Azerbaijan Journal of Mathematics
V. 2, No 2, 2012, July
ISSN 2218-6816

A Class of Small Deviation Theorems for Random Sums of Multivariate Function of mth-Order Markov Chain

Kangkang Wang*, Decai Zong

Abstract. In this paper, the notion of limit relative logarithmic likelihood ratio of stochastic sequence, as a measure of dissimilarity between the joint distribution and the mth-order nonhomogeneous Markov distribution, is introduced. A kind of strong limit theorems represented by inequalities which we call the small deviation theorems for the arbitrary stochastic sequence are obtained by constructing the consistent distribution functions. As corollaries, we obtain some small deviation theorems for the occurrence frequency of the state groups and the harmonic mean of the transitional probabilities of the mth-order nonhomogeneous Markov chain.

Key Words and Phrases: stochastic sequence; limit logarithmic likelihood ratio; small deviation theorem; mth-order nonhomogeneous Markov chain; random sum

2010 Mathematics Subject Classifications: 60F15

1. Introduction

Let $\{X_n, n \geq 0\}$ be an arbitrary stochastic sequence defined on the probability space $(\Omega, \mathcal{F}, P)$ which takes values in an alphabet set $S = \{s_0, s_1, \cdots s_{N-1}\}$, with the joint distribution:

$$P(X_0 = x_0, \cdots, X_n = x_n) = p(x_0, \cdots, x_n) > 0, \quad x_i \in S, \quad 0 \leq i \leq n. \tag{1}$$

We obtain by the definition of the conditional probability

$$p(x_0, \cdots, x_n) = p(x_0) \prod_{k=1}^{n} p_k(x_k | x_0, \cdots, x_{k-1}). \tag{2}$$

Let Q be another probability measure on $(\Omega, \mathcal{F})$, $\{X_n, n \geq 0\}$ be an mth-order nonhomogeneous Markov chain on the measure Q, with the m-dimensional initial distribution and mth-order transition probabilities as follows:

$$q_o(i_0, \cdots, i_{m-1}) = Q(X_0 = i_0, \cdots, X_{m-1} = i_{m-1}), \quad i_0, \cdots, i_{m-1} \in S, \tag{3}$$

*Corresponding author.

$$q_n(j|i_1,\cdots,i_m)=Q_n(X_n=j|X_{n-m}=i_1,\cdots,X_{n-1}=i_m),\quad i_1,\cdots,i_m,j\in S,\ n\geq m. \tag{4}$$

The joint distribution of $\{X_n,n\geq 0\}$ with respect to the measure Q is

$$Q(X_0=x_0,\cdots,X_n=x_n)=q(x_0,\cdots,x_n)=q_o(x_0,\cdots,x_{m-1})\prod_{k=m}^{n}q_k(x_k|x_{k-m},\cdots,x_{k-1}),$$
$$x_i\in S,\quad 0\leq i\leq n. \tag{5}$$

Definition 1. *Let p and q be defined as (1) and (5), $\{\sigma_n,n\geq 0\}$ be a nonnegative increasing stochastic sequence and $\sigma_n\uparrow\infty$. Denote*

$$h(P|Q)=\limsup_{n\to\infty}\frac{1}{\sigma_n}\ln[p(X_0,\cdots,X_{[\sigma_n]})\Big/q_o(X_0,,\cdots,X_{m-1})\prod_{k=m}^{[\sigma_n]}q_k(X_k|X_{k-m},\cdots,X_{k-1})], \tag{6}$$

where $[\sigma_n]$ represents the integral part of σ_n. $h(P|Q)$ is called the sample divergence rate of the measure P relative to the measure Q.

In fact, $h(P|Q)$ is also called the limit relative logarithmic likelihood ratio or asymptotic logarithmic likelihood ratio of $\{X_n,n\geq 0\}$ on the measure P relative to Q. Although $h(P|Q)$ is not an optimum metric between the two probability measures, we nevertheless think of it as a measure of "dissimilarity" between their joint distribution and the mth-order nonhomogeneous Markov distribution. Obviously, $h(P|Q)=0$ if and only if $P=Q$. It will be shown in (20) that $h(P|Q)\geq 0,\ \ a.s.(almostsurely)$ in any case. Hence, $h(P|Q)$ can be used as a random measure of the deviation between the true joint distribution $p(x_0,\cdots,x_n),(n\geq 0)$ and the mth-order nonhomogeneous Markov reference distribution $q_o(x_0,\cdots,x_{m-1})\prod_{k=m}^{n}q_k(x_k|x_{k-m},\cdots,x_{k-1})$. Roughly speaking, this deviation may be regarded as the one between $\{X_n,n\geq 0\}$ under the the measure P and the one under Q. The smaller $h(P|Q)$ is, the smaller the deviation is.

Study for strong limit properties of nonhomogeneous Markov chain is always one of central parts of the limit theory of probability theory. Many scholars have studied the subject until now. Liu and and Yang (see [1]) have studied the asymptotic equipartition properties (AEP) and limit properties of function sequence of nonhomogeneous Markov chain. Liu (see [2]) has discussed the strong limit theorems relative to the geometric average of random transition properties of finite nonhomogeneous Markov chain. Liu and Yang (see [3]) have investigated the strong deviation theorems of nonhomogeneous Markov chain relative to arbitrary stochastic sequence and AEP approximation of nonhomogeneous Markov information source. Yang (see [4]) has furthermore studied a class of small deviation theorems for the sequence of N-valued random variables with respect to mth-order nonhomogeneous Markov chains. Liu (see [5]) has discussed the strong limit theorems for the harmonic mean of random transition probabilities of nonhomogeneous Markov chain. Liu and Wang (see [6]) have proved the strong limit properties for the state couples of nonhomogeneous Markov chain on the random selection system. Wang (see [7]) has studied the AEP and limit theorems for nonhomogeneous Markov chain on the generalized gambling system. Afterward, many scholars (see [11-28]) have studied all kinds of stochastic processes and some limit properties with their applications for nonhomogeneous Markov chains on the generalized gambling system.

Many of practical information sources, such as language and image information, are often mth-order Markov chain, and always nonhomogeneous. mth-order nonhomogeneous Markov chain is a natural generalization of the general nonhomogeneous Markov chain. Hence it is of importance to study the limit properties for the mth-order nonhomogeneous Markov chain in the information theory and the probability theory. Yang and Liu (see [9]) have proved the limit theorem for averages of the functions of $m+1$ variables of mth-order nonhomogeneous Markov chain and the AEP for mth-order nonhomogeneous Markov information source. Wang (see [10]) has discussed the Shannon-McMillan theorems for mth-order nonhomogeneous Markov information source.

In this paper, our aim is to establish a class of small deviation theorems represented by inequalities for the dependent stochastic sequence with respect to mth-order nonhomogeneous Markov chain by introducing the sample relative entropy rate as a measure of deviation between the arbitrary stochastic sequence and the mth-order nonhomogeneous Markov chain. We apply a new type of techniques distinct from that of Liu and Yang (see [4] and [9]) to study the small deviation theorems for arbitrary stochastic sequence. As corollaries, a class of small deviation theorems for the occurrence frequencies of state groups and some limit properties for harmonic mean of the transitional probabilities of the mth-order Markov chain are obtained.

We denote $X_m^n = \{X_m, \cdots, X_n\}$, $i_0^m = \{i_0, \cdots, i_m\}$. x_m^n is the realization of X_m^n.

2. Main Result and Its Proof

Theorem 1. *Let $\{X_n, n \geq 0\}$ be an arbitrary stochastic sequence with the joint distribution (1), $h(P|Q)$ be defined by (6), $\{\sigma_n, n \geq 0\}$ be an increasing nonnegative stochastic sequence. Let $g(x_0, \cdots, x_m)$ be a multivariate real function defined on S^{m+1}. Denote*

$$D(c) = \{\omega : \lim_{n\to\infty} \sigma_n(\omega) = +\infty,\ h(P|Q) \leq c\}. \tag{7}$$

Then

$$\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} \{g(X_{k-m}^k) - E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\} \leq (2\sqrt{c}+c) \sum_{i_0\cdots i_m \in S} |g(i_0, \cdots, i_m)|,$$

$$P-a.s.\quad \omega \in D(c) \tag{8}$$

$$\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} \{g(X_{k-m}^k) - E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\} \geq -2\sqrt{c} \sum_{i_0\cdots i_m \in S} |g(i_0, \cdots, i_m)|,$$

$$P-a.s.\quad \omega \in D(c) \tag{9}$$

where $[\sigma_n]$ represents the integral part of σ_n, and E_Q is the expectation with respect to the measure Q.

Proof. We consider the probability space $(\Omega, \mathcal{F}, P)$. Let λ be an arbitrary real number, and $\delta_i(j)$ be the Kronecker function. Denote

$$\delta_{i_0\cdots i_m}(x_{k-m}^{k}) = \delta_{i_0}(x_{k-m})\cdots\delta_{i_m}(x_k),$$

$$\delta_{i_0\cdots i_{m-1}}(x_{k-m}^{k-1}) = \delta_{i_0}(x_{k-m})\cdots\delta_{i_{m-1}}(x_{k-1}),$$

$$M_k(x_{k-m},\cdots x_k;\lambda) = \left(\frac{1}{1+(\lambda-1)q_k(i_m|i_0^{m-1})}\right)^{\delta_{i_0\cdots i_{m-1}}(x_{k-m}^{k-1})} \lambda^{\delta_{i_0\cdots i_m}(x_{k-m}^{k})} q_k(x_k|x_{k-m}^{k-1}), \quad k \geq m. \tag{10}$$

We construct the following product distribution:

$$\mu(x_0,\cdots,x_n;\lambda) = q_o(x_0,\cdots,x_{m-1})\prod_{k=m}^{n} M_k(x_{k-m},\cdots,x_k;\lambda). \tag{11}$$

By (10) and (11) we have

$$\sum_{x_n\in S}\mu(x_0,\cdots,x_n;\lambda) =$$

$$= \sum_{x_n\in S} q_o(x_0^{m-1})\prod_{k=m}^{n}\left(\frac{1}{1+(\lambda-1)q_k(i_m|i_0^{m-1})}\right)^{\delta_{i_0\cdots i_{m-1}}(x_{k-m}^{k-1})} \lambda^{\delta_{i_0\cdots i_m}(x_{k-m}^{k})} q_k(x_k|x_{k-m}^{k-1})$$

$$= \mu(x_0,\cdots,x_{n-1};\lambda)\sum_{x_n\in S}\left(\frac{1}{1+(\lambda-1)q_n(i_m|i_0^{m-1})}\right)^{\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1})} \lambda^{\delta_{i_0\cdots i_m}(x_{n-m}^{n})} q_n(x_n|x_{n-m}^{n-1})$$

$$= \mu(x_0,\cdots,x_{n-1};\lambda)\frac{1}{[1+(\lambda-1)q_n(i_m|i_0^{m-1})]^{\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1})}}\left(\sum_{x_n=i_m}+\sum_{x_n\neq i_m}\right)$$

$$= \mu(x_0,\cdots,x_{n-1};\lambda)\frac{\lambda^{\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1})} q_n(i_m|x_{n-m}^{n-1}) + 1 - q_n(i_m|x_{n-m}^{n-1})}{[1+(\lambda-1)q_n(i_m|i_0^{m-1})]^{\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1})}}. \tag{12}$$

When $\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1}) = 0$, we obtain from (12) that

$$\sum_{x_n\in S}\mu(x_0,\cdots,x_n;\lambda) =$$

$$= \mu(x_0,\cdots,x_{n-1};\lambda)\frac{q_n(i_m|x_{n-m}^{n-1}) + 1 - q_n(i_m|x_{n-m}^{n-1})}{[1+(\lambda-1)q_n(i_m|i_0^{m-1})]^0} = \mu(x_0,\cdots,x_{n-1};\lambda). \tag{13}$$

When $\delta_{i_0\cdots i_{m-1}}(x_{n-m}^{n-1})=1$, we acquire from (12) that

$$\sum_{x_n\in S}\mu(x_0,\cdots,x_n;\lambda)$$

$$=\mu(x_0,\cdots,x_{n-1};\lambda)\frac{1+(\lambda-1)q_n(i_m|i_0^{m-1})}{1+(\lambda-1)q_n(i_m|i_0^{m-1})}=\mu(x_0,\cdots,x_{n-1};\lambda). \tag{14}$$

Therefore, $\mu(x_0,\cdots,x_n;\lambda)$, $n=1,2,\cdots$ are a family of consistent distribution functions on S^{n+1}. Denote

$$T_n(\lambda,\omega)=\frac{\mu(X_0,\cdots,X_n;\lambda)}{p(X_0,\cdots,X_n)}. \tag{15}$$

By (5), (11) and (15), we can rewrite (15) as

$$T_n(\lambda,\omega)=\lambda^{\sum_{k=m}^{n}\delta_{i_0\cdots i_m}(X_{k-m}^{k})}\prod_{k=m}^{n}[\frac{1}{1+(\lambda-1)q_k(i_m|i_0^{m-1})}]^{\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})}$$

$$\cdot q_o(X_0^{m-1})\prod_{k=m}^{n}q_k(X_k|X_{k-m}^{k-1})\Big/p(X_0,\cdots,X_n). \tag{16}$$

Since μ and p are two probability measures, we can know that $T_n(\lambda,\omega)$ is a nonnegative sup-martingale from Doob's martingale convergence theorem (see [8]). Moreover,

$$\lim_{n\to\infty}T_n(\lambda,\omega)=T_\infty(\lambda,\omega)<\infty. \qquad P-a.s. \tag{17}$$

By (7) and (17) we have

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\ln T_{[\sigma_n]}(\lambda,\omega)\le 0. \qquad P-a.s. \quad \omega\in D(c) \tag{18}$$

By (16) and (18) we can obtain

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\left\{\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_m}(X_{k-m}^{k})\ln\lambda-\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})\ln[1+(\lambda-1)q_k(i_m|i_0^{m-1})]=\right.$$

$$\left.-\ln\left[p(X_0,\cdots,X_{[\sigma_n]})\Big/q_o(X_0^{m-1})\prod_{k=m}^{[\sigma_n]}q_k(X_k|X_{k-m}^{k-1})\right]\right\}\le 0. P-a.s. \quad \omega\in D(c) \tag{19}$$

Letting $\lambda=1$ in (19), by (6) we can know

$$h(P|Q)\ge\liminf_{n\to\infty}\frac{1}{\sigma_n}\ln\left[\frac{p(X_0,\cdots,X_{[\sigma_n]})}{q_o(X_0,\cdots,X_{m-1})\prod_{k=m}^{[\sigma_n]}q_k(X_k|X_{k-m}^{k-1})}\right]\ge 0. \quad P-a.s. \quad \omega\in D(c) \tag{20}$$

By (6), (7) and (19), taking into account the property of superior limit, we can write

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\{\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_m}(X_{k-m}^k)\ln\lambda - \sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})\ln[1+(\lambda-1)q_k(i_m|i_0^{m-1})]\}\le h(P|Q)\le c.$$

$$P-a.s. \qquad \omega\in D(c) \qquad . \tag{21}$$

In the case $\lambda>1$, dividing two sides of (21) by $\ln\lambda$, we have

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\left\{\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_m}(X_{k-m}^k)-\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})\frac{\ln[1+(\lambda-1)q_k(i_m|i_0^{m-1})]}{\ln\lambda}\right\}\le\frac{c}{\ln\lambda}.$$

$$P-a.s. \qquad \omega\in D(c) \tag{22}$$

By (22), using the inequalities $1-1/x\le\ln x\le x-1,(x>0)$, $0\le\delta_{i_0,\cdots,i_{m-1}}(X_{k-m}^{k-1};\omega)\le 1$ and the properties of superior limit

$$\limsup_{n\to\infty}(a_n-b_n)\le d\Rightarrow\limsup_{n\to\infty}(a_n-c_n)\le\limsup_{n\to\infty}(b_n-c_n)+d,$$

we can write

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\\
\le\;&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})\frac{\ln[1+(\lambda-1)q_k(i_m|i_0^{m-1})]}{\ln\lambda}\\
&\qquad\qquad -\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})\}+\frac{c}{\ln\lambda}\\
\le\;&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})[\frac{(\lambda-1)q_k(i_m|i_0^{m-1})}{\ln\lambda}-q_k(i_m|i_0^{m-1})]\}+\frac{c}{\ln\lambda}\\
=\;&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})\left(\frac{\lambda-1}{\ln\lambda}-1\right)+\frac{c}{\ln\lambda}\\
\le\;&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})\left(\frac{\lambda-1}{1-1/\lambda}-1\right)+\frac{c}{1-1/\lambda}\\
\le\;&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})(\lambda-1)+c+\frac{c}{\lambda-1}
\end{aligned}$$

$$\leq \quad \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} (\lambda - 1) + c + \frac{c}{\lambda - 1} = (\lambda - 1) + c + \frac{c}{\lambda - 1}.$$

$$P - a.s. \qquad \omega \in D(c) \tag{23}$$

It is easy to show that in the case $c > 0$, the function $f(\lambda) = (\lambda - 1) + c + \frac{c}{\lambda - 1} (\lambda > 1)$ attains its smallest value $f(1 + \sqrt{c}) = 2\sqrt{c} + c$ at $\lambda = 1 + \sqrt{c}$. Hence, letting $\lambda = 1 + \sqrt{c}$ in (23), we have

$$\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} [\delta_{i_0 \cdots i_m}(X_{k-m}^k) - \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) q_k(i_m | i_0^{m-1})] \leq 2\sqrt{c} + c.$$

$$P - a.s. \qquad \omega \in D(c) \tag{24}$$

In the case $c = 0$, (24) also follows from (23) by choosing $\lambda_i \to 1 + (i \to \infty)$.

In the case $0 < \lambda < 1$, dividing two sides of (21) by $\ln \lambda$, we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} [\delta_{i_0 \cdots i_m}(X_{k-m}^k) - \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) q_k(i_m | i_0^{m-1})] \\
\geq \quad &\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) \left(\frac{\ln[1 + (\lambda - 1) q_k(i_m | i_0^{m-1})]}{\ln \lambda} - q_k(i_m | i_0^{m-1}) \right) + \frac{c}{\ln \lambda} \\
\geq \quad &\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) q_k(i_m | i_0^{m-1}) \left(\frac{\lambda - 1}{1 - 1/\lambda} - 1 \right) + \frac{c}{\lambda - 1} \\
\geq \quad &\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) (\lambda - 1) + \frac{c}{\lambda - 1} \geq (\lambda - 1) + \frac{c}{\lambda - 1}
\end{aligned}$$

$$P - a.s. \qquad \omega \in D(c) \qquad . \tag{25}$$

It is easy to show that in the case $c > 0$, the function $h(\lambda) = (\lambda - 1) + \frac{c}{\lambda - 1} (0 < \lambda < 1)$ attains its largest value $h(1 - \sqrt{c}) = -2\sqrt{c}$ at $\lambda = 1 - \sqrt{c}$. Hence, letting $\lambda = 1 - \sqrt{c}$ in (25), we obtain

$$\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} [\delta_{i_0 \cdots i_m}(X_{k-m}^k) - \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) q_k(i_m | i_0^{m-1})] \geq -2\sqrt{c}.$$

$$P - a.s. \qquad \omega \in D(c) \tag{26}$$

In the case $c = 0$, (26) also follows from (25) by choosing $\lambda_i \to 1 - (i \to \infty)$.

It follows from (24) and (26) that for any real function $g(i_0, \cdots, i_m)$,

$$\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=m}^{[\sigma_n]} g(i_0, \cdots, i_m) [\delta_{i_0 \cdots i_m}(X_{k-m}^k) - \delta_{i_0 \cdots i_{m-1}}(X_{k-m}^{k-1}) q_k(i_m | i_0^{m-1})]$$

$$\leq \quad (2\sqrt{c}+c)|g(i_0,\cdots,i_m)|, \qquad P-a.s. \quad \omega\in D(c) \tag{27}$$

$$\begin{aligned}&\liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]} g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\\ \geq\quad & -2\sqrt{c}|g(i_0,\cdots,i_m)|. \qquad P-a.s. \quad \omega\in D(c)\end{aligned} \tag{28}$$

Note that

$$\begin{aligned}& g(X_{k-m},\cdots,X_k)-E_Q(g(X_{k-m},\cdots,X_k)|X_{k-m},\cdots,X_{k-1})\\ =\quad & g(X_{k-m},\cdots,X_k)-\sum_{i_m\in S}g(X_{k-m},\cdots X_{k-1},i_m)q_k(i_m|X_{k-m},\cdots,X_{k-1})\\ =\quad & \sum_{i_0\cdots i_m\in S}\delta_{i_0\cdots i_m}(X_{k-m},\cdots,X_k)g(i_0,\cdots,i_m)\\ & \qquad -\sum_{i_0\cdots i_m\in S}\delta_{i_0\cdots i_{m-1}}(X_{k-m},\cdots,X_{k-1})g(i_0,\cdots,i_m)q_k(i_m|i_0,\cdots,i_{m-1})\\ =\quad & \sum_{i_0\cdots i_m\in S}\delta_{i_0\cdots i_{m-1}}(X_{k-m},\cdots,X_{k-1})g(i_0,\cdots,i_m)[\delta_{i_m}(X_k)-q_k(i_m|i_0,\cdots,i_{m-1})]\\ =\quad & \sum_{i_0\cdots i_m\in S}g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})].\end{aligned} \tag{29}$$

By virtue of the properties of superior limit and inferior limit, combing (27)-(29), we obtain

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{g(X_{k-m}^k)-E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\}\\ =\quad & \limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\sum_{i_0\cdots i_m\in S}g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\\ \leq\quad & \sum_{i_0\cdots i_m\in S}\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\\ \leq\quad & (2\sqrt{c}+c)\sum_{i_0\cdots i_m\in S}|g(i_0,\cdots,i_m)|,\end{aligned} \tag{30}$$

$$\begin{aligned}&\liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{g(X_{k-m}^k)-E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\}\\ =\quad & \liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\sum_{i_0\cdots i_m\in S}g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\end{aligned}$$

$$
\begin{aligned}
&\geq \sum_{i_0\cdots i_m\in S}\liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]} g(i_0,\cdots,i_m)[\delta_{i_0\cdots i_m}(X_{k-m}^k)-\delta_{i_0\cdots i_{m-1}}(X_{k-m}^{k-1})q_k(i_m|i_0^{m-1})]\\
&\geq -2\sqrt{c}\sum_{i_0\cdots i_m\in S}|g(i_0,\cdots,i_m)|. \qquad (31)
\end{aligned}
$$

(8), (9) follow from (30) and (31), respectively. The proof is finished. ◀

Corollary 1. *Let $\{X_n, n\geq 0\}$ be an mth-order nonhomogeneous Markov chain with the m dimensional initial distribution (3) and mth-order transitional probabilities (4). Let $g(x_0,\cdots,x_m)$ be a multivariate real function defined on S^{m+1}. Then*

$$
\lim_{n\to\infty}\frac{1}{n}\sum_{k=m}^{n}\{g(X_{k-m},\cdots,X_k)-E(g(X_{k-m},\cdots,X_k)|X_{k-m},\cdots,X_{k-1})\}=0. \quad P{-}a.s. \qquad (32)
$$

Proof. Letting $\sigma_n = n$, $n\geq 1$, $P\equiv Q$ in Theorem 1, we obtain $h(P|Q)\equiv 0$, $E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1}) = E(g(X_{k-m}^k)|X_{k-m}^{k-1})$ at the moment. This implies that $D(0)=\Omega$ when $c=0$. Let $c=0$. Then (32) follows from (8) and (9).◀

3. Some small deviation theorems for the occurrence frequency of the state groups

In equation (4), if for all $n\geq m$,

$$
\begin{aligned}
&Q_n(X_n=j|X_{n-m}=i_1,\cdots,X_{n-1}=i_m)\\
=\ &Q(X_n=j|X_{n-m}=i_1,\cdots,X_{n-1}=i_m), \qquad \forall i_1,\cdots,i_m,\ j\in S
\end{aligned}
$$

we call $\{X_n, n\geq 0\}$ an mth-order homogeneous Markov chain on the measure Q.

Corollary 2. *Let $\{X_n, n\geq 0\}$ be an arbitrary stochastic sequence with the joint distribution (1), $h(P|Q)$ be defined as before. Denote by $S_n(j_0,\cdots,j_m)$ the number of the state groups $(j_0,\cdots,j_m)$ occurring in $(X_0,\cdots,X_m),(X_1,\cdots,X_{m+1}),\cdots,(X_{n-m},\cdots,X_n)$, and by $S_n(j_0,\cdots,j_{m-1})$ the number of the state groups $(j_0,\cdots,j_{m-1})$ occurring in $(X_0,\cdots,X_{m-1}),(X_1,\cdots,X_m),\cdots,(X_{n-m},\cdots,X_{n-1})$. That is,*

$$
S_n(j_0,\cdots,j_m)=\sum\nolimits_{k=m}^{n}\delta_{j_0\cdots j_m}(X_{k-m},\cdots X_k), \qquad (33)
$$

$$
S_n(j_0,\cdots,j_{m-1})=\sum\nolimits_{k=m}^{n}\delta_{j_0\cdots j_{m-1}}(X_{k-m},\cdots X_{k-1}). \qquad (34)
$$

We set

$$
L(c)=\{\omega: h(P|Q)\leq c\}. \qquad (35)
$$

Then

$$\limsup_{n\to\infty}\frac{1}{n}\{S_n(j_0,\cdots,j_m)-S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})\}\le(2\sqrt{c}+c),$$

$$P-a.s. \qquad \omega\in L(c) \tag{36}$$

$$\liminf_{n\to\infty}\frac{1}{n}\{S_n(j_0,\cdots,j_m)-S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})\}\ge-2\sqrt{c}.$$

$$P-a.s. \qquad \omega\in L(c) \tag{37}$$

Proof. Letting $\sigma_n=n, n\ge 0$, $g(X_{k-m}^k)=\delta_{j_0\cdots j_m}(X_{k-m}^k)$, $q_k(X_k|X_{k-m}^{k-1})=q(X_k|X_{k-m}^{k-1}),\ k\ge m$ in Theorem 1, we can obtain $L(c)=D(c)$ and write

$$\begin{aligned}
&\sum_{k=m}^{n}\{g(X_{k-m}^k)-E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\}\\
=\;&\sum_{k=m}^{n}\{\delta_{j_0\cdots j_m}(X_{k-m}^k)-E_Q(\delta_{j_0\cdots j_m}(X_{k-m}^k)|X_{k-m}^{k-1})\}\\
=\;&\sum_{k=m}^{n}\{\delta_{j_0\cdots j_m}(X_{k-m}^k)-\sum_{x_k\in S}\delta_{j_0\cdots j_m}(X_{k-m}^{k-1},x_k)q(x_k|X_{k-m}^{k-1})\}\\
=\;&\sum_{k=m}^{n}\{\delta_{j_0\cdots j_m}(X_{k-m}^k)-\delta_{j_0\cdots j_{m-1}}(X_{k-m}^{k-1})q(j_m|X_{k-m}^{k-1})\}\\
=\;&\sum_{k=m}^{n}\left\{\delta_{j_0\cdots j_m}(X_{k-m}^k)-\delta_{j_0\cdots j_{m-1}}(X_{k-m}^{k-1})q(j_m|j_0^{m-1})\right\}\\
=\;&S_n(j_0,\cdots,j_m)-S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})
\end{aligned}\tag{38}$$

and

$$\sum_{i_0\cdots i_m\in S}|g(i_0,\cdots,i_m)|=\sum_{i_0\cdots i_m\in S}|\delta_{j_0\cdots j_m}(i_0,\cdots i_m)|=1. \tag{39}$$

Taking into account (38) and (39), we obtain (36), (37) from (8), (9) immediately. ◄

Corollary 3. *Let $\{X_n,n\ge0\}$ be an arbitrary stochastic sequence with the joint distribution (1), $h(P|Q)$ be defined as before. Denote by $S_n(j_m)$ the number of the state groups j_m occurring in $X_0,\cdots,X_n$. That is,*

$$S_n(j_m)=\sum_{k=0}^{n}\delta_{j_m}(X_k). \tag{40}$$

Then

$$\limsup_{n\to\infty}\frac{1}{n}\{S_n(j_m)-\sum_{j_0,\cdots,j_{m-1}\in S}S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})\}\le(2\sqrt{c}+c)N^m,$$

$$P - a.s. \qquad \omega \in L(c) \tag{41}$$

$$\liminf_{n\to\infty} \frac{1}{n}\{S_n(j_m) - \sum_{j_0,\cdots,j_{m-1}\in S} S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})\} \geq -2\sqrt{c}N^m.$$

$$P - a.s. \qquad \omega \in L(c) \tag{42}$$

Proof. Letting $\sigma_n = n, n \geq 0$, $g(X_{k-m}^k) = \delta_{j_m}(X_k)$, $q_k(X_k|X_{k-m}^{k-1}) = q(X_k|X_{k-m}^{k-1})$ in Theorem 1, we can write

$$\begin{aligned}
&\sum_{k=m}^{n}\{g(X_{k-m}^k) - E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\}\\
= &\sum_{k=m}^{n}\{\delta_{j_m}(X_k) - E_Q(\delta_{j_m}(X_k)|X_{k-m}^{k-1})\}\\
= &\sum_{k=m}^{n}\{\delta_{j_m}(X_k) - \sum_{x_k\in S}\delta_{j_m}(x_k)q(x_k|X_{k-m}^{k-1})\}\\
= &\sum_{k=m}^{n}\{\delta_{j_m}(X_k) - \sum_{j_0\cdots j_{m-1}\in S}\sum_{x_k\in S}\delta_{j_0\cdots j_{m-1}}(X_{k-m}^{k-1})\delta_{j_m}(x_k)q(x_k|j_0^{m-1})\}\\
= &\sum_{k=m}^{n}\{\delta_{j_m}(X_k) - \sum_{j_0\cdots j_{m-1}\in S}\delta_{j_0\cdots j_{m-1}}(X_{k-m}^{k-1})q(j_m|j_0^{m-1})\}\\
= &\sum_{k=m}^{n}\delta_{j_m}(X_k) - \sum_{j_0\cdots j_{m-1}\in S}q(j_m|j_0^{m-1})\sum_{k=m}^{n}\delta_{j_0\cdots j_{m-1}}(X_{k-m}^{k-1})\\
= &S_n(j_m) - \sum_{k=0}^{m-1}\delta_{j_m}(X_k) - \sum_{j_0\cdots j_{m-1}\in S}S_n(j_0,\cdots,j_{m-1})q(j_m|j_0^{m-1})
\end{aligned} \tag{43}$$

and

$$\sum_{i_0\cdots i_m\in S}|g(i_0,\cdots,i_m)| = \sum_{i_0\cdots i_m\in S}|\delta_{j_m}(i_m)| = \sum_{i_0\cdots i_{m-1}\in S}1 = N^m. \tag{44}$$

Combining (43), (44) with (8) and (9), we obtain (41), (42) immediately.◄

We present a small deviation theorem for harmonic mean of the transitional probability of the mth-order Markov chain as follows:

Theorem 2. *Let $\{X_n, n \geq 0\}$ be an arbitrary stochastic sequence with the joint distribution (1), $h(P|Q)$ be defined by (6), $\{\sigma_n, n \geq 0\}$ be an increasing nonnegative stochastic sequence. Then*

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}q(X_k|X_{k-m}^{k-1})^{-1} \leq N + (2\sqrt{c}+c)\sum_{i_0\cdots i_m\in S}q(i_m|i_0^{m-1})^{-1},$$

$$P-a.s.\quad \omega\in D(c), \tag{45}$$

$$\liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}q(X_k|X_{k-m}^{k-1})^{-1}\geq N-2\sqrt{c}\sum_{i_0\cdots i_m\in S}q(i_m|i_0^{m-1})^{-1},$$

$$P-a.s.\quad \omega\in D(c). \tag{46}$$

Proof. Letting $g(X_{k-m}^k)=q(X_k|X_{k-m}^{k-1})^{-1}$ in Theorem 1, by (8) we can write

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{g(X_{k-m}^k)-E_Q(g(X_{k-m}^k)|X_{k-m}^{k-1})\}$$

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{q(X_k|X_{k-m}^{k-1})^{-1}-E_Q(q(X_k|X_{k-m}^{k-1})^{-1}|X_{k-m}^{k-1})$$

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{q(X_k|X_{k-m}^{k-1})^{-1}-\sum_{x_k\in S}q(x_k|X_{k-m}^{k-1})^{-1}q(x_k|X_{k-m}^{k-1})\}$$

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}\{q(X_k|X_{k-m}^{k-1})^{-1}-N\}=$$

$$=\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}q(X_k|X_{k-m}^{k-1})^{-1}-\limsup_{n\to\infty}\frac{\sigma_n-m+1}{\sigma_n}N$$

$$=\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}q(X_k|X_{k-m}^{k-1})^{-1}-N$$

$$\leq (2\sqrt{c}+c)\sum_{i_0\cdots i_m\in S}q(i_m|i_0^{m-1})^{-1}. \tag{47}$$

Analogously, from (9) we can get

$$\liminf_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=m}^{[\sigma_n]}q(X_k|X_{k-m}^{k-1})^{-1}-N\geq -2\sqrt{c}\sum_{i_0\cdots i_m\in S}q(i_m|i_0^{m-1})^{-1}. \tag{48}$$

(45), (46) follow from (47) and (48), respectively.◀

Acknowledgements. Authors would like to thank the referee for his valuable suggestions. The work is supported by the Project of Higher Schools' Natural Science Basic Research of Jiangsu Province of China (09KJD110002). Kangkang Wang is the corresponding author.

References

[1] W. Liu, W.G. Yang. An extension of Shannon-Mcmillan theorem and some limit properties for nonhomogeneous Markov chains. Stochastic. Process. Appl. 61(1): 129-145, 1996.

[2] W. Liu. Some strong limit theorems relative to the geometric average of random transition probabilities of arbitrary finite nonhomogeneous Markov chains. Stat. Probab. Letts. 21(1): 77-83 1994.

[3] W. Liu, W.G. Yang. The Markov approximation of the sequences of N-valued random variables and a class of small deviation theorems. Stochastic. Process. Appl. 89(1): 117-130, 2000.

[4] W.G. Yang. A class of small deviation theorems for the sequence of N-valued random variabls with respect to mth-order nonhomogeneous Markov chains[J]. Acta. Mathematica. Scientia. 29(2): 517-527. 2009. (in Chinese)

[5] W. Liu. A strong limit theorem for the harmonic mean of random transition probabilities of finite nonhomogeneous Markov chains[J]. Acta. Mathematica. Scientia. 20(1): 81-84, 2000.(in Chinese)

[6] Z.Z. Wang, W. Liu. A strong limit theorem on random selection forcountable nonhomogeneous Markov chains. Chinese J. Math. 24(2): 187-197, 1996.

[7] K. K. Wang. A Class of strong limit theorems for countable nonhomogeneous Markov chains on the generalized gambling system. Czechoslovak Mathematical Journal. 59 (1): 23-37, 2009.

[8] J.L. Doob. stochastic Process. Wiley, New York, 1953, USA.

[9] W.G. Yang, W. Liu. The asymptotic equipartition property for Mth-order nonhomogeneous Markov information source. IEEE Trans.Inform.Theory. 50(3): 3326-3330, 2004.

[10] K. K. Wang. Some Research on Shannon-McMillan Theorems for mth-order nonhomogeneous Markov information source. Stochastic Analysis and Applications. 27(6): 1117-1128, 2009.

[11] Z.R. Xu, Y.J. Zhu. Flow controled quene with negative customers and preemptive priority. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24(4): 400-404, 2010.

[12] K.K. Wang, D.C. Zong, A class of strong limit theorems for random sum of Cantor-like random sequence on gambling system. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24(3): 305-308, 2010.

[13] Z.R. Xu, M.J. Li, Geo Geo 1 queue model with RCH strategy of negative customers and single vacation. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25(2): 191-194, 2011.

[14] K.K. Wang, and D.C. Zong, A class of strong deviation theorems on generalized gambling system for the sequence of arbitrary continuous random variables. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25 (2): 195-199, 2011.

[15] K.K. Wang, H. Ye and Y. Ma, A class of strong deviation theorems for multivariate function sequence of mth-order countable nonhomogeneous Markov chains. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25 (1): 93-96, 2011

[16] K.K. Wang, A class of local strong limit theorems for Markov chains field on arbitrary Bethe tree. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (2): 205-209, 2010.

[17] K.K. Wang, and D.C. Zong, Strong limit theorems of mth-order nonhomogeneous Markov chains on fair gambling system . J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (4): 410-413, 2010.

[18] K.K. Wang, and F. Li, Strong deviation theorems for the sequence of arbitrary random variables with respect to product distribution in random selection system. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (1): 91-94. 2010.

[19] K.K. Wang, and F. Li, A class of Shannon-McMillan theorems for mth-order nonhomogeneous Markov information source on generalized gambling system. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25 (4): 396-400. 2011

[20] K.K. Wang, and F. Li, Y. Ma, A class of Shannon-McMillan theorems for nonhomogeneous Markov information source on random selection system. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25(3): 299-302. 2011

[21] S.H. Li and Z.G. Zhou, Shrunken estimators of seemingly unrelated regression model. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (2): 197-200. 2010

[22] Y.X. Yuan and J.S. Jiang, A direct updating method for the stiffness matrix. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (2): 193-196. 2010

[23] M. Hu, Feasibility of classification on two kinds of 2-person noncooperative finite games. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (4): 405-409, 2010

[24] X.Q. Hua, D.P. Xu, Two groups of three-level high-precision explicit schemes for dispersive equation. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24(4): 395-399. 2010

[25] Y.X. Yuan and J.S. Jiang, Generalized inverse eigenvalue problems for tridiagonal symmetric matrices. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (1): 88-90. 2010

[26] J.E. Wu and T.C. Wu, Regularized Symmlq for solving Symm integral equation. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25(3): 294-298. 2011

[27] S.H. Li and Z.G. Zhou, Linear minimax estimators of estimable function in multiple Gauss-Markov model with matrix loss function. J. Jiangsu Univ. Sci-tech. Nat. Sci. 24 (1): 95-98. 2010

[28] C.J. Fang, D.C. Zong Strong deviation theorems for the stochastic sequence on the Poisson distribution in generalized gambling systems. J. Jiangsu Univ. Sci-tech. Nat. Sci. 26 (1): 100-104. 2012

Kangkang Wang
School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China
E-mail: wkk.cn@126.com

Decai Zong
Department of Computer Science and Engineering, Changshu Institute of Technology, Changshu 215500, China

Received 21 September 2011
Accepted 9 March 2012