LIMITING DISTRIBUTION OF A SEQUENCE OF FUNCTIONS DEFINED ON A MARKOV CHAIN

ZURAB KVATADZE\textsuperscript{1} AND TSIALA KVATADZE\textsuperscript{2}

Abstract. The present article shows the limiting distribution of partial sums of a functional sequence defined on a Markov Chain in case the chain is ergodic, with one class of ergodicity and contains cyclical subclasses.

Limiting behavior of sums of random variables is a classical problem in the probability theory, which is intensely studied by contemporaneous researchers both for independent variables and for the case of certain relationships between the terms of sequences. There exists a rich theory of sums of independent random variables (see, e.g., [5, 12, 13]). The problem of extending this case to the sums of dependent random variables introduces naturally the Markovian dependence, which in turn represents particular type of a weak dependency. The limiting theorems by Rosenblatt, Ibragimov and others concerning weakly dependent sequences are usually stated in terms of $\sigma$-algebras generated by asymptotically separable intervals of the sequence. The process of their investigation involves the so-called S. Bernstein’s “sectioning” method based on the weakening effect taking place during separation of groups of dependent variables (see [6]). Contemporaneous situation in the theory of sums of dependent random variables is expressed by using limiting theorems for martingales and semi-martingales (see [7]).

Different authors considered sums of random variables, whose joint distribution is determined by the controlling sequence of random variables (see [2, 3, 9]). An important part of these comprise problems regarding the sums of variables is defined directly on a chain (see [1, 3, 8, 11]). This paper considers the limiting theorem for functions defined on a stationary, finite, ergodic Markov chain.

We consider stationary, homogeneous, finite $\{\xi_i\}_{i\geq 1}$ ergodic Markov chain with one class of ergodicity (might containing cyclic subclasses) defined on a probability space $(\Omega, F, P)$. The chains have a set of states $\Xi = \{b_1, b_2, \ldots, b_r\}$, a matrix of transient probabilities $P = \|P_{\alpha\beta}\|_{\alpha,\beta=1}^{r}$ and a vector of limiting distribution of stationary probabilities $\pi = (\pi_1, \pi_2, \ldots, \pi_r)$ representing a solution of the following matrix equation:

$$\pi = \pi P.$$

We suggest that the initial distribution is stationary, the distribution

$$P(\xi_1 = b_\alpha) = \pi_\alpha, \quad \alpha = 1, r,$$

is based on stationarity means and the chain has the same distribution for each step

$$P(\xi_n = b_\alpha) = \pi_\alpha, \quad \alpha = 1, r, \quad n = 1, 2, \ldots.$$

Next, we introduce the Cezaro definition for convergence of the sequence and, relying on that definition, we establish all types of convergence when the chain has cyclical subclasses.

The sequence $\{t_n\}_{n\geq 1}$ is Cezaro convergent to $t$, and we write

$$\lim_{n \to \infty} t_n = t,$$

if the means of the first $n$ terms of the sequence converge to $t$:

$$\lim_{n \to \infty} T_n = t,$$

2020 Mathematics Subject Classification. 62G05, 62G07.

Key words and phrases. Markov Chain; Limiting distribution; Class of ergodicity; Cyclical subclasses.
where \( T_n = \frac{1}{n} \sum_{i=1}^{n-1} t_i \).

Cezaro convergence may be considered upon analyzing the convergence of series.

The series \( \sum_{k=1}^{\infty} a_k \) is said to be Cezaro convergent and the sum be equal to \( a \), if \( a \) is the Cezaro limit of the sequence of the partial sum \( S_n = \sum_{k=1}^{n} a_k \),

\[
\lim_{n \to \infty} (S_n)_c = a,
\]

which implies that there exists the limit of the sequence \( \tilde{a}_n \),

\[
\lim_{n \to \infty} \tilde{a}_n = a,
\]

where \( \tilde{a}_n = \frac{1}{n} \sum_{k=0}^{n} (n-k)a_k \) and, at the same time, this \( a \) represents the Cezaro sum of the series under consideration which can be written as

\[
\left( \sum_{k=1}^{n} a_k \right)_c = a.
\]

We denote by \( \Pi \) the limit (see [8])

\[
\lim_{n \to \infty} (P^n)_c = \Pi = \begin{pmatrix}
\pi_1, & \pi_2, & \ldots, & \pi_r \\
\pi_1, & \pi_2, & \ldots, & \pi_r \\
& & & \\
& & & \\
& & & \\
\pi_1, & \pi_2, & \ldots, & \pi_r
\end{pmatrix}
\]

It is obvious that

\[
\Pi = \|\pi_{\alpha\beta}\|_{\alpha,\beta=1\rightarrow r}; \quad \pi_{\alpha\beta} = \pi_{\beta\alpha}; \quad \alpha,\beta = 1\rightarrow r,
\]

\[
\lim_{n \to \infty} (P^n)_{\alpha\beta} = \pi_{\beta\alpha}, \quad \alpha,\beta = 1\rightarrow r.
\]

Let the fundamental matrix of the chain be

\[
Z = \|z_{\alpha\beta}\|_{\alpha,\beta=1\rightarrow r},
\]

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

where \( I \) is the identity matrix of \( r \times r \) dimensions. For the regular chain, the convergence of series is implied to be a standard convergence.

Let us consider a vector function defined on the \( \Xi \) space

\[
f(\xi_i) : \Xi \to R^k,
\]

\[
f(\xi_i) = (f_1(\xi_i), f_2(\xi_i), \ldots, f_k(\xi_i))
\]

and introduce the notation:

\[
f(b_\alpha) = f(\alpha) = (f_1(\alpha), f_2(\alpha), \ldots, f_k(\alpha)), \quad \alpha = 1\rightarrow r,
\]

\[
f_i(\alpha) = f_i(b_\alpha), \quad i = 1\rightarrow k, \quad \alpha = 1\rightarrow r.
\]

Theorem 1. When \( \{\xi_i\}_{i \geq 1} \) is the above-mentioned Markov chain and \( f = (f_1, f_2, \ldots, f_k) \) is the \( k \)-dimensional vector function from \( \Xi \) to \( R^k \), then if the limiting covariance matrix of the sum is

\[
U_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} [f(\xi_j) - Ef(\xi_j)]
\]

\[
T_f = \|t_{f_{i,j}}\|_{i,j=1\rightarrow k},
\]

\[
t_{f_{i,j}} = \sum_{\alpha,\beta} (\pi_{\alpha} z_{\alpha\beta} + \pi_{\beta} z_{\beta\alpha} - \pi_{\alpha} \pi_{\beta} - \pi_{\alpha} \delta_{\alpha\beta}) f_i(\alpha) f_j(\beta) \quad i, j = 1\rightarrow k
\]
(where $\delta_{\alpha\beta}$ is the Kronecker symbol) is positively defined, as $n \to \infty$, there is a convergence

$$P_{U_n} \xrightarrow{W} \Phi_{T_f}.$$ 

The case for $k = 1$, when $\lim_{n \to \infty} D(U_n) > 0$ (where $D(\cdot)$ denotes variance), is a famous fact (see [4, 10]) (when $\lim_{n \to \infty} D(U_n) = 0$, then $U_n$ converges to zero in probability) and $T_f$ can be written explicitly as a sum of components of the chain (see [8])

$$t = \lim_{n \to \infty} D(U_n) = \sum_{\alpha, \beta = 1}^{r} (\pi_\alpha z_{\alpha\beta} + \pi_\beta z_{\beta\alpha} - \pi_\alpha \pi_\beta - \pi_\alpha \delta_{\alpha\beta}) f(\alpha) f(\beta).$$

**Proof.** Using the Kramer-Wold method, we can derive the multidimensional case. Using the chain characteristic, we derive a matrix representation of the matrix $T_f$. Let us introduce a $k \times r$ matrix $F$,

$$F = \begin{pmatrix}
    f_1(b_1), & f_1(b_2), & \ldots, & f_1(b_r) \\
    f_2(b_1), & f_2(b_2), & \ldots, & f_2(b_r) \\
    \vdots & \vdots & \ddots & \vdots \\
    f_k(b_1), & f_k(b_2), & \ldots, & f_k(b_r)
\end{pmatrix} = \begin{pmatrix}
    f_1(1), & f_1(2), & \ldots, & f_1(r) \\
    f_2(1), & f_2(2), & \ldots, & f_2(r) \\
    \vdots & \vdots & \ddots & \vdots \\
    f_k(1), & f_k(2), & \ldots, & f_k(r)
\end{pmatrix},$$

and denote

$$V_0 = \text{cov}[f(\xi_1)] = E\{[f(\xi_1) - Ef(\xi_1)]^T [f(\xi_1) - Ef(\xi_1)]\},$$

$$V_j = E\{[f(\xi_1) - Ef(\xi_1)]^T [f(\xi_{1+j}) - Ef(\xi_{1+j})]\}, \quad j > 0,$$

$$V_{-j} = E\{[f(\xi_{1+j}) - Ef(\xi_{1+j})]^T [f(\xi_1) - Ef(\xi_1)]\}, \quad j > 0.$$

Based on the stationarity of the sequence $\{\xi_i\}_{i \geq 1}$, as $n \to \infty$, we have

$$E[U_n^T U_n] = \frac{1}{n} \left[ nV_0 + \sum_{j=1}^{n-1} (n-j) (V_j + V_{-j}) \right]$$

$$= V_0 + \frac{1}{n} \sum_{j=1}^{n} (n-j) V_j + \frac{1}{n} \sum_{j=1}^{n} (n-j) V_{-j} \xrightarrow{n \to \infty} V_0 + \left( \sum_{j=1}^{\infty} V_j \right)_c + \left( \sum_{j=1}^{\infty} V_{-j} \right)_c,$$

(2)

where $(\cdot)_c$ denotes the Czearo convergence of the sum in the parenthesis. It is obvious that if the chain is regular, this convergence is equivalent to the standard case of convergence of partial sums.

Thus, $T_f$ represents the limiting covariance of the sum $U_n$ and we have

$$T_f = V_0 + \left( \sum_{j=1}^{\infty} V_j \right)_c + \left( \sum_{j=1}^{\infty} V_{-j} \right)_c.$$

(3)

In the right-hand side, the convergence of matrix series is equivalent to that of a regular chain by virtue of a common definition of the convergence.

We now express $V_0$ and $V_j$ matrices based on the components of the chain

$$Ef(\xi_1) = \sum_{\alpha=1}^{r} \pi_\alpha f(\alpha) = \left( \sum_{\alpha=1}^{r} \pi_\alpha f_1(\alpha), \ldots, \sum_{\alpha=1}^{r} \pi_\alpha f_k(\alpha) \right)$$

$$= (\pi_1, \pi_2, \ldots, \pi_r) \begin{pmatrix}
    f_1(1), & f_2(1), & \ldots, & f_k(1) \\
    f_1(2), & f_2(2), & \ldots, & f_k(2) \\
    \vdots & \vdots & \ddots & \vdots \\
    f_1(r), & f_2(r), & \ldots, & f_k(r)
\end{pmatrix} = \pi F^T,$$
\[ E\{f(\xi_1)^T f(\xi_1)\} = E \left \{ \begin{pmatrix} f_1(\xi_1) \\ f_2(\xi_1) \\ \vdots \\ f_k(\xi_1) \end{pmatrix} (f_1(\xi_1), f_2(\xi_1), \ldots, f_k(\xi_1)) \right \} \]

\[ = \|E f_i(\xi_1) f_j(\xi_1)\|_{i,j=1}^\infty = \left \| \sum_{\alpha=1}^r \pi_\alpha f_i(\alpha) f_j(\alpha) \right \|_{i,j=1}^\infty \]

\[ = \left( \sum_{\alpha=1}^r \pi_\alpha f_1(\alpha) f_1(\alpha), \sum_{\alpha=1}^r \pi_\alpha f_1(\alpha) f_2(\alpha), \ldots, \sum_{\alpha=1}^r \pi_\alpha f_1(\alpha) f_k(\alpha) \right) \]

\[ \sum_{\alpha=1}^r \pi_\alpha f_2(\alpha) f_1(\alpha), \sum_{\alpha=1}^r \pi_\alpha f_2(\alpha) f_2(\alpha), \ldots, \sum_{\alpha=1}^r \pi_\alpha f_2(\alpha) f_k(\alpha) \]

\[ \ldots \]

\[ \sum_{\alpha=1}^r \pi_\alpha f_k(\alpha) f_1(\alpha), \sum_{\alpha=1}^r \pi_\alpha f_k(\alpha) f_2(\alpha), \ldots, \sum_{\alpha=1}^r \pi_\alpha f_k(\alpha) f_k(\alpha) \]

\[ = F\Pi_{dg} F^T; \]

\[ E\{f(\xi)^T f(\xi+j)\} = \|E f_i(\xi) f_s(\xi+j)\|_{i,s=1}^\infty \]

\[ = \left\| \sum_{\alpha,\beta=1}^r \pi_\alpha f_i(\alpha) \Pi_{\alpha\beta}^j f_s(\beta) \right\|_{i,s=1}^\infty = F\Pi_{dg} \Pi^j F^T, \]

where \((\cdot)_{dg}\) denotes the matrix obtained by replacing each element of the matrix in the parenthesis by zero, except ones located on the main diagonal.

The following equality

\[ \pi^T \pi = \Pi_{dg} \Pi \]

holds and the derived equations will be taken into account in the expression for \(V_j\). When \(j = 0\), we obtain

\[ V_0 = E \{ f^T(\xi_1) f(\xi_1) \} - E \{ f^T(\xi_1) \} E \{ f(\xi_1) \} \]

\[ = F\Pi_{dg} F^T - (\pi F^T)^T \pi F^T = F\Pi_{dg} F^T - F\pi^T \pi F^T \]

\[ = F\Pi_{dg} F^T - F\Pi_{dg} \Pi F^T = F(\Pi_{dg} - \Pi_{dg} \Pi) F^T. \]

By the stationarity \(E\{f(\xi_{j+1})\} = E\{f(\xi_1)\}\), when \(j > 0\), the equalities

\[ V_j = E \{ f^T(\xi_1) f(\xi_{j+1}) \} - E \{ f^T(\xi_1) \} E \{ f(\xi_{j+1}) \} \]

\[ = \left\{ f^T(\xi_1) f(\xi_{j+1}) \right\} - E \left\{ f^T(\xi_1) \right\} E \left\{ f(\xi_1) \right\} \]

\[ = F\Pi_{dg} F^T - (\pi F^T)^T \pi F^T = F\Pi_{dg} F^T - F\pi^T \pi F^T \]

\[ = F\Pi_{dg} P^j F^T \]

are true.

Thus, the sum in the right-hand side of (2) can be expressed as

\[ \left( \sum_{j=1}^\infty V_j \right)_c = \left( \sum_{j=1}^\infty F\Pi_{dg} (P^j - \Pi) F^T \right)_c = F\Pi_{dg} \left( \sum_{j=1}^\infty (P^j - \Pi) \right)_c F^T. \]
By the property of the fundamental matrix, we have
\[
\left( \sum_{j=1}^{\infty} (P^j - \Pi) \right)_c = Z - I.
\]

Thus we get the following equation:
\[
\left( \sum_{j=1}^{\infty} V_j \right)_c = F\Pi_d g (Z - I)F^T = F(\Pi_d g Z - \Pi_d g)F^T.
\]

Like equation (4), the following sum can be computed by using stationarity of the chain
\[
\left( \sum_{j=1}^{\infty} V_{-j} \right)_c = \left( \sum_{j=1}^{\infty} E\{[f(\xi_{1+j}) - Ef(\xi_{1+j})]^T[f(\xi_1) - Ef(\xi_1)]\} \right)_c
\]
\[
= \left( \sum_{j=1}^{\infty} E\{[f(\xi_1) - Ef(\xi_1)]^T[f(\xi_{1+j}) - Ef(\xi_{1+j})]\} \right)_c
\]
\[
= \left[ \left( \sum_{j=1}^{\infty} E\{[f(\xi_1) - Ef(\xi_1)]^T[f(\xi_{1+j}) - Ef(\xi_{1+j})]\} \right) \right]_c = \left[ \left( \sum_{j=1}^{\infty} V_j \right)_c \right]_c^T = F(\Pi_d g Z - \Pi_d g)F^T.
\]

Substituting the obtained results into (3) and using characteristic matrices corresponding to the chain, we get the following matrix expression for \( T_f \),
\[
T_f = F(\Pi_d g Z + (\Pi_d g Z)^T - \Pi_d g)F^T.
\]

Obviously, the \( t_{f_{i,j}} \) elements of the matrix \( T_f \) can be expressed by virtue of (1).

Next, we introduce a characteristic of time moments quantity elapsed by the chain at the first \( n \) steps in different \( b_\alpha \), \( \alpha = \bar{1}, \bar{r} \) positions.

Let \( \nu_n(\alpha) = \nu_n(b_\alpha) \), \( \alpha = \bar{1}, \bar{r} \) be a random variable representing the amount of time intervals during the first \( n \) steps when the chain is in position \( b_\alpha \), \( \alpha = \bar{1}, \bar{r} \) on a fixed trajectory \( \bar{\xi}_{1n} = (\xi_1, \xi_2, \ldots, \xi_n) \). Then it is obvious that the equation
\[
\nu_n(1) + \nu_n(2) + \cdots + \nu_n(r) = n
\]
holds.

The quantity \( \frac{\nu_n(\alpha)}{n} \) is a part of time \( n \) during which the chain at the first \( n \) steps spends in condition \( b_\alpha \). \( \square \)

**Theorem 2.** The \( \nu_n(\alpha) \), \( \alpha = \bar{1}, \bar{r} \), random variable is measurable with respect to the sigma algebra induced by dividing the \( \Omega \) space during fixation of a \( \bar{\xi}_{1n} \) trajectory.

**Proof.** We show that a discrete random variable \( \nu_n(\alpha) \) attains constant values on sets generated by partitioning the \( \Omega \) space during fixation of a \( \bar{\xi}_{1n} \) trajectory.

Conditions set of the chain is \( \Xi = \{b_1, b_2, \ldots, b_r\} \). On a fixed \( \bar{\xi}_{1n} \) trajectory, possible values will be the Cartesian product \( \Xi^n = \Xi \times \Xi \times \cdots \times \Xi \). Let us show how the \( \Omega \) space will be partitioned.

Introduce the following sets:
\[
D_{1,2,\ldots,n,m_1,m_2,\ldots,m_n} = \{\omega \mid \xi_1 = b_{m_1}, \xi_2 = b_{m_2}, \ldots, \xi_n = b_{m_n}\}, \quad b_{m_i} \in \Xi, \quad i = \bar{1}, \bar{r}.
\]

Fixation of a \( \bar{\xi}_{1n} \) trajectory will result in a partition of the \( \Omega \) space,
\[
\bar{D} = \{D_{1,2,\ldots,n,m_1,m_2,\ldots,m_n} \mid m_i \in \{1, 2, \ldots, r\}\}.
\]
It is clear that
\[ D_{ik} = \{ \omega | \xi_i = b_k \} = \sum_{m_i \in \Xi \setminus \{ b_i \}} \{ D_{1,2,\ldots,n,m_1,m_2,\ldots,m_{i-1},k,m_{i+1},\ldots,m_n} \}, \]
\[ \xi_i = \sum_{k=1}^{r} b_k I_{(D_{ik})}. \]

To derive analytical expression for the sum \( \nu_n(i) \), consider the sets
\[ A_{n,j_1,j_2,\ldots,j_k} = \left\{ \omega \mid \begin{array}{l} j_1 < j_2 < \cdots < j_k \\ \xi_\alpha = b_i \quad \alpha \in \{ j_1, \ldots, j_k \} \\ \xi_\alpha \in \Xi \setminus \{ b_i \} \quad \alpha \notin \{ j_1, \ldots, j_k \}, \alpha = 1, n \end{array} \right\} \]
\[ = \sum_{j_1 < j_2 < \cdots < j_k} D_{j_1,\ldots,j_n,m_{j_1},\ldots,m_{j_k}}; \]
\[ A_{n,k} = \{ \nu_n(b_i) = k \} = \sum_{j_1 < j_2 < \cdots < j_k} A_{n,j_1,\ldots,j_k} \]
\[ = \sum_{j_1 < j_2 < \cdots < j_k} \sum_{\{j_1, j_2, \ldots, j_k\} \subset \{1, 2, \ldots, n\}} D_{j_1,\ldots,j_n,m_{j_1},\ldots,m_{j_k}}. \]

Clearly, the \( A_{n,j_1,j_2,\ldots,j_k} \) type sets are \((r-1)^{n-k}\) in total, while there are \( C_n^k \cdot (r-1)^{n-k} \) \( A_{n,k} \) type sets.

Relying on the above-said, we easily find that
\[ \nu_n(i) = \nu_n(b_i) = \sum_{k=0}^{n} k I_{A_{n,k}^i}. \]

Thus, the measurability of a \( \nu_n(i) \) random variable with respect to partition \( \tilde{D} \) is shown. Clearly, this implies that the variable is measurable with respect to the sigma algebra generated by that partition. Note finally that any function \( f(\nu_n(i)) \) is also measurable, where \( f(\cdot) \) is a continuous function.

REFERENCES


(Received 20.12.2019)

1Department of Mathematics, Georgian Technical University, 77 Kostava Str., Tbilisi 0175, Georgia

E-mail address: zurakvatadze@yahoo.com

2Faculty of Informatics and Management Systems, Georgian Technical University, 77 Kostava Str., Tbilisi 0175, Georgia

E-mail address: ttkvatadze@gmail.com