Association For Science

Georgian Scientists

ქართველი მეცნიერები

Vol. 5 Issue 3, 2023

https://doi.org/10.52340/gs.2023.05.03.03


# On the conditional distribution of the sum of $m$ –dependent vectors

**Zurab Kvatadze *, Beqnu Parjiani** and TSiala Kvatadze *****

* Professor of the Department of Mathematics of Georgian Technical University; ** Associate Professor of Department of Mathematics of Georgian Technical University; * Invited Lecturer in Statistics, International Black Sea University

**Abstract**

On the probability space $(\Omega, F, P)$ is considered a stationary (in the narrow sense) two-component sequence $\{\xi_i, Y_i\}_{i\ge1}$. $Y_i : \Omega \to R^k$ is conditionally $m$ –independent vectors sequence The decomposition $S_n = \sqrt{\frac{1}{n}}\sum_{i=1}^{n}[Y_i - EY_1] = S_{n1} + S_{n2}$ is applied, where $S_{n1} = \sqrt{\frac{1}{n}}\sum_{i=1}^{n}\left[Y_i - E\left(Y_i \middle| \xi_i\right)\right]$ and $S_{n2} = \sqrt{\frac{1}{n}}\sum_{i=1}^{n}\left[E\left(Y_i \middle| \xi_i\right) - EY_1\right]$. Is proved that if $F_{S_{n2}}(\cdot)\xrightarrow[n\to\infty]{}Q(\cdot)$ and $Q(\cdot)$ is nondegenerate distribution, then for each $x, y \in R^k$ $P\left(\Phi_{R_m}(x-y) \le F_{S_n|\bar{\xi}_{1n}}(x) \le \Phi_{R_m}(x+y)\right)\xrightarrow[n\to\infty]{}Q(y)-Q(-y)$, where $R_m = R_0^{(0)} + \sum_{p=1}^{m}\left[R_0^{(p)} + \left(R_0^{(p)}\right)^T\right]$, $R_0^{(p)} = E\operatorname{cov}\left(Y_1, Y_{1+p} \middle| \bar{\xi}_{1n}\right)$, $\bar{\xi}_{1n} = (\xi_1, \xi_2, ..., \xi_n)$ is fixed trajectory of control sequence. When $\{\xi_i\}_{i\ge1}$ is a finite ergodic Markov's chain with one class of ergodicity, then is shownthat $Q(\cdot) = R_{T_\mu}(\cdot)$, where $T_\mu$ is expressed by chain parameters. When a chain has cyclic subclasses, the existence of the corresponding limiting characteristics is understood in the sense of Cesaro.

When solving many practical problems, the question arises of constructing statistical estimates of parameters by dependent observations. For example, in a study conducted to determine the rating model of vocational colleges, the data obtained from a survey of students are dependent on each other, due to the fact that they are obtained from the same social group. (R. Chartolani, N. Durglishvili, Z. Kvatadze. Optimization of a State Financing Model of Vocational Colleges. Proc. A. Razmadze Math. Inst. 2015. 169.(2-15), pp. 23-31). Also, the dependence of data in various studies of the geophysical direction cannot be neglected. For example, in the pre-earthquake (preparatory) period, tectonic processes in rocks develop continuously in time, and therefore records of any characteristic taken at discrete points in time depend on each other as manifestations of one whole. Proceeding from similar considerations, when determining the law of distribution of various populations, density estimates are constructed by dependent observations. For example, nonparametric density estimates and estimates of regression coefficients are known, built by observations connected in a Markov chain (Yakowitz Sidney (1989) Nonparametric density and regression estimation for Markov sequences without mixing assumptions. 85721–Journal of Multivariate Analysis, 30: 124-136. Arisona, USA). Conditionally independent observations and observations with a chain dependence are also considered. The accuracies of density estimates constructed by such observations by the $L_2$ metric (Z, Kvatadze, B. Phardjiani. On the Exsactness of Distribution Density Estimates Constructed by Some Class of Dependent Observations. Mathematics and Statistics. 2019 Vol. 7(4), pp. 135-145. SAN JOSE) and by the $L_1$ metric (B. Parjiani, L. Labadze, T. Kvatadze; Georgian Scientists, "On the accuracyby the metric L1 of the density estimation constructed by dependent observations" Vol. 5. Issue 1, pp. 308-321, 2023) are known.

To construct statistical estimates based on dependent observations and determine their unbiasedness, it is necessary to know the asymptotic distribution of sums of dependent random variables. At the present stage, the rich theory of summation of independent random variables (Normal Approximation Some Recent Advances. Sazonov V.V. Lecture Notes in Math. V. 79, Berlin, ete.,: Springer. 1981.) is transferred to dependent random variables. In many problems, Markov dependence is used, which is one of the types of weak dependence. Questions of the limiting asymptotic behavior of sums of random variables of various types of dependency (weakly dependent, conditionally independent, connected in a Markov chain) are discussed in many papers. Sums of such random variables are often considered, the joint distribution of which is determined by some control sequence of random elements. Conditionally independent sequences (Bokuchava I. V. Limit theorems for conditionally independent sequences. (in Russian) Teor. Verojatnost. i Primenen. XXIX. (1984). №1, p. 192-193) and sequences with chain dependence are considered. For example, articles by O'Brien (O' Braien G.L. Limit Theorems for Sum of Chain Dependent Proccesses. U. Appl. Probab., 1974, 11, 582-587), Y. Aleshkevichus (G. YU. Aleshkyavichus, On the central limit theorem for sums of random variables given on a Markov chain. (Russian) Lithuanian Mathematical Collected Works, Vilnius 6 (1966), №. 1, 15–22.), R. Chitashvili, T. Shervashidze, I. Bokuchava, Z. Kvatadze (Bokuchava I., Kvatadze Z., Shervashidze T. *On Limit theorem for random vectors controlled by a Marcov chain*. Prob. Theory and Math. Stat., Vol. 1, 1986, 239-250. VNU Science Press, Utrecht.) and other authors deal with these issues.In this paper, we discuss the class of sequences of conditionally $m$-independent vectors and the class of sequences of vectors with chain $m$-dependence. (Kvatadze Z., ShervashidzeT. *Some limit theorems for I.I.D. and Conditionally independent random variables*. The second international Conference, "Problems of Cybernetics and Informatics". September 10-12, 2008. Baku. Azerbaijan. Section № 4. "Applied Stochastic Analysis". *Institute of Information Technologies of NASA*. Printing House of "Information Technology" Baku. 2008. Vol. II, 217-219). In the case of a non-regular ergodic chain, the sequence of functions is defined

on a sequence of vectors with a chain $m$-dependence. A representation in matrix form of the limit covariance matrix of the normalized sum of the sequence of these functions is obtained. The fundamental matrix and the corresponding limiting characteristics of the Markov chain in the case of cyclic subclasses are computed using the Cesaro summation.

The Levy-Prokhorov metric $\rho(F,G)=\sup_{x\in R^k}|F(x)-G(x)|$ is often used when considering the convergence of distributions on the space of distribution functions. Therefore, the question of determining the limiting probability of the distribution of the sum of functions falling into a strip consisting of two normal shifted distributions is of natural interest. The article discusses the normalized sum of a sequence of conditionally $m$-dependent vectors. The limiting probability of the conditional distribution of this sum falling into the band consisting of two shifted normal distributions is determined. The covariance matrix of these distributions is expressed by the Markov chain parameters. A similar limiting probability is set for the normalized sum of a sequence of vectors with a chain $m$-dependence.

**Keywords**: Conditionally $m$–independent sequence, conditional distribution, Markov chain, $m$-dependent sequence controlled by the Markov chain.

2010 Mathematics Subject Classification. 60F05. 60J10. 60J20.

## Introduction

When solving practical problems, it is often necessary to build statistical estimates based on dependent observations. For example, time series used in psychological research or in economic parameter estimations, in most cases, consist of dependent data. Determining the accuracy and unbiasedness of the constructed estimates is related to the knowledge of the asymptotic distribution of the sums of dependent random variables. There is a spread of methods for studying the distribution of sums of independent random variables ([1]) over dependent random variables. The Markov dependence, which is one of the types of weak dependence, naturally enters this path. Limit theorems for weakly dependent sequences are expressed in terms of sigma-algebras generated by asymptotically separable segments of the sequence. An essential role in their approval is played by Bernstein's "sectioning" method, which uses the effect of dependence weakening with increasing distance between segments. Many authors study sums of random variables whose distribution is determined by some "control" sequence of random elements. Conditionally independent sequences ([2],[3]) and sequences with chain dependence ([4],[5]) are considered. In this paper, we consider a conditionally $m$-independent sequence of vectors and a sequence of vectors with a chain $m$-dependence. Let's calculate the limiting probability of their conditional distribution falling into the band created from two shifted normal distributions.

On the probability space $(\Omega, F, P)$ we consider a two-component stationary (in the narrow sense) sequence

$$\{\xi_i, Y_i\}_{i\ge1} \qquad (1)$$

$\{\xi_i\}_{i\ge1}$, $(\xi_i : \Omega \to \Xi)$ is the control sequence. Assume that the set $\Xi$ consists of a finite number of finite elements. $\{Y_i\}_{i\ge1}$, $\left(Y_i : \Omega \to R^k\right)$ is a sequence of conditionally $m$–independent vectors.

**Definition.** (see. [6]) The sequence $\{Y_i\}_{i\ge1}$ from (1) is called a conditionally $m$ –independent sequence if the vectors $Y_1, Y_2, \dots, Y_n$ on a fixed trajectory $\bar{\xi}_{1n} = (\xi_1, \xi_2, .., \xi_n)$ for an arbitrary natural number $n$ become independent when thier index difference exceeds $m$. In this case the distribution of $Y_i$ is depends only on $\xi_i$. For arbitrary natural numbers $i$, $l$, $n$, $j_1, j_2, \dots, j_l$, $(2 \le l \le n;\ i \le n;\ 1 \le j_1 < j_2 < \dots < j_l \le n)$ the equations are fulfilled:

$$\mathcal{P}_{(Y_{j_1}, Y_{j_2}, \dots, Y_{j_l})|\bar{\xi}_{1n}} = \mathcal{P}_{Y_{j_1}|\xi_{j_1}} * \mathcal{P}_{Y_{j_2}|\xi_{j_2}} * \dots * \mathcal{P}_{Y_{j_l}|\xi_{j_l}}, \qquad \text{თუ} \quad \left|j_p - j_q\right| > m,\ \forall\, p, q \in \{1, 2, \dots, l\}$$

$$\mathcal{P}_{Y_i|\bar{\xi}_{1n}} = \mathcal{P}_{Y_i|\xi_i}, \qquad 1 < j_1 < j_2 < \dots < j_l < n, \quad i = \overline{1, n},$$

where $\mathcal{P}_X$ is the distribution of $X$.

**Remark:** If $\{\xi_i\}_{i\ge1}$ is a Markov chain with discrete time, then the sequence $\{Y_i\}_{i\ge1}$ is called a conditionally $m$-dependent sequence controlled by the Markov chain (sequence with chain $m$-dependence).

Let us give auxiliary lemmas.

Let's introduce notations :

$$\mu(\xi_j) = E\left(Y_j \left|\xi_j\right.\right), \qquad \mu = E\mu(\xi_1) = EY_1$$

$$R\left(\xi_j, \xi_r\right) = E\left\{\left[Y_j - \mu\left(\xi_j\right)\right]\left[Y_r - \mu\left(\xi_r\right)\right]^T \left|\xi_{1n}\right.\right\}, \qquad 1 \le j, r \le n$$

$$R_0^{(l)} = ER\left(\xi_1, \xi_{1+l}\right), \qquad R_0^{(-l)} = ER\left(\xi_{1+l}, \xi_1\right), \qquad l = 0, \dots, m$$

$$R_m = \sum_{l=-m}^{m} R_0^{(l)} = R_0^{(0)} + \sum_{l=1}^{m}\left[R_0^{(l)} + \left(R_0^{(l)}\right)^T\right], \qquad (1)$$

Consider sum $S_n = \sqrt{\frac{1}{n}}\sum_{i=1}^{n}\left[Y_i - \mu\right]$. Let us use the same method that was used to determine the limit distribution of the sum $S_n$ by fixing the trajectories $\bar{\xi}_{1n}$ in [5]. Let us decompose the sum $S_n$ on a fixed trajectory into two uncorrelated and asymptotically independent centered sums $S_n = S_{n1} + S_{n2}$:

$$S_{n1} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left[Y_j - \mu\left(\xi_j\right)\right], \qquad S_{n2} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\mu\left(\xi_j\right) - \mu\right].$$

**Lemma 1.** (see [5]) Suppose that $\{Y_i\}_{i\ge1}$ is a conditionally $m$–independent sequence in model (1). Let's say for arbitrary function $\Psi : \Xi \to R^k$ when $E\Psi_p\left(\xi_1\right) < \infty$ $\left(p = \overline{1, k}\right.$, $\Psi(\cdot) = \left(\Psi_1(\cdot), \Psi_2(\cdot), \dots, \Psi_k(\cdot)\right)\Big)$ almost everywhere an convergence is executed

$$\frac{1}{n}\sum_{j=1}^{n}\Psi\left(\xi_j\right)\xrightarrow{n\to\infty}E\Psi\left(\xi_1\right) \qquad \text{a. e.} \tag{2}$$

Let's say

$$sp(R_m)<\infty \tag{3}$$

Then, for $n\to\infty$, the following convergences are fulfilled:

a) $\mathcal{P}_{S_{n1}|\bar{\xi}_{1n}}\xrightarrow{W}\Phi_{R_m}$ a. e.,

b) $\mathcal{P}_{S_{n1}}\xrightarrow{W}\Phi_{R_m}$,

c) If $\mathcal{P}_{S_{n2}}\xrightarrow{W}\mathcal{P}$ then $\mathcal{P}_{S_{n1}}\xrightarrow{W}\Phi_{R_m}*\mathcal{P}$

**Lemma 2.** (see. [7]) Let's say $g:T\to T'$ is a continuous mapping of a $T$ metric space into a $T'$ metric space. If $X_n\xrightarrow{d}X$ when $n\to\infty$ then $g(X_n)\xrightarrow{d}g(X)$.

**Lemma 3.** (see. [7]) Assume that, $X_n$, $Z_n$ and $X$ are random variables with values in a separable metric space $T$. Let's say when $n\to\infty$ the limit equalities $X_n\xrightarrow{d}X$ and $\rho\left(X_n,Z_n\right)\to 0$ are satisfied, then $Z_n\xrightarrow{d}X$.

## Methodology

When proving the theorem, we use the methods applicable in article [5] by I. Bokuchava, T. Shervashidze and Z. Kvatadze. The sums $S_{n1}$ and $S_{n2}$ are uncorrelated and asymptotically independent (see. [5]). For calculated expressions in inequalities, we use the following representation $E\left(\cdot\right)=E\left\{E\left(\cdot\middle|\bar{\xi}_{1n}\right)\right\}$. This representation allows us to consider the terms of these sums on a fixed trajectory as independent quantities. We are investigating an ergodic chain that is not regular. Since there are cyclic subclasses, the existence of the fundamental matrix of a chain is understood in the sense of Cesaro (see.[8]). In [9], the representation of the covariance matrix of the limit distribution of the normalized sum of functions defined on a chain was obtained using the characteristics of the chain.

## Main Results

**Theorem 1.** Suppose that $\{Y_i\}_{i\geq 1}$ is a conditionally $m$–dependent sequence in model (1). Suppose the control sequence $\{\xi_i\}_{i\geq 1}$ satisfies condition (2) and inequality (3) is fulfilled. Then the following propositions are true:

a) Let's say $\mathcal{P}_{S_{n2}}$ is weakly converges to some non degenerate distribution $Q$ with distribution function $Q(\cdot)$, then for arbitrary vectors $x,y$, $\left(x,y\in R^k\right)$ when $n\to\infty$ there is a convergence

$$\mathrm{P}\left\{\Phi_{R_m}(x-y)\leq F_{S_n|\bar{\xi}_{1n}}(x)\leq\Phi_{R_m}(x+y)\right\}\to Q(y)-Q(-y) \tag{4}$$

b) The opposite is true, if (4) is true, then $\mathcal{P}_{S_{n2}}\xrightarrow{W}\mathcal{Q}$.

**Proof.** Let's use the following representations

$$\mathrm{P}\left\{\Phi_{R_m}(x-y)\le F_{S_n|\bar{\xi}_{1n}}(x)\le\Phi_{R_m}(x+y)\right\}=\mathrm{P}\left\{\Phi_{R_m}(x-y)\le F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})\le\Phi_{R_m}(x+y)\right\}=$$
$$=\mathrm{P}\left\{F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})\le\Phi_{R_m}(x+y)\right\}-\mathrm{P}\left\{F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})\ge\Phi_{R_m}(x-y)\right\}-1\equiv$$
$$=\mathrm{P}\{M_n\le 0\}-\mathrm{P}\{L_n\ge 0\}-1 \qquad (5)$$
$$\mathrm{P}\left\{\Phi_{R_m}(x-y)\le\Phi_{R_m}(x-S_{n2})\le\Phi_{R_m}(x+y)\right\}\equiv\mathrm{P}\left\{M_n^1\le 0\right\}-\mathrm{P}\left\{L_n^1\ge 0\right\}-1 \qquad (6)$$

where

$$M_n=F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})-\Phi_{R_m}(x+y);$$
$$M_n^1=\Phi_{R_m}(x-S_{n2})-\Phi_{R_m}(x+y);$$
$$L_n=F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})-\Phi_{R_m}(x-y);$$
$$L_n^1=\left\{\Phi_{R_m}(x-S_{n2})-\Phi_{R_m}(x-y)\right\}.$$

On the space where the sequence (1) is given, let's determine the random variable $\eta$ with distribution $Q$. Due the continuity and monotonicity of the function $\Phi_{R_m}(\cdot)$

$$\mathrm{P}\left\{\Phi_{R_m}(x-y)\le\Phi_{R_m}(x-S_{n2})\le\Phi_{R_m}(x+y)\right\}=$$
$$=\mathrm{P}\{-y\le S_{n2}\le y\}\xrightarrow{n\to\infty}\mathrm{P}\{-y\le\eta\le y\}=Q(y)-Q(-y)\ . \qquad (7)$$

Let us show that the quantities $M_n$ and $L_n$ have the same limiting distributions as the quantities $M_n^1$ and $L_n^1$ (respectively).

$\Phi_{R_m}(x-t)$ is continuous with respect to $t$, and $\mathcal{P}_{S_{n2}}\xrightarrow{W}\mathcal{P}_\eta$, therefore, by virtue of Lemma 2, when $n\to\infty$, the following convergences are fulfilled

$$M_n^1=\Phi_{R_m}(x-S_{n2})-\Phi_{R_m}(x+y)\xrightarrow{W}\Phi_{R_m}(x-\eta)-\Phi_{R_m}(x+y),$$
$$L_n^1=\Phi_{R_m}(x-S_{n2})-\Phi_{R_m}(x-y)\xrightarrow{W}\Phi_{R_m}(x-\eta)-\Phi_{R_m}(x-y).$$

Let us introduce the distance between the functions $f(\cdot)$ and $g(\cdot)$ as following:

$$\rho\left(f(\cdot),g(\cdot)\right)=\sup_{x\in R^k}\left|f(x)-g(x)\right|.$$

It is clear that

On a fixed trajectory $\bar{\xi}_{1n}=(\xi_1,\xi_2,..,\xi_n)$, the sum $S_{n2}$ becomes equal to some specific (trajectory dependent) number. Therefore, it is clear that the following equations hold

$$\rho\left(M_n,\ M_n^1\right)=\rho\left(L_n,\ L_n^1\right)=$$
$$=\sup_{x\in R^k}\left|F_{S_{n1}|\bar{\xi}_{1n}}(x-S_{n2})-\Phi_{R_m}(x-S_{n2})\right|=\sup_{x\in R^k}\left|F_{S_{n1}|\bar{\xi}_{1n}}(x)-\Phi_{R_m}(x)\right|.$$

By virtue of Lemma 1, when $n\to\infty$ we have

$$\sup_{x\in R^k}\left|F_{S_{n1}|\bar{\xi}_{1n}}(x)-\Phi_{R_m}(x)\right|\to 0 \qquad \text{a. e.}$$

Therefore

$$\rho\left(M_n,\ \mathrm{M}_n^1\right)\xrightarrow{P}0,\qquad \rho\left(L_n,\ L_n^1\right)\xrightarrow{P}0.$$

By virtue of Lemma 3, we will have

$$M_n\xrightarrow{W}\Phi_{R_m}(x-\eta)-\Phi_{R_m}(x+y),$$

$$L_n \xrightarrow{W} \Phi_{R_m}(x-\eta) - \Phi_{R_m}(x-y).$$

We have obtained that the random vectors $M_n$ and $L_n$ have the same limiting distributions as the vectors $M_n^1$ and $L_n^1$. According to equalities (5) and (6) we will have

$$\lim_{n\to\infty} \mathrm{P}\left\{\Phi_{R_m}(x-y) \le F_{S_n|\bar{\xi}_{1n}}(x) \le \Phi_{R_m}(x+y)\right\} = \lim_{n\to\infty}\left(\mathrm{P}\{M_n \le 0\} + \mathrm{P}\{L_n \ge 0\} - 1\right) =$$

$$= \lim_{n\to\infty}\left(\mathrm{P}\left\{M_n^1 \le 0\right\} + \mathrm{P}\left\{L_n^1 \ge 0\right\} - 1\right) = \lim_{n\to\infty} \mathrm{P}\left\{\Phi_{R_m}(x-y) \le \Phi_{R_m}(x-S_{n2}) \le \Phi_{R_m}(x+y)\right\} =$$

$$= Q(y) - Q(-y)$$

The proof of point b) is obtained from the fact that the random vectors $M_n$ and $L_n$ have the same limiting distributions as the vectors $M_n^1$ and $L_n^1$ respectively. According to equalities (5) and (6) we will have

$$\lim_{n\to\infty} \mathrm{P}\left\{\Phi_{R_m}(x-y) \le \Phi_{R_m}(x-S_{n2}) \le \Phi_{R_m}(x+y)\right\} \equiv$$

$$\equiv \lim_{n\to\infty}\left(\mathrm{P}\left\{M_n^1 \le 0\right\} + \mathrm{P}\left\{L_n^1 \ge 0\right\} - 1\right) = \lim_{n\to\infty}\left(\mathrm{P}\{M_n \le 0\} + \mathrm{P}\{L_n \ge 0\} - 1\right) =$$

$$= \lim_{n\to\infty} \mathrm{P}\left\{\Phi_{R_m}(x-y) \le F_{S_n|\bar{\xi}_{1n}}(x) \le \Phi_{R_m}(x+y)\right\} \to Q(y) - Q(-y).$$

On the other hand

$$\mathrm{P}\left\{\Phi_{R_m}(x-y) \le \Phi_{R_m}(x-S_{n2}) \le \Phi_{R_m}(x+y)\right\} \equiv \mathrm{P}\left\{-y \le S_{n2} \le y\right\}$$

from this we will obtain b).

Let's suppose that $\{\xi_n\}_{n\ge1}$ is a finite, homogeneous, stationary, ergodic Markov chain with one class of ergodicity, which may contain cyclic subclasses. Let's say the set of states of the Markov chain is $\Xi = \{b_1, b_2, ..., b_r\}$. The limiting stationary distribution is $\pi = (\pi_1, \pi_2, ..., \pi_r)$.

Consider a function $f : \Xi \to R^k$ defined on a chain that satisfies the conditions of Lemma 1. In this case, it is known (see [8]) that for the sum $U_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}[f(\xi_j) - Ef(\xi_j)]$ when $n \to \infty$ thefollowing convergences take place

$$\operatorname{cov}(U_n) \to T_f,$$

$$\mathcal{P}_{U_n} \xrightarrow{W} \Phi_{T_f}.$$

Matrix $T_f = \left(t_{f_{i,j}}\right)_{i,j=\overline{1,k}}$ elements are expressed by Markov chain parameters

$$t_{f_{i,j}} = \sum_{\alpha,\beta}^{r}(\pi_\alpha z_{\alpha\beta} + \pi_\beta z_{\beta\alpha} - \pi_\alpha\pi_\beta - \pi_\alpha\delta_{\alpha\beta}) f_i(\alpha) f_j(\beta),\ i, j = \overline{1,k}. \quad (8)$$

where $Z$ is the fundamental matrix of the chain

$$Z = [I - (P - \Pi)]^{-1} = I + \left(\sum_{j=1}^{\infty}(P^j - \Pi)\right)_c = \left\|z_{\alpha\beta}\right\|_{\alpha,\beta=\overline{1,r}},$$

$$\Pi = \left\|\pi_{\alpha\beta}\right\|_{\alpha,\beta=\overline{1,r}},\quad \pi_{\alpha\beta} = \pi_\beta,\quad \alpha,\beta = \overline{1,r},$$

and

$$f=(f_1, f_2,..., f_k), \quad f_i(\alpha)\equiv f_i(b_\alpha),\ i=\overline{1,k},\ \alpha=\overline{1,r}.$$

The symbol $(\cdot)_c$ stands for convergence according to Cesaro. It was shown in [8] that when the chain is irregular (contains cyclic subclasses), then both existence $Z$ and convergence $\operatorname{cov}(U_n)\to T_f$ are understood in the sense of Cesaro.

In [9], the representation of the matrix $T_f$ is accepted in the matrix form

$$T_f = F[\Pi_{dg}Z+(\Pi_{dg}Z)^T-\Pi_{dg}\Pi-\Pi_{dg}]F^T, \qquad (9)$$

where $F$ is the following matrix of order $k\times r$.

$$F=\begin{pmatrix} f_1(b_1),\ f_1(b_2),...,\ f_1(b_r) \\ f_2(b_1),\ f_2(b_2),...,\ f_2(b_r) \\ \dots\dots\dots\dots\dots\dots \\ f_k(b_1),\ f_k(b_2),...,\ f_k(b_r)\end{pmatrix}=\begin{pmatrix} f_1(1),\ f_1(2),...,\ f_1(r) \\ f_2(1),\ f_2(2),...,\ f_2(r) \\ \dots\dots\dots\dots\dots\dots \\ f_k(1),\ f_k(2),...,\ f_k(r)\end{pmatrix},$$

$$\lim_{n\to\infty}(P^n)_c=\Pi=\begin{pmatrix}\pi_1,\pi_2,...,\pi_r \\ \pi_1,\pi_2,...,\pi_r \\ \dots\dots\dots \\ \pi_1,\pi_2,...,\pi_r\end{pmatrix}.$$

Let's use this fact. Consider, as a function $f:\Xi\to R^k$, the function

$$\mu:\Xi\to R^k.\ f(\xi_j)=\mu(\xi_j)=E(Y_j|\xi_j)=(\mu_1(\xi_j),\mu_2(\xi_j),...,\mu_k(\xi_j)),$$

$$f_p(\xi_j)=\mu_p(\xi_j)=E(Y_j^p|\xi_j),\ p=\overline{1,k},\ Y_j=(Y_j^1,Y_j^2,...,Y_j^k),$$

$$f(\alpha)=\mu(\alpha)=E(Y_j|\xi_j=\alpha),\ \alpha=\overline{1,r},$$

$$f_p(\alpha)=\mu_p(\alpha)=E(Y_j^p|\xi_j=\alpha),\ p=\overline{1,k},\quad \alpha=\overline{1,r}.$$

It is clear that

$$\mu=E\{\mu(\xi_j)\}=E\{E(Y_j|\xi_j)\}=\sum_{\alpha=1}^{r}\pi_\alpha\mu(\alpha).$$

Accordingly, as a the sum $U_n$ will be considered the sum $S_{n2}=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[\mu(\xi_j)-\mu]$.

We will have the representation obtained from (9)

$$T_\mu=\bar{\mu}\cdot[\Pi_{dg}Z+(\Pi_{dg}Z)^T-\Pi_{dg}\Pi-\Pi_{dg}]\cdot\bar{\mu}^T, \qquad (10)$$

where

$$\overline{\mu}=\begin{pmatrix}\mu_1(b_1),\mu_1(b_2),\ldots,\mu_1(b_r)\\ \mu_2(b_1),\mu_2(b_2),\ldots,\mu_2(b_r)\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \mu_k(b_1),\mu_k(b_2),\ldots,\mu_k(b_r)\end{pmatrix}\equiv\begin{pmatrix}\mu_1(1),\mu_1(2),\ldots,\mu_1(r)\\ \mu_2(1),\mu_2(2),\ldots,\mu_2(r)\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \mu_k(1),\mu_k(2),\ldots,\mu_k(r)\end{pmatrix},$$

Accordingly $T_\mu=\left(t_{\mu_{i,j}}\right)_{i,j=\overline{1,k}}$, where

$$t_{\mu_{i,j}}=\sum_{\alpha,\beta}^{r}(\pi_\alpha z_{\alpha\beta}+\pi_\beta z_{\beta\alpha}-\pi_\alpha\pi_\beta-\pi_\alpha\delta_{\alpha\beta})\mu_i(\alpha)\mu_j(\beta)\,,\quad i,j=\overline{1,k}\,. \quad (11)$$

And $\delta_{\alpha\beta}$ is the Kronecker symbol.

**Theorem 2.** Suppose that $\{Y_i\}_{i\geq 1}$ is a sequence with a chain $m$-dependence in model (1). Suppose the control sequence $\{\xi_i\}_{i\geq 1}$ satisfies condition (2) and inequality (3) is fulfilled. Suppose $\{\xi_n\}_{n\geq 1}$ is a finite, homogeneous, stationary, ergodic Markov chain with one ergodicity class, which may contain cyclic subclasses. Let's say the set of states of the Markov chain is $\Xi=\{b_1,b_2,\ldots,b_r\}$. The limiting stationary distribution is $\pi=(\pi_1,\pi_2,\ldots,\pi_r)$. Then in the above notation for any vectors $\left(x,y\in R^k\right)$ when $n\to\infty$ occurs the convergence $x,y$

$$\mathrm{P}\left\{\Phi_{R_m}(x-y)\leq F_{S_n|\bar{\xi}_{1n}}(x)\leq\Phi_{R_m}(x+y)\right\}\to\Phi_{T_\mu}(y)-\Phi_{T_\mu}(-y) \quad (12)$$

**Proof.** The proof of the theorem is carried out similarly to Theorem 1. It is only necessary to take into account that the conditions of Theorem 2 include the conditions of Theorem 1. At the moment, it is known that the limit distribution of the sum $S_{n2}$ is expressed by the parameters of the Markov chain. In particular, its limiting covariance matrix is established. It can be seen from (10) that at $n\to\infty$

$$\mathcal{P}_{S_{n2}}\xrightarrow{W}\Phi_{T_\mu}.$$

This means that instead of a distribution of $Q(\cdot)$, we can consider a distribution of $\Phi_{T_\mu}(\cdot)$ . The form of the matrix $T_\mu$ in the presence of cyclic subclasses is obtained with the help of Cesaro convergence.

**Discussion**

In the particular case when the Markov chain is regular (does not contain cyclic subclasses) and $\{Y_n\}_{n\geq 1}$ are one-dimensional ($k=1$) discrete random variables, then an analogue of (11) was obtained in [10]. From here we can obtain (11) by the Cramer-Wold method. However, this does not give us the matrix representation (10). In [8], the case $k=1$ was considered, when the Markov chain already contains cyclic subclasses. In [9], a matrix representation (9) was obtained in the general case when $k > 1$ and the chain has cyclic subclasses. In this case, as in [8], the result is obtained using Cesàro convergence. An analog of Theorem 2 was obtained in [11] for one-dimensional integer random variables connected by a regular Markov chain. This result was generalized in [12] by partially removing the restriction on random variables. Conditionally $m$-independent and chain $m$-dependent one-dimensional random variables are discussed in [12]. The case when the chain has no cyclic subclasses is considered. Analogues (4) and (12) are obtained for this particular case.

**Conclusion.**

In the work, we have established the probability of the conditional distributions of partial sums of a sequence of conditionally m-independent of random vectors falling into a strip consisting of two shifted normal distributions. From this theorem, a similar result is obtained for sequences of random vectors with a chain $m$–dependence. This latter is a generalization of the result obtained by Z. Bezhaeva [11] to the multidimensional case, when the chain is not regular and has cyclic subclasses. A matrix representation of the covariance matrix of this shifted normal distribution with characteristics of a Markov chain is given. These characteristics are established in the sense of convergence and summation according to Cesàro.

**REFERENCES**


1. Normal Approximation Some Recent Advances. Sazonov V. V. Lecture Notes in Math. V. 79, Berlin, ete.,:Springer. 1981. https://www.amazon.com/Normal-Approximation-Advances-Lecture-Mathematics/dp/3540108637
2. Bokuchava I. V. Limit theorems for conditionally independent sequences. (in Russian) Teor. Verojatnost. i Primenen. XXIX. (1984). №1, p. 192-193. https://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=tvp&paperid=1992&option_lang=rus
3. Kvatadze Z., Shervashidze T. *Some limit theorems for I.I.D. and Conditionally independent random variables*. The second international Conference, "Problems of Cybernetics and Informatics". September 10-12, 2008. Baku. Azerbaijan. Section № 4. "Applied Stochastic Analysis". www.pci2008.science.az/4/12.pdf. Azerbaijan National Academy of Sciences. *Institute of Information Technologies of NASA*. Printing House of "Information Technology" Baku. 2008. Vol. II, 217-219., https://ict.az/uploads/konfrans/PCI2008/Pci%202008%20v%202/58.pdf
4. O' Brien G. L. Limit Theorems for Sum of Chain Dependent Proccesses. U. Appl. Probab., 1974, 11, 582-587. https://www.cambridge.org/core/journals/journal-of-applied-probability/article/abs/limit-theorems-for-sums-of-chaindependent-processes/279300A95BD1750B2E32FFD49AAA1B78
5. Bokuchava I., Kvatadze Z., Shervashidze T. *On Limit theorem for random vectors controlled by a Marcov chain*. Prob. Theory and Math. Stat., Vol. 1, 1986, 231-250. VNU Science Press, Utrecht. https://doi.org/10.1515/9783112319000-018 ;

6. Kvatadze Z., Pharjiani B., Kvatadze TS., One Example of $m$ – Dependent Vector's Sequence. Report of XXXV enlarged session of the seminar of Ilia Vekua Institute of Applied Mathematics (VIAM), of Ivane Javakhisvili Tbilisi State University (TSU). Sept. 16-19., 2021; Volume, 35. pp. 51-54. http://www.viam.science.tsu.ge/enl_ses/vol35/vol35.html ;
7. Billingsley P. Convergence of probability measures (in Russian). Moscow. "Nauka", 1971. http://bibliofilspb.ru/item.php?id=15498739 ;
8. J. G. Kemeny, J. L. Snell, Finite Markov chains. The university series in undergraduate mathematics. USA: Springer-Verlag; 1963, p. 272, https://www.math.pku.edu.cn/teachers/yaoy/Fall2011/Kemeny-Snell1976.pdf
9. Kvatadze Z., Kvatadze TS. Limiting Distribution of a Sequence of Functions Defined on a Markov Chain. Proceedings of A. Razmadze Mathematical institute. Vol. 174. 2020. issue 2, 199-205. https://rmi.tsu.ge/transactions/TRMI-volumes/174-2/v174(2)-8.pdf ;
10. Doob J. Probabilstic processes. (in russian). 1956, M. "Fizmatgiz". p. 772;
11. Bežaeva Z. I. Limit theorems for conditional Markov chains (in Russian). Teor. verojatn. i primen., 1971, v. XVI, № 3, p. 437–445. https://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=tvp&paperid=2257&option_lang=eng;
12. Kvatadze Z., Shervashidze T., On Limit Theorems for Conditionally Independent Random Variable Controlled by a Finite Markov Chain. Probability Theory and Mathematical Ststistics (Proc. 5th Japan-USSR Symposium on Probability Theory. Kyoto. July 8-14 1986). Lecture Notes in Mathematics. Vol. 1299, p. 250-259 Springer-Verlag, Berlin. etc., 1987. **DOI** · https://**doi**.org/**10.1007/BFb0078480**

# $m$ დამოკიდებულ ვექტორთა ჯამის პირობითი განაწილების შესახებ

**ზურაბ ქვათაძე*, ბექნუ ფარჯიანი**, ციალა ქვათაძე*****

* სტუ-ს მათემატიკის დეპარტამენტი. პროფესორი, ** სტუ-ს მათემატიკის დეპარტამენტი. ასოცირებული პროფესორი, *** სტატისტიკის ლექტორი, შავი სღვის საერთაშორისო უნივერსიტეტი,

**აბსტრაქტი**

მრავალი პრაქტიკული ამოცანის გადაჭრის დროს დგება დამოკიდებული დაკვირვებებით პარამეტრების სტატისტიკური შეფასებების აგების საკითხი. მაგალითად პროფესიული კოლეჯების რანჟირების მოდელის დასადგენად ჩატარებულ კვლევაში სტუდენტების გამოკითხვით მიღებული მონაცემები დამოკიდებულია ერთმანეთზე, გამომდინარე იქედან, რომ ისინი მიღებულია ერთი სოციუმის ჯგუფიდან (R. Chartolani, N. Durglishvili, Z. Kvatadze. Optimization of a State Financing Model of Vocational Colleges. Proc. A. Razmadze Math. Inst. 2015. 169.(2-15), pp. 23-31). ასევე ვერ უგულვებელვყოფთ მონაცემთა დამოკიდებულებას გეოფიზიკური მიმართულების სხვადასხვა კვლევების დროსაც. მაგალითად მიწისძვრის წინა (მოსამზადებელ) პერიოდში ქანებში მიმდინარე ტექტონიკური პროცესები დროში უწყვეტად ვითარდება და შესაბამისად დისკრეტულ მომენტებში აღებული ნებისმიერი მახასიათებლის ჩანაწერები როგორც ერთი მთლიანის გამოვლინებები დამოკიდებულია ერთმანეთზე. ანალოგიური მოსაზრებებიდან გამომდინარე სხვადასხვა პოპულაციების განაწილების კანონის დადგენის დროს ხდება სიმკვრივის შეფასების აგება დამოკიდებული დაკვირვებებით. მაგალითად ცნობილია სიმკვრივის არაპარამეტრული შეფასებები და რეგრესიის კოეფიციენტების შეფასებები რომლებიც აგებულია მარკოვის ჯაჭვად შეკრული დაკვირვებებით (Yakowitz Sidney (1989) Nonparametric density and regression estimation for Markov sequences without mixing assumptions. 85721–Journal of Multivariate Analysis, 30: 124-136. Arisona, USA). ასევე განიხილება პირობითად დამოუკიდებელი და ჯაჭვურად დამოკიდებული დაკვირვებები. ცნობილია მათი საშუალებით აგებული სიმკვრივის გულოვანი შეფასებები და მათი სიზუსტე $L_2$ მეტრიკით (Z, Kvatadze, B. Phardjiani. On the Exsactness of Distribution Density Estimates Constructed by Some Class of Dependent Observations. Mathematics and Statistics. 2019 Vol. 7(4), pp. 135-145. SAN JOSE) და $L_1$ მეტრიკით (B. Parjiani, L. Labadze, T. Kvatadze; Georgian Scientists, "On the accuracyby the metric L1 of the density estimation constructed by dependent observations" Vol. 5. Issue 1, pp. 308-321, 2023).

დამოკიდებული დაკვირვებებით სტატისტიკური შეფასებების ასაგებად და მათი ძალმოსილობის დასადგენად საჭიროა დამოკიდებელი შემთხვევითი სიდიდეების ჯამების განაწილების ასიმპტოტიკის ცოდნა. თანამედროვე ეტაპზე ხდება დამოუკიდებელ შემთხვევით სიდიდეთა შეკრების მდიდარი თეორიის (Normal Approximation Some Recent Advances. Sazonov V. V. Lecture Notes in Math. V. 79, Berlin, ete., :Springer. 1981.) გადატანა დამოკიდებულ შემთხვევით სიდიდეებზე. მრავალ ამოცანაში გამოიყენება მარკოვული დამოკიდებულება, რომელიც სუსტად დამოკიდებულების ერთ-ერთი სახეა. სხვადასხვა ტიპის დამოკიდებულების (სუსტად დამოკიდებული, პირობითად დამოუკიდებული,

მარკოვის ჯაჭვად შეკრული) შემთხვევითი სიდიდეების ჯამების ზღვარითი ასიმპტოტიკის საკითხები მრავალ ნაშრომშია გადმოცემული. ხშირად განიხილება ისეთ შემთხვევით სიდიდეთა ჯამები, რომელთა ერთობლივი განაწილება განისაზღვრება შემთხვევით ელემენტთა რაიმე მმართველი მიმდევრობით. განიხილება პირობითად დამოუკიდებელი (Bokuchava I. V. Limit theorems for conditionally independent sequences. (in Russian) Teor. Verojatnost. i Primenen. XXIX. (1984). №1, p. 192-193) და ჯაჭვურად დამოკიდებული მიმდევრობები. ამ საკითხებს ეხება მაგალითად გ. ობრაინის ( O' Braien G.L. Limit Theorems for Sum of Chain Dependent Proccesses. U. Appl. Probab., 1974, 11, 582-587); ი. ალეშკიავიჩუსის ( G. YU. Aleshkyavichus, On the central limit theorem for sums of random variables given on a Markov chain. (Russian) Lithuanian Mathematical Collected Works, Vilnius 6 (1966), №. 1, 15–22.); რ. ჩიტაშვილის, თ. შერვაშიძის, ი. ბოკუჩავას, ზ. ქვათაძის (Bokuchava I., Kvatadze Z., Shervashidze T. *On Limit theorem for random vectors controlled by a Marcov chain*. Prob. Theory and Math. Stat., Vol. 1, 1986, 239-250. VNU Science Press, Utrecht.) და სხვათა ნაშრომები. წინამდებარე ნაშრომში განხილულია პირობითად $m$ –დამოუკიდებულ ვექტორთა მიმდევრობების კლასი და ჯაჭვურად $m$ –დამოკიდებულ ვექტორთა მიმდევრობების კლასი (Kvatadze Z., ShervashidzeT. *Some limit theorems for I.I.D. and Conditionally independent random variables*. The second international Conference, "Problems of Cybernetics and Informatics". September 10-12, 2008. Baku. Azerbaijan. Section № 4. "Applied Stochastic Analysis". *Institute of Information Technologies of NASA*. Printing House of "Information Technology" Baku. 2008. Vol. II, 217-219). დადგენილია მათი ნორმირებული ჯამის ზღვარითი განაწილება. არარეგულარული ერგოდული ჯაჭვის შემთხვევაში ჯაჭვურად $m$ –დამოკიდებულ ვექტორთა მიმდევრობაზე განსაზღვრულია ფუნქციათა მიმდევრობა. მიღებულია ამ ფუნქციათა მიმდევრობის ნორმირებული ჯამის ზღვარითი კოვარიაციის მატრიცის წარმოდგენა მატრიცული სახით. ჯაჭვის შესაბამისი ფუნდამენტური მატრიცა და ზღვარითი მახასიათებლები ციკლური ქვეკლასების არსებობის შემთხვევაში გამოთვლილია ჩეზაროს აზრით კრებადობის გამოყენებით.

განაწილების ფუნქციათა სივრცეზე განაწილებების კრებადობის განხილვის დროს ხშირად გამოიყენება ლევი-პროხოროვის მეტრიკა $\rho(F,G)=\sup_{x\in R^k}\left|F(x)-G(x)\right|$. ამიტომ ბუნებრივია საინტერესოა საკითხი ფუნქციათა ჯამის განაწილების ორი ნორმალური წანაცვლებული განაწილებისგან შედგენილ ზოლში მოხვედრის ზღვარითი ალბათობის დადგენის შესახებ. ნაშრომში განხილულია პირობითად $m$ –დამოკიდებულ ვექტორთა მიმდევრობის ნორმირებული ჯამი. დადგენილია ამ ჯამის პირობითი განაწილების ორი წანაცვლებული ნორმალური განაწილებისგან შედგენილ ზოლში მოხვედრის ზღვარითი ალბათობა. ამ განაწილებების კოვარიაციის მატრიცა გამოსახულია ჯაჭვის პარამეტრების საშუალებით. დადგენილია ასეთივე ზღვარითი ალბათობა ჯაჭვურად $m$ –დამოკიდებულ ვექტორთა მიმდევრობის ნორმირებული ჯამისთვის.