Mathematics and Statistics 7(4): 135-145, 2019 http://www.hrpub.org
DOI: 10.13189/ms.2019.070407

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

Zurab Kvatadze[1,*], Beqnu Pharjiani[2]

[1]Department of Mathematics, Georgian Technical University, United States
[2]Faculty 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 (Ω, F, P) we consider a given two-component stationary (in the narrow sense) sequence $\{\xi_i X_i\}_{i\geq 1}$, where $\{\xi_i\}_{i\geq 1}$ ($\xi_i : \Omega \to \Xi$) is the controlling sequence and the members $X_i : \Omega \to R$ of the sequence $\{X_i\}_{i\geq 1}$ 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, $X_i \in R$, i = 1,2…, 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

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

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

$$\lim_{n\to\infty}a_n=\infty,\quad a_n=o\left(n\right) \qquad (2)$$

and the kernel $k(x)$ is a Lebesgue-integrable Borel function. Watson G. and Leadbetter M. ([5]) considered a more general class of nuclear estimates. Mania G. in 1970 ([6]) considered nuclear estimates for observations $X_i \in R^k$ and applied the multidimensional kernel ($k(x)\in R^k$). Nadaraya E. A. in 1980 ([7]) generalized the result of Rosenblatt M and established sufficient conditions for the uniform convergence $\hat{g}_n\left(x,a_n\right)$ to $g(x)$ k 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 ($s\ge 2$; s is an even number) the set of functions $k(x)$ which satisfy the conditions

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

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

$$\int_{-\infty}^{\infty}x^s k(x)\,dx\ne 0,\quad \int_{-\infty}^{\infty}x^s|k(x)|\,dx<\infty. \qquad (3)$$

Definition 2. Denote by W_s the set of functions $\varphi(x)$ having derivatives up to s-th order ($s\ge 2$) inclusive, and note that $\varphi^{(s)}(x)$ 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

$$\int_{-\infty}^{\infty}D\hat{g}_n(x,a_n)\,dx=\frac{a_n}{n}\int_{-\infty}^{\infty}k^2(x)\,dx+O\left(\frac{a_n}{n}\right) \qquad (4)$$

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

$$=a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}\left[g^{(s)}(x)\right]^2dx+O\left(a_n^{-2s}\right) \qquad (5)$$

hold, where $\{a_n\}_{n\ge1}$ 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 $\{\xi_i\}_{i\ge1}$). Thereby the observation densities $\{X_i\}_{i\ge1}$ 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 (Ω, F, P) let us consider the two-component stationary (in the narrow sense) sequence of random variables

$$\{\xi_i,X_i\}_{i\ge1} \qquad (6)$$

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

Definition 3. The sequence $\{X_i\}_{i\ge1}$ in (6) is called a conditionally independent sequence (see [12]) controlled by the sequence $\{\xi_i\}_{i\ge1}$ if for any natural n and the fixed trajectory $\overline{\xi}_{1n}=(\xi_1,\xi_2,\ldots,\xi_n)$ the values $X_1,X_2,\ldots,X_n$ become independent and for all natural numbers i, k, n, $j_1,j_2,\ldots,j_k$, $(2\le k\le n;\ i\le n;\ 1\le j_1<j_2<,\cdots,<j_k\le n)$ the equalities

$$\mathcal{P}_{\left(X_{j1},X_{j2},...,X_{jk}\right)|\xi_{1n}}=\mathcal{P}_{X_{j1}|\xi_{j1}}\times\mathcal{P}_{X_{j2}|\xi_{j2}}\times\cdots\times\mathcal{P}_{X_{jk}|\xi_{jk}},$$

$$\mathcal{P}_{X_i|\xi_{1n}}=\mathcal{P}_{X_i|\xi_i},$$

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

Definition 4. The conditionally independent sequence $\{X_i\}_{i\ge1}$ in (1) is called a sequence with chain dependence (see [12]) if $\{\xi_i\}_{i\ge1}$ is a finite Markov chain with discrete time.

Lemma 2 (see.[13]). Let $\{\xi_i\}_{i\ge1}$ be a finite stationary regular Markov chain with the set of states $\{b_1,b_2,\dots,b_k\}$. Let $\nu_n(i)$, $i=\overline{1,k}$, 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

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

are fulfilled, where $\pi=(\pi_1,\pi_2,\dots,\pi_k)$ is the initial distribution of the chain, $P=(p_{ij})_{i,j=\overline{1,k}}$ 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.

2. Methodology

The method used to validate the theorems is to transition from a mathematical wait to a conditional mathematical wait during the observation of a $\overline{\xi}_{1n}=(\xi_1,\xi_2,...,\xi_n)$ trajectory

$$(E[\widehat{f}_n(x)]=E\{E[\widehat{f}_n(x)\big|\overline{\xi}_{1n}]\}).$$

The sum to be considered on fixed trajectory is divided into several summons. One of them is the sum of independent summons on fixed trajectory.After some transformations it results as a type for which famous results are used from the theory of building of estimations with independent observation ([7], [9]). For estimations of the rest of the sums inequalities of Fubin, Helder and Jensen are and the characteristics of the governing sequences.For the estimation of all summons is based on the fact that the $\nu_n(i)$ and $f(\nu_n(i))$ functions (when $f\in L_2(-\infty,\infty)$ are measures toward σ-algebra induced by the division of the Ω space generated by the fixation of the ξ_{1n} trajectory (see [14]). Because of this, they are taken out of the conditional mathematical waiting mark.

The method used allows for the future to consider the estimation of various parameters by dependent observations.

3. Main Results

Let us consider the sequence (6), where ξ_i, $i=1,2,...$, are identically distributed independent discrete random values. Let

$$\Xi=\{b_1,b_2\};\ P(\xi_1=b_1)=p_i,\ i=\overline{1,2},$$

$$p_1+p_2=1.$$

For the fixed trajectory $\overline{\xi}_{1n}=(\xi_1,\xi_2,..,\xi_n)$ of the sequence $\{\xi_i\}_{i\ge1}$ we denote by $\nu_n(1)$ and $\nu_n(2)$ 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 $\{\xi_i\}_{i\ge1}$ $(\xi_i:\Omega\to\{b_1,b_2\})$ are independent, identically distributed discrete random values $\xi_i=b_1I_{\xi_1=b_1}+b_2I_{\xi_1=b_2}$. $i=1,2,...$. Let for every function $\Psi:\Xi\to R^1$, for which $E\Psi(\xi_i)<\infty$, the convergence

$$\frac{1}{n}\sum_{j=1}^{n}\Psi(\xi_i)\to E\Psi(\xi_1)\ \text{a. s.},\qquad (8)$$

hold as $n\to\infty$.

Let elements of the conditionally independent sequence $\{X_i\}_{i\ge1}$ be the observations of the value X. And let the conditional distributions be $\mathcal{P}_{X_i|\xi_i=b_i}$ $i=1,2$, have respectively unknown densities $f_i(x)$, $i=1,2$, from the class $f_i(x)\in W_s\cap L_2(-\infty,\infty)$, and let $k(x)\in H_s\cap L_2(-\infty,\infty)$. If the equalities

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

are fulfilled for the frequencies $\nu_n(i)$, $i=1,2$, then for any natural n the estimate of $\overline{f}(x)=p_1f_1(x)+p_2f_2(x)$ is

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

And for expression $u(a_n) = E\int_{-\infty}^{\infty}[\hat{f}_n(x,a_n) - \overline{f}(x)]^2 dx$ the following equality

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

$$+((C_1 + C_2)n^{-1/2} + p_1^{\,2} + p_2^{\,2})O(\frac{a_n}{n})$$

is valid, where

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

$$T_i = (a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_i^{(s)}(x)]^2 dx +$$

$$+O(a_n^{-2s}))(C_i n^{-1/2} + p_i^2), \quad i = 1,2$$

and

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

.

Theorem 2. Let in the sequence (6) the controlling sequence $\{\xi_i\}_{i\ge1}$ 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 $\{X_i\}_{i\ge1}$ be the sequence with the chain dependence whose elements are observations of the value X. Let further the conditional distributions $\mathcal{P}_{X_i|\xi_i=b_i}$, $i = 1,2$, have the unknown densities $f_i(x)$, $i = 1,2$, respectively. $f_i(x) \in W_s \cap L_2(-\infty,\infty)$; $k(x) \in H_s \cap L_2(-\infty,\infty)$. Then for any n the estimate of the density $\overline{f}(x) = \pi_1 f_1(x) + \pi_2 f_2(x)$ is $\hat{f}_n(x,a_n)$ and the estimate

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

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

$$+(\frac{1}{n}(C_1(\pi p) + C_2(\pi p)) + \pi_1^{\,2} + \pi_2^{\,2})O(\frac{a_n}{n})$$

holds, where

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

and

$$T_i^* = (a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_i^{(s)}(x)]^2 dx +$$

$$+O(a_n^{-2s}))(\frac{c_i(\pi,p)}{n} + \pi_i^{\,2}), \quad i = 1,2$$

.

Let us proceed to proving the theorems.

Proof of Theorem 1. When fixing the trajectory $\overline{\xi}_{1n}$ we enumerate separately the moments of time at which the first n members the sequence $\{\xi_i\}_{i\ge1}$ take the values b_i, $i = 1,2$, respectively:

$$\tau_0(i) = 0, \quad \tau_m(i) = \min\{j \mid \tau_{m-1} < j \le n; \xi_i = b_i\},$$
$$i = \overline{1,2},\ m = \overline{1,\nu_n(i)}.$$

Thus we obtain the sequence of indexes

$$\tau_1(i), \tau_2(i), \ldots, \tau_{\nu_n(i)}(i), \quad i = \overline{1,2},$$

for which the equalities

$$\xi_{\tau_m(i)} = b_i,\ i = \overline{1,2},\ m = \overline{1,\nu_n(i)},$$

are fulfilled.

For the fixed trajectory $\overline{\xi}_{1n}$ the sum (9) can be decomposed as follows

$$\hat{f}_n(x,a_n) = \frac{\nu_n(1)}{n}\hat{f}_{1n}(x,a_n) + \frac{\nu_n(2)}{n}\hat{f}_{2n}(x,a_n), \quad (10)$$

where

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

It is naturally understood that if $\nu_n(i)=0$, $i=\overline{1,2}$, then the summand $\hat{f}_{in}(x,a_n)$, $i=1,2$, 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 $\overline{\xi}_{1n}$, we represent $E\hat{f}_n(x,a_n)$ in the form of a conditional mathematical expectation

$$E\hat{f}_n(x,a_n)=E\{E(\hat{f}_n(x,a_n)\,|\,\xi_{1n})\}=E\{E(\sum_{i=1}^{2}\frac{\nu_n(i)}{n}\hat{f}_{in}(x,a_n)\,|\,\xi_{1n})\}$$

In view of the fact that the functions $\nu_n(1)$ and $\nu_n(2)$ are measurable with respect to the σ-algebra generated by the partitioning of Ω that results from the fixing of the trajectory $\overline{\xi}_{1n}$, they can be taken out outside the sign of the conditional mathematical expectation.

Here and in what follows it is assumed that the equality

$$E\frac{\nu_n(i)}{n}=p_i \tag{11}$$

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

$$E\left(\frac{\nu_n(i)}{n}\right)^2=D\left(\frac{\nu_n(i)}{n}\right)+\left(E\frac{\nu_n(i)}{n}\right)^2\le n^{-1/2}c_i+p_i^2 \tag{12}$$

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).

$$E\hat{f}_n(x,a_n)=\sum_{i=1}^{2}E\{\frac{\nu_n(i)}{n}E(\frac{a_n}{\nu_n(i)}\sum_{m=1}^{\nu_n(i)}k\left(a_n\left(x-X_{\tau_{m(i)}}\right)\right)\,|\,\xi_{1n})\}=$$

$$=\sum_{i=1}^{2}E\{\frac{\nu_n(i)}{n}E(\frac{a_n}{\nu_n(i)}\nu_n(i)k\left(a_n\left(x-X_{\tau_{m(i)}}\right)\right)\,|\,\xi_{\tau_{m(i)}})\}=\sum_{i=1}^{2}a_n\int_{-\infty}^{\infty}k(a_n(x-u))f_i(u)du\,E\frac{\nu_n(i)}{n}=$$

$$=\sum_{i=1}^{2}p_i\int_{-\infty}^{\infty}k(t)f_i(\frac{t}{a_n}+x)dt$$

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 $\overline{\xi}_{1n}$ the sums $\hat{f}_{1n}(x,a_n)$ and $\hat{f}_{2n}(x,a_n)$ as well as their constituent summands are independent, the following equalities are valid:

$$D\hat{f}_n(x,a_n)=E\{E([\hat{f}_n(x,a_n)-E\hat{f}_n(x,a_n)]^2\,|\,\xi_{1n})\}=$$

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

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

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

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

$$=\sum_{i=1}^{2}E\{(\frac{\nu_n(i)}{n})^2E[\sum_{j=1}^{\nu_n(i)}\frac{a_n}{\nu_n(i)}(k\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)-Ek\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right))]^2\,|\,\xi_{1n})\}=$$

$$=\sum_{i=1}^{2}E\{(\frac{\nu_n(i)}{n})^2(\frac{a_n}{\nu_n(i)})^2E(\sum_{j=1}^{\nu_n(i)}[k\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)-Ek\left(a_n\left(x-X_{\tau_{j(i)}}\right)\right)]^2\mid\xi_{1n})\}=$$

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

Applying (11), we obtain the following expression for $D\hat{f}_n(x,a_n)$

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

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

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

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

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

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

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

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

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

Applying the equality (4) from Lemma 1 we have

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

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

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

Now applying the inequality (12) we complete the estimation of I_1:

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

$$= \frac{a_n}{n}\int_{-\infty}^{+\infty} k^2(x)dx + o(\frac{a_n}{n})\sum_{i=1}^{2}(C_i n^{-1/2} + p_i^2). \tag{14}$$

Let us decompose I_2 from (13):

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

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

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

$$+2(E\frac{\nu_n(1)}{n}\hat{f}_{1n}(x,a_n) - p_1 f_1(x))(E\frac{\nu_n(2)}{n}\hat{f}_{2n}(x,a_n) - p_2 f_2(x))dx \mid \xi_{1n})\} =$$

$$= E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)]^2 dx \mid \xi_{1n})\} + E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]^2 dx \mid \xi_{1n})\} +$$

$$+2E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)][\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]dx \mid \xi_{1n})\} = I_{21} + I_{22} + 2I_{23}$$

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

$$I_{21} = E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - \frac{\nu_n(1)}{n}f_1(x) + \frac{\nu_n(1)}{n}f_1(x) - p_1 f_1(x)]^2 dx \mid \xi_{1n})\} =$$

$$= E\{E(\int_{-\infty}^{\infty}(\frac{\nu_n(1)}{n})^2[E\hat{f}_{1n}(x,a_n) - f_1(x)]^2 dx \mid \xi_{1n})\} + E\{E(\int_{-\infty}^{\infty}(\frac{\nu_n(1)}{n} - p_1)^2 f_1^2(x)dx \mid \xi_{1n})\} +$$

$$+2E\{E(\int_{-\infty}^{\infty}\frac{\nu_n(1)}{n}(\frac{\nu_n(1)}{n} - p_1)[E\hat{f}_{1n}(x,a_n) - f_1(x)]f_1(x)dx \mid \xi_{1n})\} = A_1 + A_2 + A_3$$

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

$$A_1 = E\{(\frac{\nu_n(1)}{n})^2 E(\int_{-\infty}^{\infty}[E\hat{f}_{1n}(x,a_n) - f_1(x)]^2 dx \mid \xi_{1n})\} =$$

$$= E\{(\frac{\nu_n(1)}{n})^2(a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_1^{(s)}(x)]^2 dx + o(a_n^{-2s}))\} =$$

$$\le (a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_1^{(s)}(x)]^2 dx + o(a_n^{-2s}))(C_1 n^{-1/2} + p_1^2) \equiv T_1$$

It is not difficult to obtain the estimate of the sum A_2

$$A_2 = E\{(\frac{\nu_n(1)}{n} - p_1)^2 E\int_{-\infty}^{\infty} f_1^2(x)dx\} \le C_1 n^{-1/2}\int_{-\infty}^{\infty} f_1^2(x)dx$$

We use the fact that the values $\frac{\nu_n(i)}{n} - p_i$, $i = \overline{1,2}$, as well as $\frac{\nu_n(i)}{n}$ are measurable with respect to the σ-algebra generated by the fixed trajectory $\overline{\xi}_{1n}$. Applying Holder's inequality we estimate A_3 as follows

$$A_3 = E\{\frac{\nu_n(1)}{n}(\frac{\nu_n(1)}{n} - p_1)E(\int_{-\infty}^{\infty}[E\hat{f}_{1n}(x,a_n) - f_1(x)]f_1(x)dx \mid \xi_{1n})\} \le$$

$$\le E\{\frac{\nu_n(1)}{n}(\frac{\nu_n(1)}{n} - p_1)E(\sqrt{\int_{-\infty}^{\infty}[E\hat{f}_{1n}(x,a_n) - f_1(x)]^2 dx}\sqrt{\int_{-\infty}^{\infty} f_1^2(x)dx \mid \xi_{1n}})\} \le$$

$$\le \sqrt{a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_1^{(s)}(x)]^2 dx + o(a_n^{-2s})}\sqrt{\int_{-\infty}^{\infty} f_1^2(x)dx} \times E\{\frac{\nu_n(1)}{n}(\frac{\nu_n(1)}{n} - p_1)\} \le$$

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

$$\le \sqrt{a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_1^{(s)}(x)]^2 dx + o(a_n^{-2s})}\sqrt{\int_{-\infty}^{\infty} f_1^2(x)dx} \cdot \sqrt{C_1 n^{-1/2}}\sqrt{C_1 n^{-1/2} + p_1^2} \equiv$$

$$\equiv \sqrt{T_1} \cdot \sqrt{C_1 n^{-1/2}\int_{-\infty}^{\infty} f_1^2(x)dx}$$

Combining the estimates A_1 A_2 and A_3 we obtain

$$I_{21} \le (\sqrt{T_1} + \sqrt{C_1 n^{-1/2}\int_{-\infty}^{\infty} f_1^2(x)dx})^2 \equiv M_1^2, \tag{15}$$

where,

$$M_1 = \sqrt{T_1} + \sqrt{C_1 n^{-1/2}\int_{-\infty}^{\infty} f_1^2(x)dx}.$$

In the same manner it can be shown that

$$I_{22} \le (\sqrt{T_2} + \sqrt{C_2 n^{-1/2}\int_{-\infty}^{\infty} f_2^2(x)dx})^2 \equiv M_2^2, \tag{16}$$

where,

$$T_2 \equiv (a_n^{-2s}\frac{\alpha^2}{(s!)^2}\int_{-\infty}^{\infty}[f_2^{(s)}(x)]^2 dx + o(a_n^{-2s}))(C_2 n^{-1/2} + p_2^2).$$

Let us consider the sum I_{22}. Using Fubini's and Holder's theorems we have

$$I_{23} = E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)][\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]dx \mid \xi_{1n})\} \le$$

$$\le E\{E\sqrt{\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)]^2 dx}\sqrt{\int_{-\infty}^{\infty}[\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]^2 dx \mid \xi_{1n})\}} \le$$

$$\le E\{\sqrt{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)]^2 dx \mid \xi_{1n})}\sqrt{E(\int_{-\infty}^{\infty}[\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]^2 dx \mid \xi_{1n})\}} \le$$

$$\le \sqrt{E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(1)}{n}E\hat{f}_{1n}(x,a_n) - p_1 f_1(x)]^2 dx \mid \xi_{1n})\}}\sqrt{E\{E(\int_{-\infty}^{\infty}[\frac{\nu_n(2)}{n}E\hat{f}_{2n}(x,a_n) - p_2 f_2(x)]^2 dx \mid \xi_{1n})\}}$$

Each of the obtained two factors is estimated as the summands I_{21} and I_{22}. We obtain

$$I_{23} \le M_1 M_2 \tag{17}$$

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 $\overline{\xi}_{1n}$ we denote by the symbols $\nu_n(i)$ $i = 1,2$, the frequencies of the chain being is 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

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

Hence there holds the following estimate

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

which we use instead of the estimate (9).

Example 1.

We may have one such interpretation of a conditional independent sequence

Consider $\{\xi_i\}_{i\ge1}$ the sequence independent discrete randomly distributed values with the $\xi_i : \Omega \to \{b_1, b_2, ..., b_r\}$, $\xi_i = \sum_{i=1}^{r} b_i I_{(\xi_1 = b_i)}$ of distribution $P(\xi_1 = b_i) = p_i$, $i = \overline{1,r}$, $p_1 + p_2 + ... + p_r = 1$.

Consider the r number of sequences $\{X_i^\alpha\}_{i\ge1}$, $\alpha = \overline{1,r}$ of independent random values. Assume that the members of the corresponding sequence of each α index have the same $\mathcal{P}_\alpha$, distribution $\alpha = \overline{1,r}$. We have r sequences.

$$X_1^1, X_2^1, X_3^1, ..., X_n^1, ...$$

$$X_1^2, X_2^2, X_3^2, ..., X_n^2, ...$$

....

$$X_1^r, X_2^r, X_3^r, ..., X_n^r, ...$$

At each $i-$ step, if each $i-$ member of the sequence gets a value of b_α $(1 \le \alpha \le r)$, then take the X_i^α vector $\alpha = \overline{1,r}$ from the column of ($X_i^1, X_i^2, ,..., X_i^r)^T$ 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” $\overline{\xi}_{1n} = (\xi_1, \xi_2, .., \xi_n)$.

If $\{\xi_i\}_{i\ge1}$ is a finite chain of Markov with $\{b_1, b_2, ..., b_r\}$ set of states, than

$$Y_i = \sum_{\alpha=1}^{r} X_i^{\alpha} I_{(\xi_i=\alpha)},$$

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

$$\{\xi_i X_i\}_{i\ge1} \qquad (18).$$

where $\{\xi_i\}_{i\ge1}$ are the days of observations, and $\{X_i\}_{i\ge1}$ 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 $p(\xi_i = 1) = \frac{1}{4}$, $p(\xi_i = 2) = \frac{3}{4}$. 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 $\bar{f}(x) = \frac{1}{4} f_1(x) + \frac{1}{4} f_2(x)$ density estimation let us use the core of Bartlett $\bar{k}(x) = \frac{3}{4}(1-x^2)\mathrm{I}_{|x|\le1}$. It is clear $\bar{k}(x) \in H_2$. $f_1(x)$ and $f_2(x)$ normal distribution densities so obviously $f_i(x)$ has a continuous derivative of any order. So, we can say that $f_i(x) \in \mathrm{W}_2 \cap L_2(-\infty,\infty)$. In the role of $\{a_n\}_{n\ge1}$ sequence we can suppose the sequence $a_n = \sqrt{n}$.

We will get the estimation of $\bar{f}(x)$ as the result.

$$\hat{f}_n(x) = \frac{3}{4\sqrt{n}} \sum_{j=1}^{n} (1 - [\sqrt{n}(x - X_j)]^2)\, \mathrm{I}_{(|(x-X_i)| \le \frac{1}{\sqrt{n}})}. \quad (19)$$

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:

$$\hat{f}_4(x) = \begin{cases} -1,5x^2 + 3x - 1,125) & \text{when } x \in [0,5;0,7); \\ -3x^2 + 6,6x - 2,91) & \text{when } x \in [0,7;1,5]; \\ -1,5x^2 + 3,6x - 1,785) & \text{when } x \in (1,5;1,7]; \\ -1,5x^2 + 8,4x - 11,385) & \text{when } x \in [2,3;2,7); \\ -3x^2 + 1,8x - 26,37) & \text{when } x \in [2,7;3,3]; \\ -1,5x^2 + 9,6x - 14,985) & \text{when } x \in (3,3;3,7]; \\ 0 & \text{when } x \in (-\infty;0,5) \cup (1,7;2,3) \cup (3,7;\infty); \end{cases},$$

The price distribution function will be $F(x) = \int_{-\infty}^{x} f_4(x)dx$.

4. Discussion

$\frac{\nu_n(i)}{n}$ and $\nu_n(i)$ 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 $\bar{\xi}_{1n} = (\xi_1, \xi_2, ..., \xi_n)$ trajectory and the product of these multiples. It has been shown that $\nu_n(i)$ and $f(\nu_n(i))$ functions (when $f \in L_2(-\infty,\infty)$ are measures toward σ-algebra induced by the division of the Ω space generated by the fixation of the ξ_{1n} 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 Helder 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-Parten 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. Dyorfi (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 Sequenges without Mixing Assumptions. Systems and Industrial Engineering Department. University of Arizona, Tucson, Arisona 85721–JOURNAL OF MULTIVARIATE ANALYSIS 30, p. 124-136. 1989;

[12] Bokuchava I. V., Kvatadze 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] Kvatadze Z., Kvatadze 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 Ivane Javakhisvili Tbilisi State University (TSU). April 23-25, 2019. Book of Abstracts. p. 82.