**Pak. J. Statist.**
**2019 Vol. 35(1), 53-63**

# GENERALIZED ABSOLUTE MEAN STRONG ERGODICITY OF NONHOMOGENEOUS MARKOV CHAINS

**Waqas Anwar**[§] and **Yang Weiguo**
Faculty of Science, Jiangsu University, Zhenjiang, Jiangsu, 212013, China
[§]Corresponding author Email: waqas.anwar2012@yahoo.com

## ABSTRACT

The aim of this paper is to create some theorems about generalized absolute mean strong ergodicity for the countable nonhomogeneous Markov chains. Firstly, we define absolute mean strong ergodicity and generalized absolute mean strong ergodicity for countable nonhomogeneous Markov chain. Then, we prove main theorem regarding generalized absolute mean strong ergodicity for countable nonhomogeneous Markov chain and give some applications.

## KEY WORDS

Nonhomogeneous Markov chains, Absolute Mean Strong Ergodicity, Generalized absolute mean strong ergodicity.

## 1. INTRODUCTION

Firstly, we intend to introduce some common notations which will be frequently used in this paper.

Consider a nonhomogeneous Markov chain $\{X_n,\ n \geq 0\}$ taking values in countable state space $S = \{1,2, \ldots\}$ having the transition matrices $P_n = \big(p_n(i,j)\big)$, where

$$p_n(i,j) = P(X_n = j | X_{n-1} = i)\ n \geq 1, i, j \epsilon\ S. \tag{1.1}$$

Let

$$P^{(m,n)} = P_{m+1} P_{m+2} \ldots P_n \tag{1.2}$$

where, any element $p^{(m,n)}(i,j)$ of $P^{(m,n)}$ is defined by using Chapman-Kolmogrov Identity as follows:

$$p^{(m,n)}(i,j) = P(X_n = j | X_m = i). \tag{1.3}$$

If $\{X_n,\ n \geq 0\}$ is a homogeneous Markov chain, then transition matrices $\{P_n, n \geq 0\}$ is simply represented by $P$ and $P^{(m,m+k)}$ by $P^k$.

Consider a matrix $U = \big(u_{ij}\big)$. The norm $\|\cdot\|$ of a matrix $U$ (see (Isaacson and Madsen 1976), Page 138) is defined by

$$\|U\| = \sup_i \sum_{j=1}^{\infty} \left|u_{ij}\right|. \tag{1.4}$$

 

The norm of a row vector $f = (f_1, f_2, \dots)$ is defined norm as $\|f\| = \sum_{i\epsilon S}|f_i|$, and the norm of a column vector $g = (g_1, g_2, \dots)'$ is defined as $\|g\| = \sup_i |g_i|$.

The following properties always hold for the above definition of norm (see (Bowerman, David et al. 1977))

(a) If $U$ and $V$ are any two matrices, then $\|UV\| \leq \|U\| \cdot \|V\|$,
(b) If $P$ is a stochastic matrix, then $\|P\| = 1$.

These two properties will be frequently used in this paper. The property (a) also holds if $f$ is any vector and $U$ is any matrix, i.e. $\|fU\| \leq \|f\| \cdot \|U\|$. Consider a distribution on a state space $S$ defined as follows:

$$f^{(k)} = \left(f_1^{(k)}, f_2^{(k)}, \dots\right) \tag{1.5}$$

where, $f_i^{(k)} = P(X_k = i)$. We call this distribution as initial distribution or starting vector. By (1.5), we have (see (Isaacson and Madsen 1976), Page 137)

$$f^{(k)} = f^{(0)}P^{(0,k)} = f^{(0)}P_1P_2 \dots P_k.$$

A stochastic matrix $Q$ is said to be constant stochastic matrix if every row of a matrix is same.

Consider a nonhomogeneous Markov chain $\{X_n,\ n \geq 0\}$ having the transition matrices $\{P_n, n \geq 0\}$, a constant stochastic matrix $Q$, then we say the Markov chain is strongly ergodic (see (Isaacson and Madsen 1976), Page 157), if $\forall m$

$$\lim_{k\to\infty}\left\|P^{(m,k)} - Q\right\| = 0. \tag{1.6}$$

If $\{X_n,\ n \geq 0\}$ is a homogeneous Markov chain having the transition matrix $P = \left(p_{ij}\right)$, then strong ergodicity becomes

$$\lim_{n\to\infty}\|P^n - Q\| = 0. \tag{1.7}$$

We say a nonhomogeneous Markov chain $\{X_n, n \geq 0\}$ having transition matrices $\{P_n, n \geq 0\}$ converges in the Cesaro sense to the matrix $Q$, if $\forall m$

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{k=1}^{n} P^{(m,m+k)} - Q\right\| = 0. \tag{1.8}$$

This is also referred as convergence of the Cesaro averages or simply C-strong ergodicity (see (Isaacson and Madsen 1976), Page 184).

We define absolute mean strong ergodicity of a nonhomogeneous Markov chain as if $\forall m$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\left\|P^{(m,m+k)} - Q\right\| = 0 \tag{1.9}$$

where, $Q$ is a constant stochastic matrix. It is easy to see that absolute mean strong ergodicity implies C-strong ergodicity.

Consider a sequence of non-negative integers $\{b_n, n \geq 0\}$ and a sequence of monotonically increasing non-negative integers $\{\psi(n), n \geq 0\}$ such that $\lim_{n\to\infty} \psi(n) = \infty$. If $\forall m$,

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} - Q \right\| = 0. \tag{1.10}$$

We call this nonhomogeneous Markov chain as generalized absolute mean strongly ergodic. Here, note that absolute mean strong ergodicity follows from generalized absolute mean strong ergodicity if $b_n = 0$ and $\psi(n) = n$.

The topic such as strong ergodicity and C-strong ergodicity for nonhomogeneous Markov chains have been studied by several authors in recent years. Iosifescu Marius (1972) also worked on some ergodicity properties for nonhomogeneous Markov chains. The rate of convergence of some homogeneous and nonhomogeneous Markov chains have been studied by Huang, Isaacson and Vinograde (1976). The convergence of Cesaro averages for certain nonhomogeneous Markov chains have been studied by Bowerman, David and Isaacson (1977). The necessary and sufficient condition for strongly ergodic nonhomogeneous Markov chains have been proved by Isaacson and Seneta (1982). It has been proved that the strong ergodicity implies weak ergodicity. The equivalence of strong ergodicity and weak ergodicity, if the Markov chain is nonhomogeneous already has been proved by Isaacson and Madsen, (see (Isaacson and Madsen 1976), Chapter V). Isaacson and Madsen have proved that if $P$ is strongly ergodic with respect to constant stochastic matrix $Q$ and $\lim_{n\to\infty} \|P_n - P\| = 0$, then the nonhomogeneous Markov chain is strongly ergodic (see (Isaacson and Madsen 1976), Page 170). The weak ergodicity of nonhomogeneous Markov chains have been studied by Tan (1996). Yang (2002) proved the C-strong ergodicity for countable nonhomogeneous Markov chains using the condition $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|P_k - P\| = 0$ which is obtained by generalizing the results of Bowerman, David and Isaacson (1977). The strong law of large numbers for countable nonhomogeneous Markov chains have been studied by Yang (2009). The Dobrushin ergodicity coefficient and the ergodicity of nonhomogeneous Markov chains have been studied by Mukhamedov (2013). Galtchouk and Pergamenshchikov (2014) have worked on geometric ergodicity for some classes of homogeneous Markov chains.

In this paper, firstly we define absolute mean strong ergodicity and generalized mean strong ergodicity for nonhomogeneous Markov chains. Then, we state a useful lemma and give its proof. Finally, we give the main theorem and its proof for the generalized absolute mean strongly ergodic nonhomogeneous Markov chains and also give some applications.

This paper has following formation: In Section 1, we introduced few notations which will be used frequently in this paper, some definitions and aim of paper. In Section 2, we state an important lemma and give its proof. Finally, we propose our main theorem about generalized absolute mean strong ergodicity and also give some applications.

## 2. GENERALIZED ABSOLUTE MEAN STRONG ERGODICITY AND ITS APPLICATIONS

Before proposing the main results, we will give a useful lemma.

**Lemma 2.1.**

*Consider a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *defined as before. Consider two sequences of non-negative integers* $\{b_n, n \geq 0\}$ *and* $\{\psi(n), n \geq 0\}$ *such that for any non-negative integers* $m, n$

$$\psi(m+n) - \psi(n) \geq m,\ \frac{\psi(m+n)}{\psi(n)} \to 1\ (n \to \infty). \tag{2.1}$$

*If*

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \|P_k - P\| = 0, \tag{2.2}$$

*then,* $\forall k$

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \|P_{m+k} - P\| = 0. \tag{2.3}$$

**Remark 2.1.**

*If* $\psi(n) = n^{\alpha}$ $(\alpha \geq 1)$, *then (2.1)holds.*

**Proof:**

$$\frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \|P_{m+k} - P\| = \frac{1}{\psi(n)} \sum_{\upsilon=b_n+m+1}^{b_n+m+\psi(n)} \|P_\upsilon - P\|$$

*where,* $m + k = \upsilon$

$$\leq \frac{1}{\psi(n)} \sum_{\upsilon=b_n+1}^{b_n+m+\psi(n)} \|P_\upsilon - P\|$$

$$\leq \frac{1}{\psi(n)} \sum_{\upsilon=b_n+1}^{b_n+\psi(m+n)} \|P_\upsilon - P\|$$

$$\leq \frac{\psi(m+n)}{\psi(n)} \cdot \frac{1}{\psi(m+n)} \sum_{\upsilon=b_n+1}^{b_n+\psi(m+n)} \|P_\upsilon - P\|. \tag{2.4}$$

*Since as* $n \to \infty$, $\frac{\psi(m+n)}{\psi(n)} \to 1$, *therefore by (2.2) and(2.4), (2.3) follows.* ■

**Theorem 2.2.**

*Consider a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *having the transition matrices* $P_n = \big(p_n(i,j)\big), n \geq 1, i, j \epsilon\ S$ *defined on the countable state space* $S = \{1,2, \ldots\}$. *Consider two sequences of non-negative integers* $\{b_n, n \geq 1\}$ *and* $\{\psi(n), n \geq 1\}$ *such that (2.1)holds. Let* $P$ *be any other transition matrix. Suppose that* $P$ *is strongly ergodic with respect to constant stochastic matrix* $Q$ *i.e. (2.2) hold, then this Markov chain is generalized absolute mean strongly ergodic, that is, (1.10) holds.*

**Proof:**

*Since*

$$\left\|P^{(m+k,m+k+2)} - P^2\right\| = \|P_{m+k+1}P_{m+k+2} - P^2\|$$
$$= \|P_{m+k+1}P_{m+k+2} - P_{m+k+1}P + P_{m+k+1}P - P^2\|$$
$$\leq \|P_{m+k+1}P_{m+k+2} - P_{m+k+1}P\| + \|P_{m+k+1}P - P^2\|$$
$$\leq \|P_{m+k+2} - P\| + \|P_{m+k+1} - P\|.$$

*Using (2.2) and Lemma 2.1,* $\forall m$

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\|P^{(m+k,m+k+2)} - P^2\right\| = 0. \tag{2.5}$$

*By induction, we have*

$$\lim_{n\to\infty} \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\|P^{(m+k,m+k+v)} - P^v\right\| = 0. \tag{2.6}$$

*Since,* $P$ *is strongly ergodic with respect to constant stochastic matrix* $Q$, *then* $\forall\ \varepsilon > 0, \exists\ v_0$, *we have*

$$\|P^{\,v_0} - Q\| < \ \varepsilon. \tag{2.7}$$

*It is easy to see that* $PQ = Q$ *where* $Q$ *is a constant stochastic matrix, and* $P$ *is any other stochastic matrix. By (2.7),* $\forall m$ *we have*

$$\frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\|P^{(m,m+k)} - Q\right\| = \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\,v_0+1} \left\|P^{(m,m+k)} - Q\right\|$$
$$+ \frac{1}{\psi(n)} \sum_{k=b_n+\,v_0+1}^{b_n+\psi(n)} \left\|P^{(m,m+k)} - Q\right\|$$
$$\leq \frac{2(\,v_0 + 1)}{\psi(n)} + \frac{1}{\psi(n)} \sum_{k=b_n+\,v_0+1}^{b_n+\psi(n)} \left\|P^{(m,m+k-\,v_0)}P^{(m+k-\,v_0,m+k)} - Q\right\|$$

$$\leq \frac{2(\upsilon_0+1)}{\psi(n)} + \frac{1}{\psi(n)} \sum_{k=b_n+\upsilon_0+1}^{b_n+\psi(n)} \left\| P^{(m+k-\upsilon_0,m+k)} - P^{\upsilon_0} \right\|`$$

$$+ \frac{1}{\psi(n)} \sum_{k=b_n+\upsilon_0+1}^{b_n+\psi(n)} \| P^{\upsilon_0} - Q \|$$

$$\leq \frac{2(\upsilon_0+1)}{\psi(n)} + \frac{1}{\psi(n)} \sum_{k'=b_n+1}^{b_n+\psi(n)-\upsilon_0} \left\| P^{(m+k',m+k'+\upsilon_0)} - P^{\upsilon_0} \right\| + \varepsilon$$

*where,* $k' = k - \upsilon_0$

$$\leq \frac{2(\upsilon_0+1)}{\psi(n)} + \frac{1}{\psi(n)} \sum_{k'=b_n+1}^{b_n+\psi(n)} \left\| P^{(m+k',m+k'+\upsilon_0)} - P^{\upsilon_0} \right\| + \varepsilon. \qquad (2.8)$$

*It is easy to see that the first term of (2.8) approaches to zero, by (2.6)*(2.6)*, second term of (2.8) also approaches to zero. By (2.8) and arbitrariness of* $\varepsilon$*, (1.10) follows. This completes the proof.* ■

**Corollary 2.3.**

*Consider a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *having the transition matrices* $P_n = (p_n(i,j)), n \geq 1, i, j \epsilon\ S$ *defined on the countable state space* $S = \{1,2,\ldots\}$*. Let* $P$ *be any other transition matrix. Suppose that* $P$ *is strongly ergodic with respect to constant stochastic matrix* $Q$*. If*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \|P_k - P\| = 0 \qquad (2.9)$$

*then, this Markov chain is absolute mean strongly ergodic, that is, (1.9) holds.*

**Proof:**

*This corollary follows from Theorem 2.2 by putting* $b_n = 0$ *and* $\psi(n) = n$*.* ■

**Remark 2.2.**

*It is easy to see that (1.10) holds if* $\lim_{n\to\infty} P_n = P$*. Notice that*

$$\frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} - P \right\| \leq \left(1 + \frac{b_n}{\psi(n)}\right) \cdot \left(\frac{1}{b_n+\psi(n)}\right) \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} - P \right\|.$$

*Furthermore, if* $\left\{\frac{b_n}{\psi(n)}\right\}$ *is bounded sequence, then (1.10) follows from (1.9). But in general, the converse of this statement may not hold i.e. (1.9) may not implies to (1.10). For Example, consider*

$$P_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}, \quad P_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

*and a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *having transition matrices given by:*

$$P_n = \begin{cases} P_1; & if\ 2^k \leq n \leq 2^k + k,\ \ k > 0 \\ P_2; & otherwise \end{cases}$$

*If* $P = P_2$, *it is easy to see that when* $2^k \leq n < 2^{k+1}$, *for every* $i, j \epsilon X$

$$\frac{1}{n}\sum_{l=1}^{n}\left\|P^{(m,m+l)} - P\right\| \leq \frac{1}{2^k}\sum_{l=1}^{2^{k+1}-1}\left\|P^{(m,m+l)} - P\right\|$$
$$\leq \frac{1 + 2 + 3 + \cdots + (k+1)}{2^k}\cdot\frac{1}{6} = \frac{1}{2^k}\frac{(k+2)(k+1)!}{2}\frac{1}{6} \to 0\ (k \to \infty).$$

*Therefore, (1.9) holds. Conversely, if we consider* $b_n = 2^n$ *and* $\psi(n) = n$, *then*

$$\frac{1}{\psi(n)}\sum_{k=b_n+1}^{b_n+\psi(n)}\left\|P^{(m,m+k)} - P\right\| \leq \frac{1}{n}\sum_{k=2^n+1}^{2^n+n}\left\|P^{(m,m+k)} - P\right\| = \frac{1}{6}\cdot$$

*So, (1.10) does not hold.*

**Theorem 2.4.**

*Consider a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *having the transition matrices* $\{P_n,\ n \geq 1\}$. *Suppose that* $\{X_n,\ n \geq 0\}$ *is generalized absolute mean strongly ergodic Markov chain, that is,* $\exists$ *a constant stochastic matrix* $Q$ *such that (1.10) holds. Let* $q = (q_{1,} q_2, \dots)$ *is the row vector of* $Q$. *Consider two functions* $g_n(x)$ *and* $g(x)$ *defined on the state space* $S$, *where* $g_n$ *and* $g$ *are column vectors with* $j^{th}$ *elements* $g_n(j)$ *and* $g(j)$ *respectively. If*

$$\lim_{n\to\infty}\frac{1}{\psi(n)}\sum_{k=b_n+1}^{b_n+\psi(n)}\|g_k - g\| = 0 \tag{2.10}$$

*where,* $\{\|g_n\|,\ n \geq 1\}$ *is a bounded sequence. Let*

$$\pi^{(m,m+k)}(i) = E[g_{m+k}(X_{m+k})|X_m = i], i\epsilon S \tag{2.11}$$

*and let* $\pi^{(m,m+k)}$ *are column vectors with* $i^{th}$ *element as* $\pi^{(m,m+k)}(i)$, $\pi$ *is another column vector with each element* $\sum_j g(j)q_j$, *then* $\forall m$, *we get*

$$\lim_{n\to\infty}\frac{1}{\psi(n)}\sum_{k=b_n+1}^{b_n+\psi(n)}\left\|\pi^{(m,m+k)} - \pi\right\| = 0. \tag{2.12}$$

**Proof:**

*It is easy to see that* $\|g\|$ *is also bounded. Using (2.11), we get*

$$\pi^{(m,m+k)} = \sum_j g_{m+k}(j) P^{(m,m+k)}(i,j).$$

*and*

$$\frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| \pi^{(m,m+k)} - \pi \right\| = \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} g_{m+k} - Qg \right\|$$

$$= \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} g_{m+k} - P^{(m,m+k)} g + P^{(m,m+k)} g - Qg \right\|$$

$$\leq \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} g_{m+k} - P^{(m,m+k)} g \right\|$$

$$+ \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} g - Qg \right\|$$

$$\leq \frac{1}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \| g_{m+k} - g \| \quad + \frac{\|g\|}{\psi(n)} \sum_{k=b_n+1}^{b_n+\psi(n)} \left\| P^{(m,m+k)} - Q \right\|. \tag{2.13}$$

*By (1.10), (2.10) and (2.13), (2.12) follows.* ■

**Corollary 2.5.**

*Consider a nonhomogeneous Markov chain* $\{X_n,\ n \geq 0\}$ *having the transition matrices* $\{P_n,\ n \geq 1\}$. *Suppose that* $\{X_n,\ n \geq 0\}$ *is absolute mean strongly ergodic Markov chain, that is,* $\exists$ *a constant stochastic matrix* $Q$ *such that (1.9) holds. Let* $q = (q_{1,} q_2, \ldots)$ *is the row vector of* $Q$. *Consider two functions* $g_n(x)$ *and* $g(x)$ *defined on state space* $S$, *where* $g_n$ *and* $g$ *are column vectors with* $j^{th}$ *elements* $g_n(j)$ *and* $g(j)$ *respectively. If*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \| g_k - g \| = 0 \tag{2.14}$$

*where,* $\{\|g_n\|,\ n \geq 1\}$ *is a bounded sequence, then* $\forall m$, *we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left\| \pi^{(m,m+k)} - \pi \right\| = 0 \tag{2.15}$$

*where,* $\pi^{(m,m+k)}$ *and* $\pi$ *are same as in Theorem 2.4.*

**Proof:**

*This corollary follows from Theorem 2.4 by putting* $b_n = 0$ *and* $\psi(n) = n$. ■

*If we consider* $g_n(i)$ *as the cost of the Markov chain at the state* $i$ *on the time* $n$, *then the limit*

$$\lim_{n\to\infty} V^{(o,n)}(i) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \pi^{(0,k)}(i)$$

*is called expected cost of the Markov chain at state $i$. From Corollary 2.5, above limit exist under the condition of Corollary 2.5.*

*Consider the sequence of random variables $\{X_n,\ n \geq 0\}$ taking values in state space $S$, $H(X_0, \dots, X_n)$ be the entropy of the random vectors $(X_0, \dots, X_n)$. If the following limit exist*

$$\lim_{n\to\infty} \frac{1}{n} H(X_0, \dots, X_n)$$

*then, this limit is called entropy rate of $\{X_n,\ n \geq 0\}$, denoted by $H(\{X_n\})$ (see (Cover and Thomas 2012), Page 63). It is easy to see that (see (Cover and Thomas 2012), Page 21)*

$$H(X_0, \dots, X_n) = H(X_0) + H(X_1|X_0) + \cdots + H(X_n|X_0, \dots, X_{n-1}). \tag{2.16}$$

*If $\{X_n,\ n \geq 0\}$ is nonhomogeneous Markov chain having the transition matrices $P_n = \big(p_n(i,j)\big)$, $n \geq 0$ then, we get*

$$\frac{1}{n} H(X_0, \dots, X_n) = \frac{1}{n} H(X_0) + \frac{1}{n}\sum_{k=1}^{n} H(X_k|X_{k-1}) \tag{2.17}$$

*and*

$$H(X_k|X_{k-1}) = -\sum_i P(X_{k-1} = i) \sum_j p_k(i,j) \log p_k(i,j). \tag{2.18}$$

**Theorem 2.6**

*Consider a nonhomogeneous Markov chain $\{X_n,\ n \geq 0\}$ having the initial distribution and transition matrices as $f^{(0)} = (f_1, f_2, \dots)$, $P_n = \big(p_n(i,j)\big)$, $n \geq 1$, $i, j \epsilon S$ respectively. Let $P = \big(p_{ij}\big)$ be any other transition matrix, and suppose that $P$ is strongly ergodic with respect to constant stochastic matrix $Q$. Let*

$$g_n(i) = \sum_j p_n(i,j) \log p_n(i,j)\, , g(i) = \sum_j p_{ij} \log p_{ij}$$

*where, $g_n$ and $g$ are all column vectors with $i^{th}$ elements $g_n(i)$ and $g(i)$ respectively. We suppose that $\{\|g_n\|,\ n \geq 1\}$ is a bounded sequence. If (2.9) and (2.14) hold, then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} |H(X_k|X_{k-1}) + qg| = 0 \tag{2.19}$$

*where, row vector of the constant stochastic matrix $Q$ is $q = \big(q_{1,} q_2, \dots\big)$.*

**Proof:**

*It is easy to see that $\|g\|$ is finite. Using (1.5), we have*

$$\begin{aligned}\left\|f^{(k-1)}-q\right\| &= \left\|f^{(0)}P^{(0,k-1)}-f^{(0)}Q\right\| \\ &\le \left\|P^{(0,k-1)}-Q\right\|. \end{aligned} \tag{2.20}$$

*By (2.9), Corollary 2.3 and (2.20), we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\left\|f^{(k-1)}-q\right\| = 0. \tag{2.21}$$

*Now,*

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}|H(X_k|X_{k-1})+qg| &= \frac{1}{n}\sum_{k=1}^{n}\left|f^{(k-1)}g_k-qg\right| \\ &= \frac{1}{n}\sum_{k=1}^{n}\left|f^{(k-1)}g_k-f^{(k-1)}g+f^{(k-1)}g-qg\right| \\ &\le \frac{1}{n}\sum_{k=1}^{n}\left|f^{(k-1)}g_k-f^{(k-1)}g\right| + \frac{1}{n}\sum_{k=1}^{n}\left|f^{(k-1)}g-qg\right| \\ &\le \frac{1}{n}\sum_{k=1}^{n}\|g_k-g\| + \frac{\|g\|}{n}\sum_{k=1}^{n}\left|f^{(k-1)}-q\right|. \end{aligned} \tag{2.22}$$

*By (2.14), (2.21) and (2.22, (2.19) follows.* ■

**Remark 2.3.**

*Consider a nonhomogeneous Markov chain $\{X_n,\ n\ge 0\}$ having the transition matrices $\{P_n,\ n\ge 1\}$. From Theorem 2.6, we know that if (2.9) and (2.14) hold, then the entropy rate of this Markov chain exists and*

$$\lim_{n\to\infty}\frac{1}{n}H(X_0,\dots,X_n) = -\sum_i q_i\sum_j p_{ij}\log p_{ij}. \tag{2.23}$$

**Corollary 2.7**

*Consider a finite nonhomogeneous Markov chain $\{X_n,\ n\ge 0\}$ taking values in state space $S=\{1,2,\dots,N\}$ having the transition matrices $P_n=\big(p_n(i,j)\big)$. Let $P=\big(p_{ij}\big)$ be any other stochastic matrix. Suppose that $P$ is ergodic. If $\forall i,j\epsilon\ S$*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\left|p_k(i,j)-p_{ij}\right| = 0 \tag{2.24}$$

*then,*

$$\lim_{n\to\infty}\frac{1}{n}H(X_0,\dots,X_n) = -\sum_i\sum_j \pi_i p_{ij}\log p_{ij}. \tag{2.25}$$

**Proof:**

*Since, if nonhomogeneous Markov chain is finite, then (2.9) is correspondent to (2.24) and (2.14) is correspondent to the following:* $\forall i, j \epsilon S$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left| p_k(i,j) \log p_k(i,j) - p_{ij} \log p_{ij} \right| = 0. \tag{2.26}$$

*By Yang (see (Yang 1998)), (2.24) follows from (2.22). In addition, if a transition matrix* $P$ *is finite, then ergodicity of* $P$ *is correspondent to strong ergodicity of* $P$*. So, by Corollary 2.5) and Remark 2.3, this corollary follows.* ■

## ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (11571142).

## REFERENCES

1. Bowerman, B., David, H.T. and Isaacson, D. (1977). The convergence of Cesaro averages for certain nonstationary Markov chains. *Stochastic Processes and their Applications*, 5(3), 221-230.
2. Cover, T.M. and Thomas, J.A. (2012). *Elements of information theory*, John Wiley & Sons, New York.
3. Galtchouk, L. and Pergamenshchikov, S. (2014). Geometric ergodicity for classes of homogeneous Markov chains. *Stochastic Processes and their Applications*, 124(10), 3362-3391.
4. Huang, C.C., Isaacson, D. and Vinograde, B. (1976). The rate of convergence of certain nonhomogeneous Markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 35(2), 141-146.
5. Iosifescu, M. (1972). On Two Recent Papers on Ergodicity in Nonhomogeneous Markov Chains. *The Annals of Mathematical Statistics*, 43(5), 1732-1736.
6. Isaacson, D. and Seneta, E. (1982). Ergodicity for countable inhomogeneous Markov chains. *Linear Algebra and its Applications*, 48, 37-44.
7. Isaacson, D.L. and Madsen, R.W. (1976). *Markov chains theory and applications*, John Wiley & Sons, New York.
8. Mukhamedov, F. (2013). The Dobrushin ergodicity coefficient and the ergodicity of noncommutative Markov chains. *Journal of Mathematical Analysis and Applications*, 408(1), 364-373.
9. Tan, C.P. (1996). On the weak ergodicity of nonhomogeneous Markov chains. *Statistics & Probability Letters*, 26(4), 293-295.
10. Yang, W. (1998). The Asymptotic Equipartition Property for a Nonhomogeneous Markov Information Source. *Probability in the Engineering and Informational Sciences*, 12(4), 509-518.
11. Yang, W. (2002). Convergence in the Cesàro sense and strong law of large numbers for nonhomogeneous Markov chains. *Linear Algebra and its Applications*, 354(1), 275-288.
12. Yang, W. (2009). Strong law of large numbers for countable nonhomogeneous Markov chains. *Linear Algebra and its Applications*, 430(11), 3008-3018.