the fixed trajectory $\overline{\xi}_{1n} = (\xi_1, \xi_2, \dots, \xi_n)$ of the sequence $\{\xi_i\}_{i\geq 1}$, we denote by $\nu_n(1), \nu_n(2), \dots,$ $\nu_n(r)$ the frequencies for which the first n members of the sequence adopt the values b_1, b_2, …, b_r.

Theorem. *Let us consider the sequence* (5). *The elements of the controlling sequence* $\{\xi_i\}_{i\geq 1}(\xi_i : \Omega \to \{b_1, b_2, \dots, b_r\})$ *are independent, identically distributed values* $\xi_i = \sum\limits_{i=1}^{r} b_i I_{(\xi_1 = b_i)}$. *Assume that for every function* $\Psi : \Xi \to R^1$, *for which* $E\Psi(\xi_i) < \infty$ *as* $n \to \infty$, *we have the convergence*

$$\frac{1}{n} \sum_{j=1}^{n} \Psi(\xi_i) \to E\Psi(\xi_1) \quad a.s. \tag{6}$$

The elements of the conditionally independent sequence $\{X_i\}_{i\geq 1}$ are observations of the variable X. The conditional distributions $\mathcal{P}_{X_1|\xi_1 = b_i}$, $i = \overline{1, r}$, have unknown densities $f_i(x)$, $i = \overline{1, r}$, respectively. Assume $f_i(x) \in W_s \cap L_2(-\infty, \infty)$ and $k(x) \in H_s \cap L_2(-\infty, \infty)$. If for the frequencies $\nu_n(i)$, $i = \overline{1, r}$, the inequalities

$$D\left(\frac{\nu_n(i)}{n}\right) \leq \frac{c_i}{\sqrt{n}}, \quad i = \overline{1, r} \tag{7}$$

are fulfilled, then for any natural n the estimator of the density $\bar{f}(x) = \sum\limits_{i=1}^{r} p_i f_i(x)$ is the sum

$$\hat{f}_n(x, a_n) = \frac{a_n}{n} \sum_{j=1}^{n} k(a_n(x - X_j)) \tag{8}$$

and the following asymptotic equality

$$u(a_n) \leq \left(\sum_{i=1}^{r} M_i\right)^2 + \frac{a_n}{n} \int\limits_{-\infty}^{\infty} k^2(x)\, dx + \left(\sum_{i=1}^{r} (C_i n^{-1/2} + {p_i}^2)\right) O\Big(\frac{a_n}{n}\Big)$$

is valid, where

$$M_i = T_i^{1/2} + \left(C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2\left(x\right) dx \right)^{\frac{1}{2}}$$

$$T_i = (a_n^{-2s} \frac{\alpha^2}{(s!)^2} \int_{-\infty}^{\infty} [f_i^{(s)}\left(x\right)]^2 dx + O(a_n^{-2s}))(C_i n^{-1/2} + p_i^2) \quad i = \overline{1,r}$$

and

$$\alpha = \int_{-\infty}^{\infty} x^s k\left(x\right) dx.$$

Proof. The proof of the theorem is based on the decomposition of $\hat{f}_n\left(x, a_n\right)$ (the trajectory $\overline{\xi}_{1n}$is assumed to be fixed) into independent sums of random variables. For the fixed trajectory $\overline{\xi}_{1n}$ we enumerate individually the moments of time at which the first n members of the sequence $\{\xi_i\}_{i\geq 1}$ take the value b_i, $i = \overline{1,r}$, respectively,

$$\tau_0(i) = 0, \quad \tau_m(i) = \min\{j | \tau_{m-1} < j \leq n; \xi_i = b_i\}; \quad m = \overline{1, \nu_n\left(i\right)}, \quad i = \overline{1,r}.$$

We obtain the sequence of indices

$$\tau_1(i), \tau_2(i), \ldots, \tau_{\nu_n(i)}(i) \quad i = \overline{1,r}$$

for which the equalities

$$\xi_{\tau_m(i)} = b_i \quad m = \overline{1, \nu_n\left(i\right)}, \quad i = \overline{1,r}$$

are valid.

When the trajectory $\overline{\xi}_{1n}$ is fixed, the sum (8) can be decomposed as follows

$$\hat{f}_n\left(x, a_n\right) = \sum_{i=1}^{r} \frac{\nu_n\left(i\right)}{n} \hat{f}_{in}\left(x, a_n\right),$$

where

$$\hat{f}_{in}\left(x, a_n\right) = \frac{a_n}{\nu_n\left(i\right)} \sum_{m=1}^{\nu_n(i)} k\left(a_n\left(x - X_{\tau_{m(i)}}\right)\right) \quad i = \overline{1,r}.$$

Naturally, if $\nu_n\left(i\right) = 0$, then the summand $\hat{f}_{in}\left(x, a_n\right)$, $i = \overline{1,r}$, does not exist. Let us prove the finiteness of $E\hat{f}_n\left(x, a_n\right)$ and $D\hat{f}_n\left(x, a_n\right)$. On the fixed trajectory $\overline{\xi}_{1n}$, we represent $E\hat{f}_n\left(x, a_n\right)$ as a conditional mathematical expectation

$$E\hat{f}_n\left(x, a_n\right) = E\{E(\hat{f}_n\left(x, a_n\right) | \xi_{1n})\} = E\left\{E\left(\sum_{i=1}^{r} \frac{\nu_n\left(i\right)}{n} \hat{f}_{in}\left(x, a_n\right) | \xi_{1n}\right)\right\}.$$

In proving the theorem, we take into account that $\nu_n\left(i\right)$, $i = \overline{1,r}$, are measurable functions with respect to the σ-algebra which is generated when the probability space Ω is partitioned as a result of fixing the trajectory $\overline{\xi}_{1n}$. Therefore these functions can be taken outside the sign of mathematical expectation. In the above equality and in the sequel we keep in mind the fact that the following equality

$$E\frac{\nu_n\left(i\right)}{n} = p_i$$

is fulfilled by virtue of conditions (6), and applying condition (7), we obtain the estimator

$$E\left(\frac{\nu_n\left(i\right)}{n}\right)^2 = D\left(\frac{\nu_n\left(i\right)}{n}\right) + \left(E\frac{\nu_n\left(i\right)}{n}\right)^2 \leq n^{-1/2} c_i + p_i^2. \tag{9}$$

Hence, using conditions (2), after replacing the variable under the integration sign, we see that the following chain of equalities

$$E\hat{f}_n(x,a_n)=\sum_{i=1}^{r}E\Big\{\frac{\nu_n(i)}{n}E\Big(\frac{a_n}{\nu_n(i)}\sum_{m=1}^{\nu_n(i)}k\left(a_n\left(x-X_{\tau_{m(i)}}\right)\right)|\xi_{1n}\Big)\Big\}$$
$$=\sum_{i=1}^{r}E\Big\{\frac{\nu_n(i)}{n}E\Big(\frac{a_n}{\nu_n(i)}\nu_n(i)\,k\left(a_n\left(x-X_{\tau_{m(i)}}\right)\right)|\xi_{\tau_{m(i)}}\Big)\Big\}$$
$$=\sum_{i=1}^{r}a_n\int_{-\infty}^{\infty}k(a_n(x-u))f_i(u)du E\frac{\nu_n(i)}{n}=\sum_{i=1}^{r}p_i\int_{-\infty}^{\infty}k(t)f_i\Big(\frac{t}{a_n}+x\Big)dt$$

is valid.

Since$f_i(x)$ is the density and $|k(t)|$ is bounded by the infinite constant A, we conclude that $E\hat{f}_n(x,a_n)$ is finite.

On the fixed trajectory $\overline{\xi}_{1n}$, the sums $\hat{f}_{in}(x,a_n)$, $i=\overline{1,r}$, and their constituent summands too, are independent and the following equalities

$$D\hat{f}_n(x,a_n)=E\{E([\hat{f}_n(x,a_n)-E\hat{f}_n(x,a_n)]^2|\xi_{1n})\}$$
$$=E\Big\{E\Big(\Big[\sum_{i=1}^{r}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n)-E\sum_{i=1}^{r}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n)\Big]^2|\xi_{1n}\Big)\Big\}$$
$$=E\Big\{E\Big(\Big[\sum_{i=1}^{r}\frac{\nu_n(i)}{n}(\hat{f}_{in}(x,a_n)-E\hat{f}_{in}(x,a_n))\Big]^2|\xi_{1n}\Big)\Big\}$$
$$=E\Big\{E\Big(\Big[\sum_{i=1}^{r}\Big(\frac{\nu_n(i)}{n}\Big)^2\hat{f}_{in}(x,a_n)-E\hat{f}_{in}(x,a_n)\Big]^2|\xi_{1n}\Big)\Big\}$$
$$=\sum_{i=1}^{r}\Big\{E\Big(\frac{\nu_n(i)}{n}\Big)^2E([\hat{f}_{in}(x,a_n)-E\hat{f}_{in}(x,a_n)]^2|\xi_{1n})\Big\}$$
$$=\sum_{i=1}^{r}E\Big\{\Big(\frac{\nu_n(i)}{n}\Big)^2E\Big[\sum_{j=1}^{\nu_n(i)}\frac{a_n}{\nu_n(i)}\left(k\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)-Ek\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)\right)\Big]^2|\xi_{1n}\Big\}$$
$$=\sum_{i=1}^{r}E\Big\{\Big(\frac{\nu_n(i)}{n}\Big)^2\Big(\frac{a_n}{\nu_n(i)}\Big)^2E\Big(\sum_{j=1}^{\nu_n(i)}[k\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)-Ek\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)]^2|\xi_{1n}\Big)\Big\}$$
$$=\sum_{i=1}^{r}E\Big(\frac{\nu_n(i)}{n}\Big)^2\Big(\frac{a_n}{\nu_n(i)}\Big)^2\nu_n(i)\int_{-\infty}^{+\infty}\Big[k(a_n(x-u))-\int_{-\infty}^{+\infty}k(a_n(x-y))f_i(y)dy\Big]^2f_i(u)\,du$$

are true.

Using equality (9), we obtain for $D\hat{f}_n(x,a_n)$ the following expression:

$$D\hat{f}_n(x,a_n)=\frac{a_n}{n}\sum_{i=1}^{2}p_i\int_{-\infty}^{\infty}[k(t)-\int_{-\infty}^{\infty}k(a_n(x-y))f_i(y)dy]^2f_i\Big(\frac{t}{a_n}+x\Big)dt.$$

Let us estimate $u(a_n)$. We apply Fubini's theorem and divide $u(a_n)$ into two parts

$$u(a_n)=\int_{-\infty}^{\infty}E\Big[\hat{f}_n(x,a_n)-\bar{f}(x)\Big]^2dx$$
$$=\int_{-\infty}^{\infty}E\Big[\hat{f}_n(x,a_n)-E\hat{f}_n(x,a_n)\Big]^2dx+\int_{-\infty}^{\infty}E\Big[E\hat{f}_n(x,a_n)-\overline{f}(x)\Big]^2dx=I_1+I_2. \qquad (10)$$

To estimate I_1, we again apply Fubini's theorem and obtain

$$I_1 = \int_{-\infty}^{\infty} E\Big[\hat{f}_n(x,a_n) - E\hat{f}_n(x,a_n)\Big]^2 dx = \int_{-\infty}^{\infty} E\{E([\hat{f}_n(x,a_n) - E\hat{f}_n(x,a_n)]^2 |\xi_{1n})\}dx$$

$$= E\int_{-\infty}^{+\infty} E\bigg(\bigg[\sum_{i=1}^{r}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n) - E\sum_{i=1}^{r}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n)\bigg]^2 |\xi_{1n}\bigg)dx$$

$$= E\bigg\{\int_{-\infty}^{+\infty} E\bigg(\sum_{i=1}^{r}\bigg(\frac{\nu_n(i)}{n}\bigg)^2 [\hat{f}_{in}(x,a_n) - E\hat{f}_{in}(x,a_n)]^2 |\xi_{1n}\bigg)dx\bigg\}$$

$$= E\bigg\{\int_{-\infty}^{+\infty}\sum_{i=1}^{r}\bigg(\frac{\nu_n(i)}{n}\bigg)^2 E([\hat{f}_{in}(x,a_n) - E\hat{f}_{in}(x,a_n)]^2 |\xi_{1n})dx\bigg\}.$$

Using equality (3) from the Lemma, we have

$$I_1 = \sum_{i=1}^{r} E\bigg\{\bigg(\frac{\nu_n(i)}{n}\bigg)^2\bigg(\frac{a_n}{\nu_n(i)}\int_{-\infty}^{+\infty} k(x)dx + o\Big(\frac{a_n}{n}\Big)\bigg)\bigg\}$$

$$= \sum_{i=1}^{r} E\bigg\{\bigg(\frac{\nu_n(i)}{n}\bigg)^2 \frac{a_n}{\nu_n(i)}\int_{-\infty}^{+\infty} k^2(x)dx\bigg\} + o\left(\frac{a_n}{n}\right)\sum_{i=1}^{r} E\bigg(\frac{\nu_n(i)}{n}\bigg)^2$$

$$= \frac{a_n}{n}\int_{-\infty}^{+\infty} k^2(x)dx \sum_{i=1}^{r} E\left(\frac{\nu_n(i)}{n}\right) + o\left(\frac{a_n}{n}\right)\sum_{i=1}^{r}\left[D\left(\frac{\nu_n(i)}{n}\right) + \left(E\frac{\nu_n(i)}{n}\right)^2\right].$$

By applying inequality (9), we complete the estimation of I_1,

$$I_1 \le \frac{a_n}{n}\int_{-\infty}^{+\infty} k^2(x)dx\sum_{i=1}^{r} p_i + o\left(\frac{a_n}{n}\right)\sum_{i=1}^{r}(C_i n^{-1/2} + {p_i}^2)$$

$$= \frac{a_n}{n}\int_{-\infty}^{+\infty} k^2(x)dx + o\left(\frac{a_n}{n}\right)\sum_{i=1}^{r}(C_i n^{-1/2} + {p_i}^2). \tag{11}$$

Applying Fubini's theorem once more and decomposing I_2 into two sums, we have

$$I_2 = \int_{-\infty}^{\infty} E\Big[E\hat{f}_n(x,a_n) - \overline{f}(x)\Big]^2 dx = E\int_{-\infty}^{\infty}\Big[E\hat{f}_n(x,a_n) - \overline{f}(x)\Big]^2 dx$$

$$= E\bigg\{E\bigg(\int_{-\infty}^{\infty}\Big[E\hat{f}_n(x,a_n) - \overline{f}(x)\Big]^2 dx|\xi_{1n}\bigg)\bigg\}$$

$$= E\bigg\{E\bigg(\int_{-\infty}^{\infty}\bigg[E\bigg(\sum_{i=1}^{r}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n)\bigg) - \sum_{i=1}^{r} p_i f_i(x)\bigg]^2 dx|\xi_{1n}\bigg)\bigg\}$$

$$= E\bigg\{E\bigg(\int_{-\infty}^{\infty}\bigg[\sum_{i=1}^{r}\Big(E\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n) - p_i f_i(x)\Big)\bigg]^2 dx|\xi_{1n}\bigg)\bigg\}$$

$$= E\bigg\{E\bigg(\int_{-\infty}^{\infty}\bigg(\sum_{i=1}^{r}\Big(E\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n) - p_i f_i(x)\Big)^2$$

$$+2\sum_{\substack{i,j=1\\i<j}}^{r}\Big(E\frac{\nu_n(i)}{n}\hat f_{in}(x,a_n)-p_if_i(x)\Big)\Big)\Big(E\frac{\nu_n(j)}{n}\hat f_{jn}(x,a_n)-p_jf_j(x)\Big)\Big)dx|\xi_{1n}\Big)\Big\}$$

$$=E\Big\{E\Big(\int_{-\infty}^{\infty}\sum_{i=1}^{r}E\frac{\nu_n(i)}{n}\hat f_{in}(x,a_n)-p_if_i(x))^2dx|\xi_{1n}\Big)\Big\}$$

$$+E\Big\{E\Big(\int_{-\infty}^{\infty}2\sum_{\substack{i,j=1\\i<j}}^{r}\Big(E\frac{\nu_n(i)}{n}\hat f_{in}(x,a_n)-p_if_i(x)\Big)\Big(E\frac{\nu_n(j)}{n}\hat f_{jn}(x,a_n)-p_jf_j(x)\Big)dx|\xi_{1n}\Big)\Big\}$$

$$=\sum_{i=1}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat f_{in}(x,a_n)-p_if_i(x)\Big]^2dx|\xi_{1n}\Big)\Big\}$$

$$+2\sum_{\substack{i,j=1\\i<j}}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat f_{in}(x,a_n)-p_if_i(x)\Big]\Big[\frac{\nu_n(j)}{n}E\hat f_{jn}(x,a_n)-p_jf_j(x)\Big]dx|\xi_{1n}\Big)\Big\}=I_{21}+I_{22}.$$

We decompose the sum I_{21} into three parts

$$I_{21}=\sum_{i=1}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat f_{in}(x,a_n)-\frac{\nu_n(i)}{n}f_i(x)+\frac{\nu_n(i)}{n}f_i(x)-p_if_i(x)\Big]^2dx|\xi_{1n}\Big)\Big\}$$

$$=\sum_{i=1}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\Big(\frac{\nu_n(i)}{n}\Big)^2[E\hat f_{in}(x,a_n)-f_i(x)]^2dx|\xi_{1n}\Big)\Big\}$$

$$+\sum_{i=1}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\Big(\frac{\nu_n(i)}{n}-p_i\Big)^2f_i^2(x)dx|\xi_{1n}\Big)\Big\}$$

$$+2\sum_{i=1}^{r}E\Big\{E\Big(\int_{-\infty}^{\infty}\frac{\nu_n(i)}{n}\Big(\frac{\nu_n(i)}{n}-p_i\Big)[E\hat f_{in}(x,a_n)-f_i(x)]f_i(x)dx|\xi_{1n}\Big)\Big\}=A_1+A_2+A_3.$$

Using equality (4) from the Lemma and the estimator (9), we obtain

$$A_1=\sum_{i=1}^{r}E\Big\{\Big(\frac{\nu_n(i)}{n}\Big)^2E\Big(\int_{-\infty}^{\infty}[E\hat f_{in}(x,a_n)-f_i(x)]^2dx|\xi_{1n}\Big)\Big\}$$

$$=\sum_{i=1}^{r}E\Big\{\Big(\frac{\nu_n(i)}{n}\Big)^2\Big(a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_i^{(s)}(x)]^2dx+o(a_n^{-2s})\Big)\Big\}$$

$$=\sum_{i=1}^{r}\Big(a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_i^{(s)}(x)]^2dx+o(a_n^{-2s})\Big)(C_in^{-1/2}+p_i^2)=\sum_{i=1}^{r}T_i,$$

where

$$T_i=\Big(a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_i^{(s)}(x)]^2dx+o(a_n^{-2s})\Big)(C_in^{-1/2}+p_i^2).$$

It is not difficult to derive the estimator for the sum A_2,

$$A_2=\sum_{i=1}^{r}E\Big\{\Big(\frac{\nu_n(i)}{n}-p_i\Big)^2E\int_{-\infty}^{\infty}f_i^2(x)dx\Big\}\le\sum_{i=1}^{r}C_in^{-1/2}\int_{-\infty}^{\infty}f_i^2(x)dx.$$

We use the fact that the values $\frac{\nu_n(i)}{n} - p_i$, $i=\overline{1,r}$, as well as $\frac{\nu_n(i)}{n}$, are measurable with respect to the σ-algebra generated by the fixed trajectory $\overline{\xi}_{1n}$. By Hölder's inequality, for the summand A_3, we obtain the following estimator

$$A_3 = 2\sum_{i=1}^{r} E\Big\{\frac{\nu_n(i)}{n}\Big(\frac{\nu_n(i)}{n} - p_i\Big)E\Big(\int_{-\infty}^{\infty}[E\hat f_{in}(x,a_n) - f_i(x)]f_i(x)dx|\xi_{1n}\Big)\Big\}$$

$$\le 2\sum_{i=1}^{r} E\left\{\frac{\nu_n(i)}{n}\Big(\frac{\nu_n(i)}{n} - p_i\Big)E\left(\sqrt{\int_{-\infty}^{\infty}[E\hat f_{in}(x,a_n) - f_i(x)]^2dx}\sqrt{\int_{-\infty}^{\infty}f_i{}^2(x)dx|\xi_{1n}}\right)\right\}$$

$$\le 2\sum_{i=1}^{r}\sqrt{a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}\big[f_i^{(s)}(x)\big]^2dx + o(a_n^{-2s})}\sqrt{\int_{-\infty}^{\infty}f_i{}^2(x)dx}\times E\Big\{\frac{\nu_n(i)}{n}\Big(\frac{\nu_n(i)}{n} - p_i\Big)\Big\}$$

$$\le 2\sum_{i=1}^{r}\sqrt{a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}\big[f_i^{(s)}(x)\big]^2dx + o(a_n^{-2s})}\sqrt{\int_{-\infty}^{\infty}f_i{}^2(x)dx}\times\sqrt{E\Big(\frac{\nu_n(i)}{n}\Big)^2}\sqrt{E\Big(\frac{\nu_n(i)}{n} - p_i\Big)^2}$$

$$\le 2\sum_{i=1}^{r}\sqrt{a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}\big[f_i^{(s)}(x)\big]^2dx + o(a_n^{-2s})}\sqrt{\int_{-\infty}^{\infty}f_i{}^2(x)dx}\cdot\sqrt{C_in^{-1/2}}\sqrt{C_in^{-1/2}+p_i^2}$$

$$\equiv 2\sum_{i=1}^{r}\sqrt{T_i}\cdot\sqrt{C_in^{-1/2}\int_{-\infty}^{\infty}f_i{}^2(x)dx}.$$

The summation of the estimators A_1 A_2 and A_3 gives

$$I_{21} \le \sum_{i=1}^{r}\Big(\sqrt{T_i} + \sqrt{C_in^{-1/2}\int_{-\infty}^{\infty}f_i{}^2(x)dx}\Big)^2 \equiv \sum_{i=1}^{r}M_i^2, \tag{12}$$

where

$$M_i = \sqrt{T_i} + \sqrt{C_in^{-1/2}\int_{-\infty}^{\infty}f_i{}^2(x)dx}.$$

Consider the sum I_{22}

$$I_{22} = 2\sum_{\substack{i,j=1\\ i<j}}^{r} E\Big\{E\Big(\int_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat f_{in}(x,a_n) - p_if_i(x)\Big]\Big[\frac{\nu_n(j)}{n}E\hat f_{jn}(x,a_n) - p_jf_j(x)\Big]dx|\xi_{1n}\Big)\Big\}$$

$$= 2\sum_{\substack{i,j=1\\ i<j}}^{r} B_{ij}.$$

The summands of this sum are estimated in the same manner as above. Let us estimate one of them by applying Fubini's and Hölder's theorems:

$$B_{ij} = E\Big\{E\Big(\int_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat f_{in}(x,a_n) - p_if_i(x)\Big]\Big[\frac{\nu_n(j)}{n}E\hat f_{jn}(x,a_n) - p_jf_j(x)\Big]dx|\xi_{1n}\Big)\Big\}$$

$$\leq E\left\{E\sqrt{\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat{f}_{in}(x,a_n)-p_if_i(x)\Big]^2dx}\sqrt{\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(j)}{n}E\hat{f}_{jn}(x,a_n)-p_jf_j(x)\Big]^2dx|\xi_{1n}}\right\}$$

$$\leq E\left\{\sqrt{E\Big(\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat{f}_{in}(x,a_n)-p_if_i(x)\Big]^2dx|\xi_{1n}\Big)}\right.$$

$$\left.\times\sqrt{E\Big(\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(j)}{n}E\hat{f}_{jn}(x,a_n)-p_jf_j(x)\Big]^2dx|\xi_{1n}\Big)}\right\}$$

$$\leq\sqrt{E\Big\{E\Big(\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(i)}{n}E\hat{f}_{in}(x,a_n)-p_if_i(x)\Big]^2dx|\xi_{1n}\Big)\Big\}}$$

$$\times\sqrt{E\Big\{E\Big(\int\limits_{-\infty}^{\infty}\Big[\frac{\nu_n(j)}{n}E\hat{f}_{jn}(x,a_n)-p_jf_j(x)\Big]^2dx|\xi_{1n}\Big)\Big\}}.$$

Each of the obtained two multipliers is estimated like a summand of the sum I_{21}. Thus we obtain the estimate B_{ij}

$$B_{ij}\leq M_iM_j.$$

Hence an estimator of the sum I_{22} has the form

$$I_{22}\leq 2\sum_{\substack{i,j=1\\ i<j}}^{r}M_iM_j. \tag{13}$$

Thus, in view of the decomposition (10) and the derived estimators (11), (12) and (13), the theorem is proved. □

Note that the proposed method enables one to construct density estimators for other types of dependence of observations too, for example, when the controlling sequence $\{\xi_i\}_{i\geq 1}$ is a Markov chain, i.e., $\{X_i\}_{i\geq 1}$ are observations with the chain dependence (see [4]).

References

1. I. V. Bokuchava, Z. A. Kvatadze, T. L. Shervashidze, On limit theorems for random vectors controlled by a Markov chain. *Probability theory and mathematical statistics, vol. I (Vilnius, 1985)*, 231–250, *VNU Sci. Press, Utrecht*, 1987.
2. N. N. Čencov, A bound for an unknown distribution density in terms of the observations. (Russian) *Dokl. Akad. Nauk SSSR* **147** (1962), 45–48.
3. L. Devroye, L. Györfi, *Nonparametric Density Estimation.* The L_1 view. Wiley Series in Probability and Mathematical Statistics: Tracts on Probability and Statistics. John Wiley & Sons, Inc., New York, 1985.
4. Z. Kvatadze, T. Shervashidze, On limit theorems for conditionally independent random variables controlled by a finite Markov chain. *Probability theory and mathematical statistics (Kyoto, 1986)*, 250–258, Lecture Notes in Math., 1299, *Springer, Berlin*, 1988.
5. G. M. Mania, *Statistical Estimation of a Probability Distribution Function.* (Russian) Izdat. Tbilis. Univ., Tbilisi, 1974.
6. R. M. Mnacakanov, E. B. Khmaladze, L_1-convergence of statistical kernel estimates of probability density-functions. *Dokl. Akad. Nauk SSSR* **258** (1981), no. 5, 1052–1055.
7. E. A. Nadaraya, *Nonparametric Estimation of a Probability Density Function and a Regression Curve.* Tbilisi univ. publishing house, Tbilisi, 1983.
8. E. Parzen, On the estimation of a probability density function and mode. *Ann. Math. Statist.* **33** (1962), no. 3, 1065–1076.
9. M. Rosenblatt, Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27** (1956), 832–837.

10. G. S. Watson, M. R. Leadbetter, On the estimation of the probability density. *I. Ann. Math. Statist.* **34** (1963), 480–491.

(Received 23.01.2019)

Department of Mathematics, Georgian Technical University, 77 Kostava Str., Tbilisi, Georgia
E-mail address: zurakvatadze@yahoo.com
E-mail address: beqnufarjiani@yahoo.com