Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2013, Article ID 645151, 6 pages
http://dx.doi.org/10.1155/2013/645151

Research Article

The Central Limit Theorem for mth-Order Nonhomogeneous Markov Information Source

Huilin Huang

College of Mathematics and Information Science, Wenzhou University, Zhejiang 325035, China

Correspondence should be addressed to Huilin Huang; huilin_huang@sjtu.org

Received 22 August 2013; Accepted 17 October 2013

Academic Editor: Shein-chung Chow

We prove a central limit theorem for mth-order nonhomogeneous Markov information source by using the martingale central limit theorem under the condition of convergence of transition probability matrices for nonhomogeneous Markov chain in Cesàro sense.

1. Introduction

Let $X = \{X_n, n \geq 0\}$ be an arbitrary information source taking values on alphabet set $S = \{1, 2, \ldots, N\}$ with the joint distribution

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = p(x_0, x_1, \ldots, x_n), \quad x_i \in S, \tag{1}$$

for $\forall 0 \leq i \leq n,\ n \geq 0$. If $X = \{X_n, n \geq 0\}$ is an mth-order nonhomogeneous Markov information source, then, for $n \geq m$,

$$\begin{aligned} &P(X_n = x_n \mid X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) \\ &\quad = P(X_n = x_n \mid X_{n-m} = x_{n-m}, \\ &\qquad X_{n-m+1} = x_{n-m+1}, \ldots, X_{n-1} = x_{n-1}). \end{aligned} \tag{2}$$

Denote

$$\begin{aligned} &\mu(i_0, i_1, \ldots, i_{m-1}) \\ &\quad = P(X_0 = i_0, X_1 = i_1, \ldots, X_{m-1} = i_{m-1}), \end{aligned} \tag{3}$$

$$\begin{aligned} &p_n(j \mid i_1, i_2, \ldots, i_m) \\ &\quad = P(X_n = j \mid X_{n-m} = i_1, \ldots, X_{n-1} = i_m), \end{aligned} \tag{4}$$

where $\mu(i_0, i_1, \ldots, i_{m-1})$ and $p_n(j \mid i_1, i_2, \ldots, i_m)$ are called the m-dimensional initial distribution and the mth-order transition probabilities, respectively. Moreover,

$$P_n = (p_n(j \mid i_1, i_2, \ldots, i_m)) \tag{5}$$

are called the mth-order transition probability matrices. In this case,

$$\begin{aligned} &p(x_0, x_1, \ldots, x_n) \\ &\quad = \mu(x_0, x_1, \ldots, x_{m-1}) \prod_{k=m}^{n} p_k(x_k \mid x_{k-m}, \ldots, x_{k-1}). \end{aligned} \tag{6}$$

There are many of practical information sources, such as language and image information, which are often mth-order Markov information sources and always nonhomogeneous. So it is very important to study the limit properties for the mth-order nonhomogeneous Markov information sources in information theory. Yang and Liu [1] proved the strong law of large numbers and the asymptotic equipartition property with convergence in the sense of a.s. the mth-order nonhomogeneous Markov information sources. But the problem about the central limit theorem for the mth-order nonhomogeneous Markov information sources is still open.

The central limit theorem (CLT) for additive functionals of stationary, ergodic Markov information source has been studied intensively during the last decades [2–9]. Nearly fifty years ago, Dobrushin [10, 11] proved an important central

limit theorem for nonhomogeneous Markov information resource in discrete time. After Dobrushin's work, some refinements and extensions of his central limit theorem, some of which are under more stringent assumptions, were proved by Statuljavicius [12] and Sarymsakov [13]. Based on Dobrushin's work, Sethuraman and Varadhan [14] gave shorter and different proof elucidating more the assumptions by using martingale approximation. Those works only consider the case about 1th-order nonhomogeneous Markov chain. In this paper, we come to study the central limit theorem for mth-order nonhomogeneous Markov information sources in Cesàro sense.

Let $X = \{X_n, n \geq 0\}$ be an mth-order nonhomogeneous Markov information source which is taking values in state space $S = \{1, 2, \ldots, N\}$ with initial distribution of (3) and mth order transition probability matrices (5). Denote

$$X_m^n = \{X_m, X_{m+1}, \ldots, X_n\}. \tag{7}$$

We also denote the realizations of X_m^n by x_m^n. We denote the mth-order transition matrix at step k by

$$P_k = \left(p_k\left(j \mid i_1^m\right)\right), \quad j \in S,\ i_1^m \in S^m, \tag{8}$$

where $P_k(j \mid i_1^m) = P(X_k = j \mid X_{k-m}^{k-1} = i_1^m)$.

For an arbitrary stochastic square matrix A whose elements are $A_{i,j}$, we will set the ergodic δ-coefficient equal to

$$\delta(A) = \sup_{i,j\in S} \sum_{k\in S} \left[A_{i,k} - A_{j,k}\right]^+, \tag{9}$$

where $[a]^+ = \max\{0, a\}$. Now we extend this idea to the mth-order stochastic matrix Q whose elements are $q(i_1^m, j) = q(j \mid i_1^m)$, and we will introduce the ergodic δ-coefficient equal to

$$\delta(Q) = \sup_{i_1^m, j_1^m \in S^m} \sum_{k\in S} \left[q\left(i_1^m, k\right) - q\left(j_1^m, k\right)\right]^+. \tag{10}$$

Now we define another stochastic matrix as follows:

$$\overline{P} = \left(\overline{p}\left(j_1^m \mid i_1^m\right)\right) \quad i_1^m, j_1^m \in S^m, \tag{11}$$

where

$$\overline{p}\left(j_1^m \mid i_1^m\right) = \begin{cases} p\left(j_m \mid i_1^m\right), & \text{as } j_v = i_{v+1}, \\ & v = 1, 2, \ldots, m-1; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

$\overline{P}$ is called the m-dimensional stochastic matrix determined by the mth-order transition matrix.

Let $S_n(i_1^m) = S_n(i_1, i_2, \ldots, i_m)$ be the number of $(i_1, i_2, \ldots, i_m)$ in the sequence of $X_0^{m-1}, X_1^m, \ldots, X_{n-m}^{n-1}$; that is,

$$S_n\left(i_1^m\right) = \sum_{k=m}^{n} I\left\{X_{k-m}^{k-1} = i_1^m\right\}. \tag{13}$$

Lemma 1 (see [1]). *Let $X = \{X_n, n \geq 0\}$ be an mth-order nonhomogeneous Markov information source which is taking values in state space $S = \{1, 2, \ldots, N\}$ with initial distribution of* (3) *and mth-order transition probability matrices* (5). *$S_n(i_1^m)$ is defined as* (13). *Let $P = (p(j \mid i_1^m))$ be another m-order transition matrix, and let $\overline{P}$ be the m-dimensional stochastic matrix determined by the mth-order transition matrix P, that is, $\pi = \pi\overline{P}$. Suppose that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=m}^{n} \left|p_k\left(j \mid i_1^m\right) - p\left(j \mid i_1^m\right)\right| = 0, \quad \forall j \in S,\ i_1^m \in S^m. \tag{14}$$

Then one has

$$\lim_{n\to\infty} \frac{S_n\left(i_1^m\right)}{n} = \pi\left(i_1^m\right), \quad a.s. \tag{15}$$

2. Statement of the Main Result

Let $f(x_0^m)$ be any Borel function defined on product space S^{m+1}. Denote

$$W_n = \sum_{k=m}^{n} D_k, \tag{16}$$

where

$$D_k = f\left(X_{k-m}^k\right) - E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right], \quad D_{m-1} = 0. \tag{17}$$

Obviously, $\{W_n, \mathscr{F}_n, n \geq m\}$ is a martingale, so that $\{D_n, \mathscr{F}_n, n \geq m\}$ is the associated martingale difference sequence. Denote

$$S_n = \sum_{k=m}^{n} f\left(X_{k-m}^k\right). \tag{18}$$

Our main result is describing conditions on X and f under in which the central limit theorem holds for the stochastic sequence S_n.

Theorem 2. *Let $X = \{X_n, n \geq 0\}$ be an mth-order nonhomogeneous Markov information source which is taking values in state space $S = \{1, 2, \ldots, N\}$ with initial distribution of* (3) *and mth-order transition probability matrices* (5). *Let $P = (P(j \mid i_1^m))_{j\in S, i_1^m \in S^m}$ be an m-th order transition matrix. Let f be any function defined on the state space S^{m+1} and let $\{S_n, n \geq 0\}$ be defined as* (13). *If* (14) *holds and the sequence of δ-coefficients for the mth-order stochastic matrices $\{P_k, k \geq m\}$ satisfies that*

$$\lim_{n\to\infty} \frac{\sum_{k=m}^{n} \delta\left(P_k\right)}{\sqrt{n}} = 0, \tag{19}$$

then one has

$$\frac{S_n}{\sqrt{n}\sigma} \overset{D}{\Rightarrow} N(0, 1), \tag{20}$$

where $\overset{D}{\Rightarrow}$ *denotes the convergence in distribution and*

$$\sigma^2 = \sum_{i_1^m \in S^m} \pi(i_1^m) \left\{ \sum_{j\in S} f^2(i_1^m, j)\, p(j \mid i_1^m) - \left[\sum_{j\in S} f(i_1^m, j)\, p(j \mid i_1^m) \right]^2 \right\} > 0. \tag{21}$$

Remark 3. The sequence $\{P_k = (p_k(j \mid i_1^m)), k \geq m\}$ is said to converge in the Cesàro sense to constant matrix $P = (p(j \mid i_1^m))$ if (14) holds.

3. Proof of Theorem 2

Let $\{\Omega, \mathscr{F}, \mathbf{P}\}$ be a probability space and let $\{M_n, n = 1, 2, \ldots\}$ be a sequence of random variables which is defined on $\{\Omega, \mathscr{F}, \mathbf{P}\}$. Let $\{\mathscr{F}_n, n = 1, 2, \ldots\}$ be an increasing sequence of σ-fields of $\mathscr{F}$ sets. Now let $\{M_n, \mathscr{F}_n, n = 1, 2, \ldots\}$ be a sequence of martingale, so that

$$D_0 = 0, \quad D_n = M_n - M_{n-1}, \quad n = 1, 2, \ldots. \tag{22}$$

is a martingale difference. $\mathscr{F}_0$ is a trivial σ field. For $n = 1, 2, \ldots$, denote

$$\begin{aligned} \sigma_n^2 &= E\left(D_n^2 \mid \mathscr{F}_{n-1}\right), \\ V_n^2 &= \sum_{j=1}^{n} \sigma_j^2, \\ v_n^2 &= E\left(V_n^2\right) = E\left(M_n^2\right). \end{aligned} \tag{23}$$

Lindeberg Condition. For $\forall \epsilon > 0$,

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} ED_j^2 I\left\{\left|D_j\right| \geq \epsilon v_n\right\}}{v_n^2} = 0, \tag{24}$$

where $I\{\cdot\}$ denotes the index function.

In our proof, we will use the central limit theorem of martingale sequences as the technical tool.

Lemma 4 (see [15]). *Suppose that the sequence of martingale* $\{M_n, \mathscr{F}_n, n = 1, 2, \ldots\}$ *satisfies the following condition:*

$$\frac{V_n^2}{v_n^2} \overset{p}{\Longrightarrow} 1. \tag{25}$$

Moreover, if the Lindeberg condition holds, then one has

$$\frac{M_n}{v_n} \overset{D}{\Longrightarrow} N(0, 1), \tag{26}$$

where $\overset{p}{\Rightarrow}$ *and* $\overset{D}{\Rightarrow}$ *denote convergence in probability and in distribution, respectively.*

Before we prove our main result Theorem 2, we at first come to prove Theorem 5.

Theorem 5. *Let* $X = \{X_n, n \geq 0\}$ *be an m-order nonhomogeneous Markov information source which is taking values in state space* $S = \{1, 2, \ldots, N\}$ *with initial distribution of* (3) *and mth-order transition probability matrices* (5). *Let* f *be any function defined on the state space* S^{m+1}. *Suppose that the function* f *satisfies condition* (21). *Let* $\{W_n, n \geq 0\}$ *be defined as* (16). *If* (14) *holds, then*

$$\frac{W_n}{\sqrt{n}\sigma} \overset{D}{\Rightarrow} N(0, 1), \tag{27}$$

where $\overset{D}{\Rightarrow}$ *denotes the convergence in distribution.*

Proof of Theorem 5. Noting that by using the property of the conditional expectation and Markov property, it follows from (17) that

$$\begin{aligned} \frac{V_n^2}{n} &= \frac{1}{n}\sum_{k=m}^{n} E\left[D_k^2 \mid \mathscr{F}_{k-1}\right] \\ &= \frac{1}{n}\sum_{k=m}^{n} \left\{E\left[f^2\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right] - \left(E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right]\right)^2\right\} \\ &:= I_1(n) - I_2(n), \end{aligned} \tag{28}$$

where

$$\begin{aligned} I_1(n) &= \frac{1}{n}\sum_{k=m}^{n} E\left[f^2\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right] \\ &= \frac{1}{n}\sum_{k=m}^{n}\sum_{j\in S}\sum_{i_1^m\in S^m} f^2(i_1^m, j)\, p_k(j \mid i^m)\, I\left\{X_{k-m}^{k-1} = i_1^m\right\} \\ &= \sum_{j\in S}\sum_{i_1^m\in S^m} f^2(i_1^m, j)\, \frac{1}{n}\sum_{k=m}^{n} p_k(j \mid i_1^m)\, I\left\{X_{k-m}^{k-1} = i_1^m\right\}, \end{aligned} \tag{29}$$

$$\begin{aligned} I_2(n) &= \frac{1}{n}\sum_{k=m}^{n}\left(E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right]\right)^2 \\ &= \frac{1}{n}\sum_{k=m}^{n}\sum_{i_1^m\in S^m}\left[\sum_{j\in S} f(i_1^m, j)\, p_k(j \mid i_1^m)\right]^2 \\ &\quad \times I\left\{X_{k-m}^{k-1} = i_1^m\right\}. \end{aligned} \tag{30}$$

Noting that, on the one hand,

$$\begin{aligned} &\left|\frac{1}{n}\sum_{k=m}^{n} I\left\{X_{k-m}^{k-1} = i_1^m\right\}\left[p_k(j \mid i_1^m) - p(j \mid i_1^m)\right]\right| \\ &\quad \leq \frac{1}{n}\sum_{k=m}^{n}\left|p_k(j \mid i_1^m) - p(j \mid i_1^m)\right| \end{aligned} \tag{31}$$

which tends to zero as n tends to infinity by using (14). Thus we have

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\sum_{k=m}^{n} I\left\{X_{k-m}^{k-1}=i_1^m\right\} p_k\left(j\mid i_1^m\right)\\
&\quad=\lim_{n\to\infty}\frac{1}{n}\sum_{k=m}^{n} I\left\{X_{k-m}^{k-1}=i_1^m\right\} p\left(j\mid i_1^m\right)\\
&\quad=\lim_{n\to\infty}\frac{1}{n}S_n\left(i_1^m\right) p\left(j\mid i_1^m\right)\\
&\quad=\pi\left(i_1^m\right) p\left(j\mid i_1^m\right)\quad \text{a.s.},
\end{aligned}\tag{32}$$

where the third equation holds because of (15). Combining (29) and (32), we get

$$\lim_{n\to\infty} I_1(n)=\sum_{j\in S}\sum_{i_1^m\in S^m}\pi\left(i_1^m\right) p\left(j\mid i_1^m\right) f^2\left(i_1^m,j\right)\ \text{a.s.}\tag{33}$$

On the other hand, let us come to compute the limit of $I_2(n)$ as n tends to infinity. By using (14) again, we have

$$\begin{aligned}
&\left|I_2(n)-\frac{1}{n}\sum_{k=m}^{n}\sum_{i_1^m\in S^m}\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2 I\left\{X_{k-m}^{k-1}=i_1^m\right\}\right|\\
&\le\frac{1}{n}\sum_{k=m}^{n}\sum_{i_1^m\in S^m}\left[\sum_{j\in S} f\left(i_1^m,j\right)\left|p_k\left(j\mid i_1^m\right)-p\left(j\mid i_1^m\right)\right|\right]\\
&\qquad\times\left[\sum_{j\in S} f\left(i_1^m,j\right)\left(p_k\left(j\mid i_1^m\right)+p\left(j\mid i_1^m\right)\right)\right]\\
&\le 2\left(\max_{i_1^m\in S^m, j\in S} f\left(i_1^m,j\right)\right)^2\\
&\quad\times\sum_{i_1^m\in S^m}\sum_{j\in S}\frac{\sum_{k=m}^{n}\left|p_k\left(j\mid i_1^m\right)-p\left(j\mid i_1^m\right)\right|}{n}\longrightarrow 0.\\
&\qquad\qquad\text{as } n\longrightarrow\infty.
\end{aligned}\tag{34}$$

Thus by Lemma 1, we easily arrive at

$$\begin{aligned}
&\lim_{n\to\infty} I_2(n)\\
&\quad=\sum_{i_1^m\in S^m}\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\frac{1}{n}\sum_{k=m}^{n} I\left\{X_{k-m}^{k-1}=i_1^m\right\}\\
&\quad=\sum_{i_1^m\in S^m}\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\frac{S_n\left(i_1^m\right)}{n}\\
&\quad=\sum_{i_1^m\in S^m}\pi\left(i_1^m\right)\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\quad\text{a.s.}
\end{aligned}\tag{35}$$

Combining (28), (33), and (35), we arrive at

$$\begin{aligned}
&\lim_{n\to\infty}\frac{V_n^2}{n}\\
&\quad=\sum_{i_1^m\in S^m}\pi\left(i_1^m\right)\left\{\sum_{j\in S} f^2\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right.\\
&\qquad\left.-\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\right\}\quad\text{a.s.},
\end{aligned}\tag{36}$$

which implies that

$$\begin{aligned}
&\lim_{n\to\infty}\frac{V_n^2}{n}\\
&\quad=\sum_{i_1^m\in S^m}\pi\left(i_1^m\right)\left\{\sum_{j\in S} f^2\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right.\\
&\qquad\left.-\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\right\}\\
&\qquad\qquad\text{in probability.}
\end{aligned}\tag{37}$$

Note that

$$\begin{aligned}
\frac{V_n^2}{n}&\le\max_{m\le k\le n} E\left[D_k^2\mid X_{k-m}^{k-1}\right]\\
&=\max_{m\le k\le n}\left\{E\left[f^2\left(X_{k-m}^{k}\right)\mid X_{k-m}^{k-1}\right]\right.\\
&\qquad\left.-\left(E\left[f\left(X_{k-m}^{k}\right)\mid X_{k-m}^{k-1}\right]\right)^2\right\}\\
&\le\max_{i_1^m\in S^m, j\in S} f^2\left(i_1^m,j\right).
\end{aligned}\tag{38}$$

Since S is a finite set, then the random sequence $\{V_n^2/n, n\ge 1\}$ is uniformly integrable. Combining above two facts, we arrive at

$$\begin{aligned}
&\lim_{n\to\infty}\frac{E\left[V_n^2\right]}{n}\\
&\quad=\sum_{i_1^m\in S^m}\pi\left(i_1^m\right)\left\{\sum_{j\in S} f^2\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right.\\
&\qquad\left.-\left[\sum_{j\in S} f\left(i_1^m,j\right) p\left(j\mid i_1^m\right)\right]^2\right\}\\
&\quad=\sigma^2>0.
\end{aligned}\tag{39}$$

It follows that

$$\frac{V_n^2}{v_n^2} \xrightarrow{p} 1, \tag{40}$$

where $v_n^2 = E[V_n^2] = E[W_n^2]$. On the other hand, similar to the analysis of inequality (38), we also have that $\{D_n^2 = [f(X_{k-m}^k) - E[f(X_{k-m}^k) \mid X_{k-m}^{k-1}]]^2\}$ is uniformly integrable, so that

$$\lim_{n\to\infty} \frac{\sum_{j=m}^n \mathbf{E} D_j^2 I\left(\left|D_j\right| \geq \epsilon\sqrt{n}\right)}{n} = 0, \tag{41}$$

which implies that the Lindeberg condition holds, and then we can easily get our conclusion by using Lemma 4. □

Now let us come to prove our main result of Theorem 2.

Proof of Theorem 2. Note that

$$\begin{aligned} S_n - E\left[S_n\right] &= W_n + \sum_{k=m}^n \left[E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right] - E\left[f\left(X_{k-m}^k\right)\right]\right]. \end{aligned} \tag{42}$$

Denote

$$P\left(X_{k-m}^{k-1} = s_1^m, X_k = j\right) = P_k\left(s_1^m, j\right). \tag{43}$$

and $M = \sup_{s_1^m \in S^m, j\in S} f(s_1^m, j)$. Let us come to estimate the upper bound of $|E[f(X_{k-m}^k) \mid X_{k-1}] - E[f(X_{k-m}^k)]|$. In fact, it follows from the C-K formula

$$\begin{aligned}
&\left|E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right] - E\left[f\left(X_{k-m}^k\right)\right]\right| \\
&= \left|\sum_{j\in S} f\left(X_{k-m}^{k-1}, j\right) P_k\left(j \mid X_{k-m}^{k-1}\right) - \sum_{s_1^m \in S^m, j\in S} f\left(s_1^m, j\right) P_k\left(s_1^m, j\right)\right| \\
&\leq \sup_{i_1^m} \left|\sum_{j\in S} f\left(i_1^m, j\right)\left[p_k\left(j \mid i_1^m\right) - \sum_{s_1^m} P\left(X_{k-m}^{k-1} = s_1^m\right) p_k\left(j \mid s_1^m\right)\right]\right| \\
&\leq M \sup_{i_1^m} \sum_j \left|p_k\left(j \mid i_1^m\right) - \sum_{s_1^m} P\left(X_{k-m}^{k-1} = s_1^m\right) p_k\left(j \mid s_1^m\right)\right| \\
&= M \sup_{i_1^m} \sum_j \left|\sum_{s_1^m} P\left(X_{k-m}^{k-1} = s_1^m\right) p_k\left(j \mid i_1^m\right) - \sum_{s_1^m} P\left(X_{k-m}^{k-1} = s_1^m\right) p_k\left(j \mid s_1^m\right)\right| \\
&\leq M \sum_{s_1^m} P\left(X_{k-m}^{k-1} = s_1^m\right) \times \sup_{i_1^m \in S^m} \sum_{j\in S} \left|p_k\left(j \mid i_1^m\right) - p_k\left(j \mid s_1^m\right)\right| \\
&\leq M \sup_{l_1^m, k_1^m} \sum_{j\in S} \left|p_k\left(j \mid l_1^m\right) - p_k\left(j \mid k_1^m\right)\right| \\
&= 2M\delta\left(P_k\right),
\end{aligned} \tag{44}$$

where $\delta(P_k) = \sup_{l_1^m, k_1^m} \sum_{j\in S} [p_k(j \mid l_1^m) - p_k(j \mid k_1^m)]^+ = (1/2)\sup_{l_1^m, k_1^m} \sum_{j\in S} |p_k(j \mid l_1^m) - p_k(j \mid k_1^m)|$. By using condition (19), we get

$$\lim_{n\to\infty} \frac{\sum_{k=m}^n \left[E\left[f\left(X_{k-m}^k\right) \mid X_{k-m}^{k-1}\right] - E\left[f\left(X_{k-m}^k\right)\right]\right]}{\sqrt{n}} = 0. \tag{45}$$

Then, by using (27), (42), and (45), we can arrive at our conclusion (20). Thus the proof of Theorem 2 is completed. □

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China no. 11201344.

References

[1] W. G. Yang and W. Liu, "The asymptotic equipartition property for mth-order nonhomogeneous Markov information sources," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3326–3330, 2004.

[2] M. I. Gordin and B. A. Lifšic, "Central limit theorem for stationary Markov processes," *Doklady Akademii Nauk SSSR*, vol. 239, no. 4, pp. 766–767, 1978.

[3] Y. Derriennic and M. Lin, "The central limit theorem for Markov chains with normal transition operators, started at a point," *Probability Theory and Related Fields*, vol. 119, no. 4, pp. 508–528, 2001.

[4] Y. Derriennic and M. Lin, "The central limit theorem for Markov chains started at a point," *Probability Theory and Related Fields*, vol. 125, no. 1, pp. 73–76, 2003.

[5] M. I. Gordin and H. Holzmann, "The central limit theorem for stationary Markov chains under invariant splittings," *Stochastics and Dynamics*, vol. 4, no. 1, pp. 15–30, 2004.

[6] C. Kipnis and S. R. S. Varadhan, "Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions," *Communications in Mathematical Physics*, vol. 104, no. 1, pp. 1–19, 1986.

[7] M. Maxwell and M. Woodroofe, "Central limit theorems for additive functionals of Markov chains," *The Annals of Probability*, vol. 28, no. 2, pp. 713–724, 2000.

[8] M. Woodroofe, "A central limit theorem for functions of a Markov chain with applications to shifts," *Stochastic Processes and Their Applications*, vol. 41, no. 1, pp. 33–44, 1992.

[9] H. Holzmann, "Martingale approximations for continuous-time and discrete-time stationary Markov processes," *Stochastic Processes and Their Applications*, vol. 115, no. 9, pp. 1518–1529, 2005.

[10] R. Dobrushin, "Central limit theoremsfor non-stationary Markov chains. I," *Theory of Probability & Its Applications*, vol. 1, pp. 65–80, 1956.

[11] R. Dobrushin, "Central limit theorem for nonstationary Markov chains. II," *Theory of Probability & Its Applications*, vol. 1, no. 4, pp. 329–383, 1956.

[12] V. Statuljavicius, "Limit theorems for sums of random variables connected in Markov chains," *Lietuvos Matematikos Rinkinys*, vol. 9, pp. 346–362, 635–672, *Lietuvos Matematikos Rinkinys*, vol. 10, pp. 583–592, 1969.

[13] T. A. Sarymsakov, "Inhomogeneous Markov chains," *Theory of Probability and Its Applications*, vol. 6, no. 2, pp. 178–185, 1961.

[14] S. Sethuraman and S. R. S. Varadhan, "A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains," *Electronic Journal of Probability*, vol. 10, article 36, pp. 1221–1235, 2005.

[15] B. M. Brown, "Martingale central limit theorems," *Annals of Mathematical Statistics*, vol. 42, pp. 59–66, 1971.