On the Exactness of Distribution Density Estimates Constructed by Some Classes of Dependent Observations

Zurab Kvavadze1,*, Beqnu Pharjiani2

1Department of Mathematics, Georgian Technical University, United States
2Faculty of Informatics and Management Systems, Georgian Technical University, United States

Received May 6, 2019; Revised September 16, 2019; Accepted September 25, 2019

Abstract On the probabilistic space \((\Omega, F, P)\) we consider a given two-component stationary (in the narrow sense) sequence \(\{\xi_i, X_i\}_{i \in \mathbb{Z}}\), where \(\{\xi_i\}_{i \in \mathbb{Z}} (\xi_i : \Omega \rightarrow \Xi)\) is the controlling sequence and the members \(X_i : \Omega \rightarrow R\) of the sequence \(\{X_i\}_{i \in \mathbb{Z}}\) are the observations of some random variable \(X\) which are used in the construction of kernel estimates of Rosenblatt-Parzen type for an unknown density \(f(x)\) of the variable \(X\). The cases of conditional independence and chain dependence of these observations are considered. The upper bounds are established for mathematical expectations of the square of deviation of the obtained estimates from \(f(x)\).

Keywords Conditionally Independent Sequence, Sequence with Chain Dependence, Kernel Estimate, Markov Chain

1. Introduction

Nonparametric estimates of a distribution density are one of the intensively studied issues of mathematical statistics. The previously studied estimates were constructed using independent samples. In recent years the construction of estimates has been done by dependent observations.

Our present study deals with two types of dependence: conditional independence and chain dependence.

Let us present some definitions and auxiliary facts concerning the nonparametric estimation of the distribution density by independent observations.

Assume that the variables \(X_i, X_j \in R, i = 1,2,\ldots\), are independent observations of the random variable \(X\) with an unknown density \(g(x)\). The first attempts to construct nonparametric density estimates were made by V. Glivenko ([1]) and N. Smirnov ([2]), who built histogram-type estimates. In the works of M. Rosenblatt and E. Parzen (see [3], [4]) the density \(g(x)\) was estimated by considering the class of estimates

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

where \(\{a_n\}_{n \in \mathbb{Z}}\) is the sequence of positive numbers such that

\[
\lim_{n \to \infty} a_n = \infty, \quad a_n = o(n)
\]

and the kernel \(k(x)\) is a Lebesgue-integrable Borel function. Watson G. and Leadbetter M. ([5]) considered a more general class of kernel estimates. Mania G. in 1970 ([6]) considered kernel estimates for observations \(X_i \in R^d\) and applied the multidimensional kernel \((k(x) \in R^d)\). Nadaraya E. A. in 1980 ([7]) generalized the result of Rosenblatt M and established sufficient conditions for the uniform convergence \(\hat{g}_n(x,a_n)\) to \(g(x)\) with probability 1 to the metric \(L_2\). Along with Rosenblatt and Parzen type estimates, projection type estimates were also considered (see [7], [8]) when the spectral decomposition of the kernel \(k(x)\) with respect to the basis of orthonormal functions. L. Devroye (see [9])
constructed by means of smoothing functions the density estimate with a finite number of discontinuity points. The measure of variance of estimates from \( g(x) \) considered different characteristics: L. Devroy ([9]) considered the accuracy of density approximation by the constructed estimate by metric \( L^1 \). Also, by metric \( L^1 \) was considered the accuracy of density approximation by the built by estimates of Khmaladze E. B and Mnacakanov R. M. ([10]), although they considered a more general class of nuclear estimations. In the works ([3], [7], [8]) the accuracy of density approximation by metric \( L^2 \) was considered.

E. Nadaraya [7] obtained the sufficient conditions for the uniform convergence of \( \hat{g}_n(x,a_n) \) to \( g(x) \) with probability 1. The measure of divergence between them was the value

\[
E \int_{-\infty}^{\infty} [\hat{g}_n(x,a_n) - g(x)]^2 \, dx.
\]

**Definition 1.** Denote by \( H_s \) \((s \geq 2; \ S \) is an even number\) the set of functions \( k(x) \) which satisfy the conditions

\[
k(-x) = k(x), \quad \int_{-\infty}^{\infty} k(x) \, dx = 1, \quad \sup_{-\infty}^{\infty} |k(x)| \leq A < \infty,
\]

\[
\int_{-\infty}^{\infty} x^i k(x) \, dx = 0, \quad i = 1, 2, \ldots, s - 1;
\]

\[
\int_{-\infty}^{\infty} x^i k(x) \, dx \neq 0, \quad \int_{-\infty}^{\infty} x^i |k(x)| \, dx < \infty. \tag{3}
\]

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

**Lemma 1** (see [7]). If the values \( X_i, x_i \in R, \ i = 1, 2, \ldots \), are independent observations of some random variable \( X \) with an 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

\[
\int_{-\infty}^{\infty} D\hat{g}_n(x,a_n) \, dx = \frac{a_n}{n} \int_{-\infty}^{\infty} k^2(x) \, dx + O(\frac{a_n}{n}) \tag{4}
\]

\[
\int_{-\infty}^{\infty} \left[ E\hat{g}_n(x,a_n) - g(x) \right]^2 \, dx = \frac{a_n^2}{n} \frac{\alpha^2}{(s!)^2} \int_{-\infty}^{\infty} \left[ g^{(s)}(x) \right]^2 \, dx + O(a_n^2) \tag{5}
\]

hold, where \( \{a_n\}_{n \geq 1} \) is the sequence (is defined by) \( (2) \) and

\[
\alpha = \int_{-\infty}^{\infty} x^s k(x) \, dx.
\]

Let us present our result.

In practice we encounter cases where the distribution of the observed value \( X \) changes at random moments of time depending on the conditions (the controlling sequence \( \{\xi_i\}_{i \geq 1} \)). Thereby the observation densities \( \{X_i\}_{i \geq 1} \) also undergo changes. For example, in stock-exchange transactions the price of some goods changes according to a season though it had been previously fixed at the auction. Hence the flow of revenue due to such transactions also changes and so on.

In such cases to estimate the density \( X \) it is advisable to consider dependent observations. It should be noted one interesting work by Sidney Yakowitz ([11]) concerning density estimation by observations bound to Markov chain, which deals with Markov chain by general phase space of states.

On the probabilistic space \( (\Omega, F, P) \) let us consider the two-component stationary \( (\text{in the narrow sense}) \) sequence of random variables

\[
\{\xi_i, X_i\}_{i \geq 1} \tag{6}
\]

Where, \( \xi_i : \Omega \to \Xi \), \( X_i : \Omega \to R^n \) and \( \Xi \) is some space.

**Definition 3.** The sequence \( \{X_i\}_{i \geq 1} \) in (6) is called a conditionally independent sequence (see [12]) controlled by the sequence \( \{\xi_i\}_{i \geq 1} \) if for any natural \( n \) and the fixed trajectory \( \overline{\xi_{ln}} = (\xi_{i1}, \xi_{i2}, \ldots, \xi_{in}) \) the values \( X_1, X_2, \ldots, X_n \) become independent and for all natural numbers \( i, k, n \), \( j_1, j_2, \ldots, j_k \), \( 2 \leq k \leq n; \ n \leq i \leq n; \ 1 \leq j_1 < j_2 < \cdots < j_k \leq n \) the equalities
are fulfilled, where $P_{X|Y}$ is the conditional distribution of the value $X$ under the condition $Y$.

**Definition 4.** The conditionally independent sequence $\{X_i\}_{i=1}^\infty$ in (1) is called a sequence with chain dependence (see [12]) if $\{\xi_i\}_{i=1}^\infty$ is a finite Markov chain with discrete time.

**Lemma 2** (see [13]). Let $\{\xi_i\}_{i=1}^\infty$ be a finite stationary regular Markov chain with the set of states $\{b_1, b_2, \ldots, b_k\}$. Let $v_n(i)$, $i = \overline{1,k}$, be the frequencies of those moments of time at which during the first $n$ steps the chain is in the states $b_i$, $i = \overline{1,k}$, respectively. Then the relations

$$E\frac{v_n(i)}{n} = \pi_i, \quad D\left(\frac{v_n(i)}{n}\right) \leq c_i(\pi, p), \quad i = \overline{1,k}, \quad (7)$$

are fulfilled, where $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$ is the initial distribution of the chain, $P = (p_{ij})_{i,j=1,k}$ is the matrix of transient probabilities and $c_i(\pi, p)$, $i = \overline{1,k}$, are the constants which depend on the parameters of the chain.

**3. Main Results**

Let us consider the sequence (6), where $\xi_i$, $i = 1,2,\ldots$, are identically distributed independent discrete random values. Let $\Xi = \{b_1, b_2\}$; $P(\xi_i = b_i) = p_i$, $i = \overline{1,2}, \quad p_1 + p_2 = 1$.

For the fixed trajectory $\overline{\xi}_{i=1}^n = (\xi_1, \xi_2, \ldots, \xi_n)$ of the sequence $\{\xi_i\}_{i=1}^\infty$ we denote by $v_n(1)$ and $v_n(2)$ the frequencies when the first $n$ members of the sequence takes the values $b_1$ and $b_2$ respectively.

**Theorem 1.** Let us consider the sequence (6). $\Xi = \{b_1, b_2\}$. The elements of the controlling sequence $\{\xi_i\}_{i=1}^\infty$ ($\xi_i : \Omega \rightarrow \{b_1, b_2\}$) are independent, identically distributed discrete random values $\xi_i = b_1 I_{\xi_i = b_1} + b_2 I_{\xi_i = b_2}$, $i = 1,2,\ldots$. Let for every function $\Psi : \Xi \rightarrow R^1$, for which $E\Psi(\xi_i) < \infty$, the convergence

$$\frac{1}{n} \sum_{j=1}^{n} \Psi(\xi_j) \rightarrow E\Psi(\xi_1) \quad a.s., \quad (8)$$

hold as $n \rightarrow \infty$.

Let elements of the conditionally independent sequence $\{X_i\}_{i=1}^\infty$ be the observations of the value $X$. And let the conditional distributions be $P_{X_i|\xi_i = b_i}$, $i = 1,2$, have respectively unknown densities $f_i(x)$, $i = 1,2$, from the class $f_i(x) \in W_2' \cap L_2(\Omega, \infty, \infty)$, and let $k(x) \in H_2 \cap L_2(\Omega, \infty, \infty)$. If the equalities

$$D\left(\frac{v_n(i)}{n}\right) \leq \frac{c_i}{\sqrt{n}}, \quad i = 1,2, \quad (9)$$

are fulfilled for the frequencies $v_n(i)$, $i = 1,2$, then for any natural $n$ the estimate of $\overline{f}(x) = p_1 f_1(x) + p_2 f_2(x)$ is...
\[ \hat{f}_n(x, a_n) = \frac{a_n}{n} \sum_{j=1}^{n} k \left( a_n (x - X_j) \right) \]

And for expression

\[ u(a_n) = E \int_{-\infty}^{\infty} \left[ \hat{f}_n(x, a_n) - \bar{f}(x) \right]^2 \, dx \]

holds, where

\[ u(a_n) \leq (M_1 + M_2)^2 + \frac{a_n}{n} \int_{-\infty}^{\infty} k^2(x) \, dx + \]

\[ + \left( (C_1 + C_2) n^{-1/2} + p_1^2 + p_2^2 \right) O\left( \frac{a_n}{n} \right) \]

is valid, where

\[ M_i = T^{1/2}_i + C n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) \, dx \]

\[ T_i = (a_n^{-2s} \alpha^2 / (s!)^2) \int_{-\infty}^{\infty} \left[ f_i^{(s)}(x) \right]^2 \, dx + \]

\[ + O\left( a_n^{-2s} \right) (C n^{-1/2} + p_i^2) \]

and

\[ \alpha = \int_{-\infty}^{\infty} x^k(x) \, dx \]

\[ E \int_{-\infty}^{\infty} \left[ \hat{f}_n(x, a_n) - (\pi_1 f_1(x) + \pi_2 f_2(x)) \right]^2 \, dx \leq \]

\[ \leq (M_1^* + M_2^*)^2 + \frac{a_n}{n} \int_{-\infty}^{\infty} k^2(x) \, dx + \]

\[ + \left( \frac{1}{n} \left( C_1(\pi p) + C_2(\pi p) \right) + \pi_1^2 + \pi_2^2 \right) O\left( \frac{a_n}{n} \right) \]

Theorem 2. Let in the sequence \( \{ \xi_i \}_{i \geq 1} \) be a finite, homogeneous regular Markov chain, \( \Xi = \{ b_1, b_2 \} \), with the initial distributions \( \pi = (\pi_1, \pi_2) \), \( \pi_i = P(\xi_1 = b_i) \), \( i = 1, 2 \), and the matrix of transient probabilities \( P = (p_{ij})_{i,j=1,2} \). Let \( \{ X_i \}_{i \geq 2} \) be the sequence with the chain dependence whose elements are observations of the value \( X \). Let further the conditional distributions \( \mathcal{P}_{X_i|X_i=b_i} \), \( i = 1, 2 \), have the unknown densities \( f_i(x) \), \( i = 1, 2 \), respectively. \( f_1(x) \in W_s \cap L_2(-\infty, \infty) \); \( k(x) \in H_s \cap L_2(-\infty, \infty) \). Then for any \( n \) the estimate of the density \( \bar{f}(x) = \pi_1 f_1(x) + \pi_2 f_2(x) \) is \( \hat{f}_n(x, a_n) \) and the estimate

\[ \hat{f}_n(x, a_n) = \frac{a_n}{n} \sum_{m=1}^{n} k \left( a_n (x - X_m(i)) \right) \]
It is naturally understood that if \( V_n(i) = 0, \ i = 1, 2, \) then the summand \( \hat{f}_{in}(x,a_i), \ i = 1, 2, \) does not exist.

Let us first prove the finiteness of the values \( E\hat{f}_{n}(x,a_n) \) and \( D\hat{f}_{n}(x,a_n). \)

For the fixed trajectory \( \xi_{1n}, \) we represent \( E\hat{f}_{n}(x,a_n) \) in the form of a conditional mathematical expectation

\[
E\hat{f}_{n}(x,a_n) = E\{E(\hat{f}_{n}(x,a_n)|\xi_{1n})\} = E\{E(\sum_{i=1}^{v_n(i)} f_{in}(x,a_n)|\xi_{1n})\}
\]

In view of the fact that the functions \( V_n(i) \) and \( V_n(2) \) are measurable with respect to the \( \sigma \)-algebra generated by the partitioning of \( \Omega \) that results from the fixing of the trajectory \( \xi_{1n}, \) they can be taken out outside the sign of the conditional mathematical expectation.

Here and in what follows it is assumed that the equality

\[
E \frac{V_n(i)}{n} = p_i
\]

is fulfilled by virtue of the condition (8) and the following estimate

\[
E \left( \frac{V_n(i)}{n} \right)^2 = D \left( \frac{V_n(i)}{n} \right) + E \left( \frac{V_n(i)}{n} \right)^2 \leq n^{-1/2}c_i + p_i^2
\]

Taking the conditions (3) into account and changing the variable under the sign of the integral we obtain the following chain of equalities is valid by virtue of the condition (9).

\[
E\hat{f}_{n}(x,a_n) = \sum_{i=1}^{2} E\left\{ \frac{V_n(i)}{n} E\left( \frac{a_n}{V_n(i)} \sum_{m=1}^{v_n(i)} k\left( a_n \left( x - X_{\tau_{m(i)}} \right) \right) \right) \mid \xi_{1n} \right\} = \\
= 2 \sum_{i=1}^{2} E\left\{ \frac{V_n(i)}{n} E\left( \frac{a_n}{V_n(i)} \sum_{m=1}^{v_n(i)} k\left( a_n \left( x - X_{\tau_{m(i)}} \right) \right) \right) \mid \xi_{1n} \right\}
\]

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

Taking into account the fact that on the fixed trajectory \( \xi_{1n} \) the sums \( \hat{f}_{in}(x,a_n) \) and \( \hat{f}_{2n}(x,a_n) \) as well as their constituent summands are independent, the following equalities are valid:

\[
D\hat{f}_{n}(x,a_n) = E\{E([\hat{f}_{n}(x,a_n) - E\hat{f}_{n}(x,a_n)]^2 \mid \xi_{1n})\} = \\
= E\{E([\sum_{i=1}^{2} \frac{V_n(i)}{n} \hat{f}_{in}(x,a_n)]^2) \mid \xi_{1n})\} = \\
= E\{E([\sum_{i=1}^{2} \frac{V_n(i)}{n} (\hat{f}_{in}(x,a_n) - E\hat{f}_{in}(x,a_n))]^2) \mid \xi_{1n})\} = \\
= \sum_{i=1}^{2} E\{\left( \frac{V_n(i)}{n} \right)^2 E([\hat{f}_{in}(x,a_n) - E\hat{f}_{in}(x,a_n)]^2) \mid \xi_{1n})\} = \\
= \sum_{i=1}^{2} E\{\left( \frac{V_n(i)}{n} \right)^2 E([\hat{f}_{in}(x,a_n) - E\hat{f}_{in}(x,a_n)]^2) \mid \xi_{1n})\} = \\
= \sum_{i=1}^{2} E\{\left( \frac{V_n(i)}{n} \right)^2 E\left( \sum_{j=1}^{v_n(i)} a_n \left( x - X_{\tau_{j(i)}} \right) \right) \mid \xi_{1n})\} = \\
= \sum_{i=1}^{2} E\{\left( \frac{V_n(i)}{n} \right)^2 E\left( \sum_{j=1}^{v_n(i)} a_n \left( x - X_{\tau_{j(i)}} \right) \right) \mid \xi_{1n})\}
\]
On the Exactness of Distribution Density Estimates Constructed by Some Classes of Dependent Observations

\[ E\{V_n(i)\}^2 \cdot (\frac{\alpha_n}{V_n(i)})^2 \cdot E\left(\sum_{j=1}^{V_n(i)} k(a_n(x - X_{ji})) - Ek(a_n(x - X_{ji}))\right)^2 | \xi_{1n}) = \]

\[ = \sum_{i=1}^{2} E\{(V_n(i)) (a_n)^2 \cdot V_n(i) \left[ k(a_n(x - u)) \right] \int_{-\infty}^{+\infty} k(a_n(x - y)) f_i(y) dy \cdot f_i(u) du \]

Applying (11), we obtain the following expression for \( Df_n(x, a_n) \)

\[ Df_n(x, a_n) = \frac{a_n}{n} \sum_{i=1}^{2} p_i \int\int k(t) k(a_n(x - y)) f_i(y) dy f_i(t + x) dt \]

Let us now estimate \( u(a_n) \). We use Fubini’s theorem and divide \( u(a_n) \) into two parts:

\[ u(a_n) = \int_{-\infty}^{\infty} E\left(\hat{f}_n(x, a_n) - \bar{f}(x)\right)^2 dx = \]

\[ = \hat{f}_n(x, a_n) - E\hat{f}_n(x, a_n) \] \[ dx + \int_{-\infty}^{\infty} E\left[E\hat{f}_n(x, a_n) - \bar{f}(x)\right]^2 dx = I_1 + I_2. \] (13)

To estimate \( I_1 \), we again use Fubini’s theorem and the independence of the sums \( \hat{f}_{1n}(x, a_n) \) and \( \hat{f}_{2n}(x, a_n) \). We obtain

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

\[ = E \int E(\sum_{i=1}^{2} \frac{V_n(i)}{n} (\hat{f}_{1in}(x, a_n) - E\hat{f}_{1in}(x, a_n)))^2 | \xi_{1n})dx = \]

\[ = E \{ E(\sum_{i=1}^{2} \left( \frac{V_n(i)}{n} \right)^2 [\hat{f}_{1in}(x, a_n) - E\hat{f}_{1in}(x, a_n)]^2 | \xi_{1n})dx \} = \]

\[ = E \{ \sum_{i=1}^{2} \left( \frac{V_n(i)}{n} \right)^2 E([\hat{f}_{1in}(x, a_n) - E\hat{f}_{1in}(x, a_n)]^2 | \xi_{1n})dx \} \]

Applying the equality (4) from Lemma 1 we have

\[ I_1 = \sum_{i=1}^{2} E\left\{ \left( \frac{V_n(i)}{n} \right)^2 \left( \frac{a_n}{V_n(i)} \int_{-\infty}^{+\infty} k(x) dx + o(a_n) \right) \right\} = \]

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

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

Now applying the inequality (12) we complete the estimation of \( I_1 \):
\[ I_1 \leq \frac{a_n}{n} \int_{-\infty}^{+\infty} k^2(x)dx \sum_{i=1}^{2} p_i + o\left(\frac{a_n}{n}\right) \sum_{i=1}^{2} (C_i n^{-1/2} + p_i^2) = \]
\[ = \frac{a_n}{n} \int_{-\infty}^{+\infty} k^2(x)dx + o\left(\frac{a_n}{n}\right) \sum_{i=1}^{2} (C_i n^{-1/2} + p_i^2). \]

Let us decompose \( I_2 \) from (13):

\[ I_2 = \int_{-\infty}^{+\infty} E\left[ \widehat{E f_n}(x,a_n) - \overline{f}(x) \right]^2 dx = E \int_{-\infty}^{+\infty} \left[ \widehat{E f_n}(x,a_n) - \overline{f}(x) \right]^2 dx = \]
\[ = E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \sum_{i=1}^{2} \frac{V_n(i)}{n} \widehat{f}_{in}(x,a_n) - \sum_{i=1}^{2} p_i f_i(x) \right)^2 dx \right| \xi_{1,n} \right\} = E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \sum_{i=1}^{2} \frac{V_n(i)}{n} \widehat{f}_{in}(x,a_n) - \sum_{i=1}^{2} p_i f_i(x) \right)^2 dx \right| \xi_{1,n} \right\} = \]
\[ = E \left\{ E \left( \sum_{i=1}^{2} \left( \frac{V_n(i)}{n} \widehat{f}_{in}(x,a_n) - p_i f_i(x) \right)^2 dx \right| \xi_{1,n} \right\} + 2 E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} \widehat{f}_{1n}(x,a_n) - p_1 f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} + \]
\[ = E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} \widehat{f}_{1n}(x,a_n) - p_1 f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} + 2 E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} \widehat{f}_{1n}(x,a_n) - p_1 f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} = I_{21} + I_{22} + 2 I_{23} \]

The first two summands from the sum are estimated in the same manner:

\[ I_{21} = E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} \widehat{f}_{1n}(x,a_n) - p_1 f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} = \]
\[ = E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} \widehat{f}_{1n}(x,a_n) - f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} + E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} - p_1 \right)^2 f_1(x) dx \right| \xi_{1,n} \right\} + \]
\[ + 2 E \left\{ E \left( \int_{-\infty}^{+\infty} \left( \frac{V_n(1)}{n} - p_1 \right) f_1(x) dx \right| \xi_{1,n} \right\} = A_1 + A_2 + A_3 \]

Using the equality (5) from Lemma 1 and the estimate (12) we obtain

\[ A_1 = E \left\{ \left( \frac{V_n(1)}{n} \right)^2 E \left( \int_{-\infty}^{+\infty} \left( \widehat{f}_{1n}(x,a_n) - f_1(x) \right)^2 dx \right| \xi_{1,n} \right\} = \]
\[ = E \left\{ \left( \frac{V_n(1)}{n} \right)^2 \left( a_n^{-2s} \frac{\alpha^2}{(s!)^2} \int_{-\infty}^{+\infty} \left[ f^{(s)}(x) \right]^2 dx + o(a_n^{-2s}) \right) \right\} = \]
\[ \leq \left( a_n^{-2s} \frac{\alpha^2}{(s!)^2} \right) \int_{-\infty}^{+\infty} \left[ f^{(s)}(x) \right]^2 dx + o(a_n^{-2s}) \right) (C_i n^{-1/2} + p_i^2) = T_1 \]

It is not difficult to obtain the estimate of the sum \( A_2 \)
On the Exactness of Distribution Density Estimates Constructed by Some Classes of Dependent Observations

\[ A_2 = E\{ \left( \frac{V_n(1)}{n} - p_i \right)^2 \} E \int_{-\infty}^{\infty} f_i^2(x) dx \leq C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) dx \]

We use the fact that the values \( \frac{V_n(i)}{n} - p_i, \quad i = 1, 2 \), as well as \( \frac{V_n(i)}{n} \) are measurable with respect to the \( \sigma \)-algebra generated by the fixed trajectory \( \xi_{1n} \). Applying Holder’s inequality we estimate \( A_3 \) as follows

\[ A_3 = E\{ \left( \frac{V_n(1)}{n} - p_i \right)^2 \} E \int_{-\infty}^{\infty} [E_{\xi_{1n}} f_i(x) - f_i(x)] f_i(x) dx \leq \]

\[ \leq E\{ \left( \frac{V_n(1)}{n} - p_i \right)^2 \} E \left( \int_{-\infty}^{\infty} [E_{\xi_{1n}} f_i(x) - f_i(x)]^2 dx \right) \left( \int_{-\infty}^{\infty} f_i^2(x) dx \right) \]

\[ \leq a_n^{-2s} \frac{\alpha^2}{(s!)^2} \int_{-\infty}^{\infty} \left[ f_i^{(s)}(x) \right]^2 dx + \alpha(a_n^{-2s}) \int_{-\infty}^{\infty} f_i^2(x) dx \cdot \left( \frac{V_n(1)}{n} - p_i \right)^2 \leq \]

\[ \leq \sqrt{T_1} \cdot \left( C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) dx \right) \]

Combining the estimates \( A_1, A_2 \) and \( A_3 \) we obtain

\[ I_{21} \leq (\sqrt{T_1} + \sqrt{C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) dx})^2 \equiv M_1^2, \quad (15) \]

where,

\[ M_1 = \sqrt{T_1} + \sqrt{C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) dx} \]

In the same manner it can be shown that

\[ I_{22} \leq (\sqrt{T_2} + \sqrt{C_i n^{-1/2} \int_{-\infty}^{\infty} f_i^2(x) dx})^2 \equiv M_2^2, \quad (16) \]

where,

\[ T_2 \equiv a_n^{-2s} \frac{\alpha^2}{(s!)^2} \int_{-\infty}^{\infty} \left[ f_i^{(s)}(x) \right]^2 dx + \alpha(a_n^{-2s})(C_i n^{-1/2} + p_2^2) \]

Let us consider the sum \( I_{22} \). Using Fubini’s and Holder’s theorems we have
Each of the obtained two factors is estimated as the summands $I_{21}$ and $I_{22}$. We obtain

$$I_{23} \leq M_1 M_2 \quad (17)$$

By the decomposition (13) and the obtained estimates (14-17) the theorem is proved.

**Proof of Theorem 2.** The proof is carried out in the same way as that of Theorem 1, but we should note that there is some difference.

For the fixed trajectory $\xi_{1n}$ we denote by the symbols $\nu_n(i) \ i = 1, 2$, the frequencies of the chain being in the states $b_1$ and $b_2$ (respectively) during the first $n$ steps.

We do not need the conditions (8) and (9) of Theorem 1. As we know (see [10]), if the so-called condition of ergodicity is fulfilled for regular Markov chains, then the condition (8) is fulfilled for any continuous function of the chain states. For regular Markov chains the relations (7) from Lemma 2 are valid

$$E\left(\frac{\nu_n(i)}{n}\right) = \pi_i, \ D\left(\frac{\nu_n(i)}{n}\right) \leq \frac{c_i(\pi, \rho)}{n} \quad i = 1, k.$$  

Hence there holds the following estimate

$$E\left(\frac{\nu_n(i)^2}{n}\right) = D\left(\frac{\nu_n(i)}{n}\right) + E\left(\frac{\nu_n(i)}{n}\right)^2 \leq \frac{c_i(\pi, \rho)}{n} + \pi_i^2,$$

which we use instead of the estimate (9).

**Example 1.**

We may have one such interpretation of a conditional independent sequence

Consider $\{\xi_i\}_{i=1}^r$ the sequence independent discrete randomly distributed values with the distribution $P(\xi_i = b_i) = p_i \ i = 1, r \ p_1 + p_2 + \ldots + p_r = 1$.

Consider the $r$ number of sequences $\{X^\alpha\}_{i=1}^r$, $\alpha = 1, r$ of independent random values. Assume that the members of the corresponding sequence of each $\alpha$ index have the same distribution $P_\alpha$, distribution $\alpha = 1, r$. We have $r$ sequences.

$$X_1^\alpha, X_2^\alpha, X_3^\alpha, \ldots, X_n^\alpha,$$

$$X_1^r, X_2^r, X_3^r, \ldots, X_n^r,$$

At each step, if each member of the sequence gets a value of $b_\alpha$ ($1 \leq \alpha \leq r$), then take the $X_i^\alpha$ vector $\alpha = 1, r$ from the column of $(X_1^\alpha, X_2^\alpha, \ldots, X_n^\alpha)^T$ vectors. Call it $Y_i$. We get a conditional independent sequence

$$Y_{i} = \sum_{\alpha=1}^{r} X_{i}^\alpha I(\xi_i = \alpha),$$

where $I_A$ is an indicator of $A$ set. It is clear that $Y_1, Y_2, \ldots, Y_n$ represent conditionally independent random vectors, with a fixed “trajectory” $\overline{\xi}_{1n} = (\xi_1, \xi_2, \ldots, \xi_n)$. 

$$\xi_i : \Omega \rightarrow \{b_1, b_2, \ldots, b_r\}, \ \xi_i = \sum_{j=1}^{r} b_j I(\xi_i = b_j) \text{ of distribution } P(\xi_i = b_i) = p_i \ i = 1, r,$$
If \( \{\xi_j\}_{j \geq 1} \) is a finite chain of Markov with \( \{b_1, b_2, \ldots, b_r\} \) set of states, then

\[
Y_i = \sum_{\alpha=1}^{r} X_i^{\alpha} I(\xi_i = \alpha),
\]

sequence will be chain defendant sequence.

Example 2.

Let us take an example of the practical application of the theorem approved in the article.

At the Frankfurt Stock Exchange in the winter season (Börse Frankfurt) 2018, the fruit brokerage agreements set the price of apples at 3.9 euros.

Naturally, let us consider the price of apples to be normally distributed random variables. Suppose the price of apples during the winter season is normally distributed with parameters \( N(2.5, 1.44) \), and let us assume that during the summer season we do not know these parameters \( N(1, 1.21) \). Let us estimate the price density of apples by observing contract prices.

Let us consider the sequence

\[
\{\xi_j, X_j\}_{j \geq 1}
\]

where \( \{\xi_j\}_{j \geq 1} \) are the days of observations, and \( \{X_j\}_{j \geq 1} \) is the apple price for brokerage transactions during the relevant days.

Let winter days be 1 and the summer days be 2.

Suppose the desire to trade apples during winter is three times as high as summer season, therefore \( p(\xi_1 = 1) = \frac{3}{4} \), \( p(\xi_1 = 2) = \frac{1}{4} \). Suppose, the sequence (18) is stationary in the narrow sense. For simplicity of computation, consider only 4 observations: \( X_1 = 1.2 \) (EUR); \( X_2 = 1 \) (EUR); \( X_3 = 2.8 \) (EUR); \( X_4 = 3.2 \) (EUR).

For building the \( \hat{f}(x) = \frac{1}{4} f_1(x) + \frac{1}{4} f_2(x) \) density estimation let us use the core of Bartlett

\[
\hat{f}_n(x) = \frac{3}{4\sqrt{n}} \sum_{j=1}^{n} (1 - \sqrt{n}(x - X_j)^2) I_{|x - X_j| \leq \sqrt{n}},
\]

(19)

Suppose that the terms of (8) and (9) are fulfilled. Consider that \( n = 4 \), insert in (19) the values of \( X_1; X_2; X_3 \) and \( X_4 \). The result of calculation will be:

\[
\begin{align*}
& -1.5x^2 + 3x - 1.125 \quad \text{when} \ x \in [0.5; 0.7]; \\
& -3x^2 + 6.6x - 2.91 \quad \text{when} \ x \in [0.7; 1.5]; \\
& -1.5x^2 + 3.6x - 1.785 \quad \text{when} \ x \in (1.5; 1.7]; \\
& \hat{f}_n(x) = -1.5x^2 + 8.4x - 11.385 \quad \text{when} \ x \in [2.3; 2.7]; \\
& -3x^2 + 1.8x - 26.37 \quad \text{when} \ x \in [2.7; 3.3]; \\
& -1.5x^2 + 9.6x - 14.985 \quad \text{when} \ x \in (3.3; 3.7]; \\
& 0 \quad \text{when} \ x \in (-\infty; 0.5) \cup (1.7; 2.3) \cup (3.7; \infty).
\end{align*}
\]

The price distribution function will be

\[
F(x) = \int_{-\infty}^{x} f_4(x) dx.
\]

4. Discussion

\[
\frac{v_n(i)}{n} \quad \text{and} \quad v_n(i) \quad \text{multiples made it difficult to obtain estimates when proving theorems. It was necessary to calculate the mathematical expectation of the estimated value on the fixed \( \{\xi_1^{(n)} = (\xi_1, \xi_2, \ldots, \xi_n) \) trajectory and the product of these multiples. It has been shown that} \quad (v_n(i)) \quad \text{and} \quad f(v_n(i)) \quad \text{functions (when} \quad f \in L_1(\infty, \infty) \quad \text{are measures toward} \quad \sigma \text{-algebra induced by the division of the} \quad \Omega \quad \text{space generated by the fixation of the} \quad \xi_{1n} \quad \text{trajectory (see [14]). This fact has allowed us to extend these values beyond the conditional mathematical expectation.}}
\]

The \( E \hat{f}_n(x, a_n) \) and \( D \hat{f}_n(x, a_n) \) values show that the remainder is finite in the \( I_1 \) addend, the remainder is estimated by Lemma 1, the remainder in the \( A_2 \) addend is estimated using the Hölder inequality.

5. Conclusions

Until now, the estimates were based on independent observations. Their accuracy was also established in the case of independent observations.

The obtained results allow to construct the Rosenblatt-Parzen density estimation by means of some type of dependent observations (conditionally independent and chain dependent).
Therefore, we built nonparametric density estimates with conditional observations (conditionally independent and chain-dependent observations) and determined the approximation densities with these estimates.

REFERENCES

[1] Glivenko V. I Probability theory course (Russian) M. "The science" 1939.218 s.;

[2] Smirnov N. V O. Random distribution density approximation. (Russian) Uch. Zapisi M.G.P.I im V.I. Potiomkina 1951. 16. 3 p. 69-96. s.,

[3] Rosenblatt M. (1956), Remarks on some nonparametric estimates of a density function. Ann. Math. Statist.27: 832-837. Chicago;

[4] Parzen E.(1962) On estimation of a probability density function and mode. Ann. Math. Statist. 33: 1065-1076. Stanford;

[5] Watson G. S, Leadbetter M. R. (1963) on the estimation of the probability density. I. Ann. Math. Statist. 34: 480-491 Toronto;

[6] G.M. Mania (1974), Statistical estimation of the probability distribution. (Russian) Tbilisi State Univ. Press, Tbilisi, 1974;

[7] E.A. Nadaraya (1983), Non parametric estimation of the probability density and regression curve. (Russian) Tbilisi State Univ. Press, 1983;

[8] N.N. Chentsov (1952), Estimation of an unknown distribution density in terms of the observations. (Russian) Dokl. Akad. Nauk SSSR 147, 45-48;

[9] L. Devroye, L. Dyoří (1988), Nonparametric density estimation. L1-approach. Translated from English, edited by M.B. Malyutov, Mir, Moscow, 407 p.;

[10] R.M. Mnatsakanov, E.M. Khmaladze (1981), On the L1-convergence of statistical kernel estimates of distribution densities. (Russian) Dokl. Akad. Nauk SSSR 258, no. 5, Moscow, 1052-1055;

[11] Sidney Yakowitz. Nonparametric Density and Regression Estimation for Markov Sequences without Mixing Assumptions. Systems and Industrial Engineering Department. University of Arizona, Tucson, Arizona 85721–JOURNAL OF MULTIVARIATE ANALYSIS 30, p. 124-136. 1989;

[12] Bokuehava I. V., Kvadadze Z. A., and Shervashidze T. L., (1985) on limit theorems for random vectors controlled by a Markov chain. Probability theory and mathematical statistics, Vol. I: p. 231-250 Vilnius (Lithuania);

[13] Dj.Kemeny., Dj. Snell. (1970) Finite Markov chains. (Russian) Moscow. 272 p.;

[14] Kvadadze Z., Kvadadze TS., Maisuradze A., Limiting Distribution of a Sequence of Functions Defined on a Markov Chain. XXXIII Enlarged Sessions of the Seminar of Ilia Vekua Institute of Applied Mathematics (VIAM), of