On the strong law of large numbers for functionals of countable nonhomogeneous Markov chains

Guoxin Liu and Wen Liu

Department of Mathematics, Hebei Institute of Technology, Tianjin, People's Republic of China

Received 31 March 1993
Revised 8 June 1993

We prove a strong law of large numbers for functionals of nonhomogeneous Markov chains. The approach is analytic and different from the usual one.

AMS 1980 Subject Classifications: 60F15, 60F10.

functional of a nonhomogeneous Markov chain * strong law of large numbers * a.s. convergence * conditional expectation

Let \( \{X_n, n \geq 0\} \) be a nonhomogeneous Markov chain with the state space \( S = \{1, 2, \ldots\} \) and transition matrices \( P_n, n \geq 0 \). If \( f_n (n \geq 0) \) is a real function defined on \( S \), the process \( \{f_n(X_n), n \geq 0\} \) is called a functional of the Markov chain \( \{X_n, n \geq 0\} \).

There have been some researches on the strong law of large numbers for Markov chains. Chung (1967) presents some remarkable results in the case of a homogeneous Markov chain with \( f_n = f (n \geq 0) \). Rosenblatt-Roth (1964), by means of the ergodic coefficient of a stochastic transition function, tries to give us some necessary and sufficient conditions for the strong law of large numbers in the case of a nonhomogeneous Markov chain. It should be mentioned that some of the results by Rosenblatt-Roth [7] are based on an incorrect proof (cf. M. Iosifescu's review in Zentralblatt für Mathematik 127 (1967) 354–355).

Here we prove that, under some conditions of classical form, the stochastic processes \( \{E(f_{m+k}(X_{m+k}) | X_m), m \geq 0\}, k \geq 0 \), simultaneously obey or do not obey the strong law of large numbers, and give some sufficient conditions for the strong law of large numbers to hold for a functional of a nonhomogeneous Markov chain. The approach used in this paper is rather different from the usual one. The main idea is to construct a suitable monotone function and to use Lebesgue's theorem on a.s. differentiability of monotone functions.

Correspondence to: Dr. Guoxin Liu, Department of Mathematics, Hebei Institute of Technology, Tianjin 300130, People's Republic of China.
Let \( \{g_n, n \geq 0\} \) be a sequence of positive, even and continuous functions on \( \mathbb{R} \) such that, as \( |x| \) increases,
\[
g_n(x) / |x| \uparrow, \quad g_n(x) / x^2 \downarrow, \quad n \geq 0.
\]

**Theorem 1.** Let \( \{f_n(X_n), n \geq 0\} \) be a functional of a nonhomogeneous countable Markov chain \( \{X_n, n \geq 0\} \) and \( 0 < a_n \rightarrow \infty \). If \( \{g_n, n \geq 0\} \) satisfies condition (1) and
\[
\sum_n \frac{E(g_n(f_n(X_n)))}{g_n(a_n)} < \infty,
\]
then
\[
\sum_n \frac{f_n(X_n) - E(f_n(X_n) \mid X_{n-k})}{a_n} \text{ converges a.s. ,}
\]
and
\[
\frac{1}{a_n} \sum_{n=1}^m (f_n(X_n) - E(f_n(X_n) \mid X_{n-k})) \rightarrow 0 \quad \text{a.s.}
\]
for all \( k \geq 1 \) (where \( X_{-n} = \text{const.}, n \geq 1 \)).

**Proof.** Let the initial distribution and transition matrices of \( \{X_n, n \geq 0\} \) be, respectively,
\[
q(1), q(2), q(3), \ldots,
\]
\[
P_n = (p_n(i,j)), \quad i, j \in S, \quad n \geq 0,
\]
where \( p_n(i,j) = P(X_{n+1} = j \mid X_n = i) \). Throughout the proof we shall deal with the underlying probability space \( (\mathcal{F}, \mathcal{B}, \mathbb{P}) \), where \( \mathcal{B} \) is the class of Borel sets in the interval \([0, 1)\), and \( \mathbb{P} \) is the Lebesgue measure. We first give, in the above probability space, a realization of the Markov chain with initial distribution (5) and transition matrices (6).

Let
\[
q(n_i), \quad i = 1, 2, \ldots,
\]
be the positive terms of (5), where \( n_1 < n_2 < n_3 < \cdots \). Divide the interval \([0, 1)\) into countably (possibly finitely) many right-semiopen intervals: \( d_{n_1}, x_0 = n_1, n_2, \ldots \), i.e.
\[
d_{n_1} = [0, q(n_1)) \quad \text{and} \quad d_{n_2} = [q(n_1), q(n_1) + q(n_2)), \ldots,
\]
which will be called intervals of order 0. In general, suppose \( d_{n_0} \cdots d_{n_k} \) is an interval of order \( n \), and the positive elements of the \( x_{th} \) row of the matrice \( P_n \), in their natural order, are
\[
p_n(x_n, m_i), \quad i = 1, 2, \ldots.
\]
By dividing \( d_{n_0} \cdots d_{n_k} \) into countably many right-semiopen intervals \( d_{n_0} \cdots d_{n_{k+1}} \) (where \( x_{n+1} = m_i, \)
\[
i = 1, 2, \ldots \)) at the ratio
\[
p_n(x_n, m_1) : p_n(x_n, m_2) : p_n(x_n, m_3) : \cdots.
\]
the intervals of order \( n + 1 \) are constructed. It is easy to see that
\[
P(d_{x_0} \cdots x_n) = q(x_0) \prod_{m=0}^{n-1} p_m(x_m, x_{m+1}).
\]  
(8)

For \( n \geq 0 \), define a random variable \( X_n : [0, 1) \to S \) by
\[
X_n(\omega) = x_n, \quad \text{if} \ \omega \in d_{x_0} \cdots x_n.
\]  
(9)

By (8) and (9) we have
\[
P(X_0 = x_0, \ldots, X_n = x_n) = P(d_{x_0} \cdots x_n) = q(x_0) \prod_{m=0}^{n-1} p_m(x_m, x_{m+1}).
\]  
(10)

Therefore \( \{X_n, n \geq 0\} \) be a Markov chain with initial distribution (5) and transition matrices (6).

Since hypothesis (2) can be rewritten as
\[
\sum_n E\left( E(g_n(f_n(X_n)) | X_{n-1}) / g_n(a_n) \right) < \infty,
\]  
(11)

and \( E(g_n(f_n(X_n)) | X_{n-1}) \) is a nonnegative random variable \( n \geq 1 \), by (11), we have
\[
\sum_n E\left( g_n(f_n(X_n)) | X_{n-1} \right) / g_n(a_n) \text{ converges a.s.}
\]  
(12)

For \( n \geq 0 \), define the function \( f_n^* \) by
\[
f_n^*(x) = \begin{cases} f_n(x), & \text{if} \ |f_n(x)| \leq a_n, \\ 0, & \text{if} \ |f_n(x)| > a_n. \end{cases}
\]  
(13)

By (1), \( g_n(x) \uparrow \) as \( |x| \) increases, and by (2), we have
\[
\sum_n P(f_n^*(X_n) \neq f_m(X_m)) = \sum_n \int_{|f_m(X_m)| > a_m} P(d\omega) 
\leq \sum_n \int_{|f_m(X_m)| > a_m} \frac{g_m(f_m(X_m))}{g_m(a_m)} P(d\omega)
\leq \sum_n E\left( g_m(f_m(X_m)) \right) / g_m(a_m) P(d\omega) < \infty.
\]

Hence \( \{f_n(X_n)\} \) and \( \{f_n^*(X_n)\} \) are equivalent sequences and, in particular,
\[
\sum_n (f_n^*(X_n) - f_m(X_m)) / a_m \text{ converges a.s.}
\]  
(14)

Let \( i \in S \), and let \( \lambda \) be a nonzero constant. If \( P(X_{n-1} = i) > 0 \), then let
\[
Q_n(\lambda, i) = E(\exp(\lambda(f_n(X_n) - f_n(X_n) | X_{n-1} = i)) / a_n | X_{n-1} = i))
\]
\[
(n \geq 1).
\]  
(15)

It is easy to see that \( 0 < Q_n(\lambda, i) < \infty \) \( n \geq 1 \).
Let \( \mathcal{A} \) denote the collection consisting of the interval \([0, 1]\) and the intervals of all orders. Define a set function \( \mu \) on \( \mathcal{A} \) by

\[
\mu(d_{x_0 \cdots x_n}) = \sum_{\lambda_1} \mu(d_{x_0x_1})
\]

where \( \lambda_1 \) denotes the sum taken over all possible values of \( x_i \). By (10), (15) and (16), it is easy to see that

\[
\mu(d_{x_0 \cdots x_n}) = \mu(d_{x_0 \cdots x_{n-1}}) .
\]

By (16), (17) and (19), \( \mu \) is additive. Hence there exists an increasing function \( f_\lambda \) defined on \([0, 1]\) such that

\[
\mu(d_{x_0 \cdots x_n}) = f_\lambda(d_{x_0 \cdots x_n}) - f_\lambda(d_{x_0 \cdots x_n}) ,
\]

for any \( d_{x_0 \cdots x_n} \), where \( d_{x_0 \cdots x_n}^+ \) and \( d_{x_0 \cdots x_n}^- \) denote respectively the right and left endpoints of \( d_{x_0 \cdots x_n} \). Let

\[
t_n(\lambda, \omega) = \frac{f_\lambda(d_{x_0 \cdots x_n}^+) - f_\lambda(d_{x_0 \cdots x_n}^-)}{d_{x_0 \cdots x_n}^+ - d_{x_0 \cdots x_n}^-} = \frac{\mu(d_{x_0 \cdots x_n})}{P(d_{x_0 \cdots x_n})} , \quad \omega \in d_{x_0 \cdots x_n} , n \geq 1 .
\]

If \( \lim_n P(d_{x_0 \cdots x_n}) = d > 0 \), then

\[
\lim_n t_n(\lambda, \omega) = (1/d) \lim_n \mu(d_{x_0 \cdots x_n}) \text{ exists and is finite} ;
\]

and if \( \lim_n P(d_{x_0 \cdots x_n}) = 0 \), according to a property of the derivative (see [1, p. 423]), by (20) we have

\[
\lim_n t_n(\lambda, \omega) = f'_\lambda(\omega) ,
\]

if \( f'_\lambda(\omega) \) exists. Since \( f_\lambda \) is a monotone function, it is a.s. differentiable. Therefore by (21) and (22) we have

\[
\lim_n t_n(\lambda, \omega) \text{ exists and is finite a.s.}
\]
By (9), (10), (16) and (20) we have
\[
\exp\left(\lambda \sum_{m=1}^{n} \frac{[f^*_m(X_m) - E(f^*_m(X_m) \mid X_{m-1})]}{a_m}\right)
\]
\[
t_n(\lambda, \omega) = \frac{1}{\prod_{m=1}^{n} Q_n(\lambda, X_{m-1})}
\]
for all \(\omega\). Hence, by (23),
\[
\exp\left(\lambda \sum_{m=1}^{n} \frac{[f^*_m(X_m) - E(f^*_m(X_m) \mid X_{m-1})]}{a_m}\right)
\]
\[
limit_{n} \frac{1}{\prod_{m=1}^{n} Q_n(\lambda, X_{m-1})}
\]
exists and is finite a.s. (24)

Noticing that
\[
Q_m(\lambda, X_{m-1})
\]
\[
- E(\exp(\lambda(f^*_m(X_m) - E(f^*_m(X_m) \mid X_{m-1}) \mid a_m \mid X_{m-1})) (m \geq 1)
\]
and
\[
0 \leq e^x - 1 - x \leq x^2 e^x \quad \text{for } -t < x < t,
\]
by (13) we have
\[
0 \leq Q_m(\lambda, X_{m-1}) - 1
\]
\[
= E(\exp(\lambda(f^*_m(X_m) - E(f^*_m(X_m) \mid X_{m-1}) \mid a_m \mid X_{m-1})) - 1
\]
\[
- \lambda(f^*_m(X_m) - E(f^*_m(X_m) \mid X_{m-1}) \mid a_m \mid X_{m-1}))
\]
\[
\leq \lambda^2 e^{2\lambda |x|} E(\frac{f^*_m(X_m)}{a_m})^2 |X_{m-1}) - (E(f^*_m(X_m) \mid X_{m-1}))^2 / a_m^2
\]
\[
\leq \lambda^2 e^{2\lambda |x|} E(\frac{f^*_m(X_m)}{a_m})^2 |X_{m-1}) / a_m^2.
\]
By (1), if \(|x| \leq a_m\), then
\[
x^2 / a_m^2 \leq g_m(x) / g_m(a_m) \quad \text{for } m \geq 1.
\]
Hence
\[
\frac{(f^*_m(X_m))^2}{a_m^2} \leq \frac{g_m(f^*_m(X_m))}{g_m(a_m)} \leq \frac{g_m(f_m(X_m))}{g_m(a_m)} \quad m \geq 1.
\]
Therefore
\[
0 \leq Q_m(\lambda, X_{m-1}) - 1 \leq \lambda^2 e^{2\lambda |x|} E(\frac{f_m(X_m)}{g_m(a_m)}) / g_m(a_m) \mid X_{m-1})
\]
By (12) we have
\[
\sum_{m} \left( Q_m(\lambda, X_{m-1}) - 1 \right) \text{ converges a.s.},
\]
or, equivalently,
\[
\prod_{m=1}^{\infty} Q_m(\lambda, X_{m-1}) \text{ converges a.s.} \quad (26)
\]

Using (24), by (26) we have
\[
\lim_{n} \exp \left( \lambda \sum_{m=1}^{n} \frac{(f_m^*(X_m) - E(f_m^*(X_m) \mid X_{m-1}))/a_m}{a_m} \right)
\]
exists and is finite a.s.  

Because (27) holds for both \( \lambda = 1 \) and \( \lambda = -1 \), we get
\[
\sum_{m} \left( f_m^*(X_m) - E(f_m^*(X_m) \mid X_{m-1}) \right)/a_m \text{ converges and is finite a.s.} \quad (28)
\]

By (1), we have \(|x|/a_n < g_n(x)/g_n(a_n)\) if \(|x| > a_n\). Hence
\[
\left| \left( E(f_n(X_n) \mid X_{n-1}) - E(f_n^*(X_n) \mid X_{n-1}) \right)/a_n \right|
\]
\[
= \left| E(f_n(X_n) - f_n^*(X_n) \mid X_{n-1})/a_n \right|
\]
\[
\leq E(|f_n(X_n) - f_n^*(X_n) \mid X_{n-1})/a_n \mid X_{n-1})
\]
\[
= E((|f_n(X_n)|/a_n)I(|f_n(X_n)| > a_n) \mid X_{n-1})
\]
\[
\leq E((g_n f_n(X_n))/g_n(a_n))I(|f_n(X_n)| > a_n) \mid X_{n-1})
\]
\[
\leq E(g_n f_n(X_n)) \mid X_{n-1})/g_n(a_n). \quad (29)
\]

Therefore, by (12) and (29) we have
\[
\sum_{n} \left( E(f_n(X_n) \mid X_{n-1}) - E(f_n^*(X_n) \mid X_{n-1}) \right)/a_n \text{ converges a.s.} \quad (30)
\]

By (14), (28) and (30) we conclude that
\[
\sum_{n} (f_n(X_n) - E(f_n(X_n) \mid X_{n-1}))/a_n \text{ converges a.s.} \quad (31)
\]

This completes the proof of (3) for \( k = 1 \).

For \( k > 1 \), \([X_{nk+m}, n \geq 0]\) is a nonhomogeneous Markov chain too, and by (2) we have
\[
\sum_{n} E(g_{nk+m}(f_{nk+m}(X_{nk+m}))) < \infty \quad (m = 0, 1, \ldots, k-1).
\]

Hence
\[
\sum_{n} \frac{f_{nk+m}(X_{nk+m}) - E(f_{nk+m}(X_{nk+m}) \mid X_{(n-1)k+m})}{a_{nk+m}}
\]
converges a.s. \((m=0, 1, \ldots, k-1)\).

Therefore,

\[
\sum_n \frac{f_n(X_n) - E(f_n(X_n) \mid X_{n-k})}{a_n} = \sum_{m=0}^{k-1} \sum_n \frac{f_{nk+m}(X_{nk+m}) - E(f_{nk+m}(X_{nk+m}) \mid X_{(n-1)k+m})}{a_{nk+m}}
\]

converges a.s.,

i.e. (3) holds for \(k>1\).

Applying Kronecker's lemma to (3) for each \(\omega\) in a set of probability one, we obtain (4) for all \(k \geq 1\). This completes the proof of the theorem.

It is obvious that, in the independent case, \(E(f_n(X_n) \mid X_{n-1}) = Ef_n(X_n)\) for all \(n \geq 0\).

Hence, Theorem 1 implies Loève's generalization of Kolmogorov's result for sequences of independent random variables.

For a stochastic matrix \(P = (p(i,j))\), denote

\[
\delta(P) = \frac{1}{2} \sup_{i,k} \sum_j |p(i,j) - p(k,j)|.
\]

The \(\delta(P)\) is called the \(\delta\)-coefficient of \(P\). It is well known that a nonhomogeneous Markov chain with transition matrices \(\{P_n\}\) is weakly ergodic if and only if

\[
\delta(P^{(m,m+k)}) \to 0 \quad (k \to \infty) \quad \text{for all } m \geq 0,
\]

where \(P^{(m,m+k)} = P_m \cdots P_{m+k-1}\) (see [4, p. 149]).

**Corollary 1.** If

\[
|f_n(X_n)| \leq M, \quad n \geq 0, \quad \text{(32)}
\]

and

\[
\sup_m \delta(P^{(m,m+k)}) \to 0 \quad (k \to \infty), \quad \text{(33)}
\]

then

\[
\lim_n \frac{1}{n} \sum_{m=0}^n (f_m(X_m) - Ef_m(X_m)) = 0 \quad \text{a.s.} \quad \text{(34)}
\]

(Condition (33) above can be thought as uniform weak ergodicity.)

**Proof.** By (32) we have
\[
\sum_n \frac{E(f_n(X_n))^2}{(n+1)^2} \leq \sum_n \frac{M^2}{(n+1)^2} < \infty,
\]
i.e., \([f_n(X_n), n \geq 0]\) satisfies the hypothesis of Theorem 1 for \(g_n(x) = x^2, \alpha_n = n + 1 (n \geq 0)\).

Hence
\[
\lim n \frac{1}{n+1} \sum_{m=1}^n (f_m(X_m) - E(f_m(X_m) \mid X_{m-k})) = 0 \text{ a.s.}
\]

For all \(k \geq 1 (X_{-m} = 0 \text{ for } m \geq 1)\). Since
\[
\left| \frac{1}{n+1} \sum_{m=0}^n (f_m(X_m) - E_f m(X_m)) \right| 
\leq \left| \frac{1}{n+1} \sum_{m=0}^n (f_m(X_m) - E(f_m(X_m) \mid X_{m-k})) \right|
\]
\[+ \left| \frac{1}{n+1} \sum_{m=0}^n (E(f_m(X_m) \mid X_{m-k}) - E_f m(X_m)) \right|,
\]
and the left side of the inequality above is independent of \(k\), we need only prove that
\[
\lim \sup n \frac{1}{n+1} \sum_{m=0}^{n-k} (E(f_m(X_m) \mid X_{m-k}) - E_f m(X_m)) = 0 \text{ a.s.}
\]
or, equivalently
\[
\lim \sup \frac{1}{n+1} \sum_{m=0}^{n-k} (E(f_{m+k}(X_{m+k}) \mid X_m) - E_{f_{m+k}}(X_{m+k})) = 0 \text{ a.s.}
\]

Denoting \(p^{(n)}(j) = P(X_n = j), p^{(m,m+k)}(i,j) = P(X_{m+k} = j \mid X_{m-i})\), we have
\[
|E(f_{m+k}(X_{m+k}) \mid X_m) - E_{f_{m+k}}(X_{m+k})| = \sum_j \sum_j p^{(m,m+k)}(X_m, j) f_{m+k}(j) - \sum_j p^{(m+k)}(j) f_{m+k}(j)
\]
\[
= \sum_j (p^{(m,m+k)}(X_m, j) - p^{(m+k)}(j)) f_{m+k}(j)
\]
\[
\leq \sup_i \sum_j (p^{(m,m+k)}(i, j) - p^{(m+k)}(j)) f_{m+k}(j),
\]
and, by (32),
\[ \sup_{i \leq j} \left| \sum_{m} (p^{(m,m+k)}(i,j) - p^{(m+k)}(j)) f_{m+k}(j) \right| \]

\[ \leq \sup_{i \leq j} \sum_{m} \left| (p^{(m,m+k)}(i,j) - \sum_{s} p^{(m,s)} p^{(m,m+k)}(s,j)) M \right| \]

\[ \leq \sup_{i \leq j} \sum_{m} p^{(m)}(s) \sup_{i \leq j} \left| p^{(m,m+k)}(i,j) - p^{(m,m+k)}(s,j) \right| M \]

\[ = \sup_{i \leq j} \sum_{m} \left| p^{(m,m+k)}(i,j) - p^{(m,m+k)}(l,j) \right| M \leq 2M \sup_{m} \delta(p^{(m,m+k)}) . \]

Hence

\[ \lim \sup_{k} \lim \sup_{n} \frac{1}{n+1} \sum_{m=0}^{n-k} (E(f_{m+k}(X_{m+k}) \mid X_{m}) - Ef_{m+k}(X_{m+k})) \]

\[ \leq \lim \sup_{k} \lim \sup_{n} \frac{1}{n+1} \sum_{m=0}^{n-k} 2M \sup_{m} \delta(p^{(m,m+k)}) \]

\[ = 2 \lim \sup_{k} M \sup_{m} (\delta(p^{(m,m+k)})) = 0 \quad \text{a.s.} \]

The last equality follows from (33). This completes the proof of (35) and the proof of the corollary too. \( \square \)

**Corollary 2.** Under the hypothesis of Theorem 1 for \( a_n = n + 1 \), if

\[ \sup_{m,i \leq j} \sum_{j} \left| p^{(m,m+k)}(i,j) - p^{(m+k)}(j) \right| \left| f_{m+k}(j) \right| \to 0 \quad (k \to \infty) . \]

(37)

where \( p^{(m,m+k)}(i,j) = P(X_{m+k} = j \mid X_{m} = i) \), \( p^{(m+k)}(j) = P(X_{m+k} = j) \), then

\[ \lim_{n} \frac{1}{n+1} \sum_{m=0}^{n} (f_{m}(X_{m}) - E(f_{m}(X_{m}))) = 0 \quad \text{a.s.} \]

(38)

(Condition (37) is something of a uniform weak weighted ergodicity of the chain.)

**Proof.** Similarly to the proof of Corollary 1, we need only prove (35). By (36) we have

\[ \left| \frac{1}{n+1} \sum_{m=0}^{n-k} (E(f_{m+k}(X_{m+k}) \mid X_{m}) - Ef_{m+k}(X_{m+k})) \right| \]

\[ \leq \frac{1}{n+1} \sum_{m=0}^{n-k} \sup_{i \leq j} \left| \sum_{j} (p^{(m,m+k)}(i,j) - p^{(m+k)}(j)) f_{m+k}(j) \right| \]

\[ \leq \frac{1}{n+1} (n-k+1) \sup_{m,i \leq j} \sum_{j} \left| p^{(m,m+k)}(i,j) - p^{(m+k)}(j) \right| \left| f_{m+k}(j) \right| . \]
Hence, by (37),
\[
\limsup_k \limsup_n \left| \frac{1}{n} \sum_{m=0}^{n-k} \left( E(f_{m+k}(X_{m+k}) \mid X_m) - E_{m+k}(X_{m+k}) \right) \right|
\leq \limsup_k \sup_{m,j} \sum_j |p^{(m,m+k)}(i,j) - p^{(m+k)}(j)| \mid f_{m+k}(j) | = 0 \quad \text{a.s.}
\]

This completes the proof of the corollary. \qed

**Theorem 2.** Let \( \{f_n(X_n), n \geq 0\} \) be a functional of a nonhomogeneous Markov chain \( \{X_n, n \geq 0\} \) and \( 0 < a_n \uparrow \infty \) such that

\[
\limsup \mathbb{E}\left[ \left| f_n(X_n) \right|^r \right] < \infty \quad (39)
\]

If

\[
\sum_n \frac{\mathbb{E}[f_n(X_n)]}{a_n^r} < \infty \quad (40)
\]

for some \( r \ (1 \leq r \leq 2) \), then

\[
\lim_n \left( \frac{1}{a_n} \sum_{m=0}^{n} \left( f_m(X_m) - Ef_m(X_m) \right) \right)
- \frac{1}{a_n} \sum_{m=0}^{n} \left( \frac{1}{N} \sum_{k=1}^{N} \left( E(f_{m+k}(X_{m+k}) \mid X_m) - E_{m+k}(X_{m+k}) \right) \right) = 0 \quad \text{a.s.}
\]

for all \( N \ (N \geq 1) \).

**Proof.** Letting \( g_n(x) = |x|^r \ (1 \leq r \leq 2) \), by (37) we have that \( \{f_n(x_n), n \geq 0\} \) satisfies the condition of Theorem 1. Hence, for \( N \geq 1 \),

\[
\lim_n \left( \frac{1}{a_n} \sum_{m=0}^{n} \left( f_m(X_m) - Ef_m(X_m) \right) \right)
- \frac{1}{a_n} \sum_{m=0}^{n} \left( E(f_m(X_m) \mid X_{m-k}) - Ef_m(X_m) \right) = 0 \quad \text{a.s.}
\]

for \( k = 1, 2, \ldots, N \). Therefore,

\[
\lim_n \left( \frac{1}{a_n} \sum_{m=0}^{n} \left( f_m(X_m) - Ef_m(X_m) \right) \right)
- \frac{1}{a_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{m=0}^{n} \left( E(f_m(X_m) \mid X_{m-k}) - Ef_m(X_m) \right) = 0 \quad \text{a.s.}
\]
Noticing that \( E(f_m(X_m) \mid X_{m-k}) = Ef_m(X_m) \) as \( m < k \), we have

\[
\frac{1}{\alpha_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{m=0}^{n} \left( E(f_m(X_m) \mid X_{m-k}) - Ef_m(X_m) \right)
\]

\[
= \frac{1}{\alpha_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{m=0}^{n} \left( E(f_m(X_m) \mid X_{m-k}) - Ef_m(X_m) \right)
\]

\[
= \frac{1}{\alpha_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{m=0}^{n-k} \left( E(f_{m+k}(X_{m+k}) \mid X_m) - Ef_{m+k}(X_{m+k}) \right)
\]

\[
= \frac{1}{\alpha_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{m=0}^{n} \left( E(f_{m+k}(X_{m+k}) \mid X_m) - Ef_{m+k}(X_{m+k}) \right) - I
\]

where

\[
I = \frac{1}{\alpha_n} \frac{1}{N} \sum_{k=1}^{N} \sum_{s=1}^{k} \left( E(f_{n+s}(X_{n+s}) \mid X_{n+s-k}) - Ef_{n+s}(X_{n+s}) \right).
\]

Now, we shall prove that for all \( N \geq 1 \) we have

\[
\lim_n I = 0 \quad \text{a.s.} \tag{43}
\]

Let \( \varepsilon > 0, s \geq 1 \). By Chebyshev’s inequality and Jensen’s inequality of conditional expectation, we have

\[
P( |E(f_{n+s}(X_{n+s}) \mid X_{n+s-k}) | \geq \varepsilon a_{n+s} )
\]

\[
\leq \frac{1}{\varepsilon^2 a_{n+s}} E[|E(f_{n+s}(X_{n+s}) \mid X_{n+s-k})|^r]
\]

\[
\leq \frac{1}{\varepsilon^2 a_{n+s}^r} E[f_{n+s}(X_{n+s})]^r \quad (k \geq 1).
\]

Hence, by (40) we get

\[
\sum_n P( |E(f_{n+s}(X_{n+s}) \mid X_{n+s-k}) | \geq \varepsilon a_{n+s} ) \leq \frac{1}{\varepsilon^r} \sum_n \frac{E[f_{n+s}(X_{n+s})]^r}{a_{n+s}^r} < \infty.
\]

By the formula above and the arbitrariness of \( \varepsilon > 0 \), we have

\[
\lim_n \frac{E(f_{n+s}(X_{n+s}) \mid X_{n+s-k})}{a_{n+s}} = 0 \quad \text{a.s.},
\]
and by (39), we get

$$\lim_{n} \frac{E(f_{n+s}(X_{n+s}) \mid X_{n+s-k})}{a_n} = 0 \quad \text{a.s.}$$

(44)

for all $k \geq 1$, $s \geq 1$. For $\varepsilon > 0$, $s \geq 1$, we have

$$\left| \frac{E_{f_{n+s}(X_{n+s})}}{a_{n+s}} \right| \leq \frac{E|f_{n+s}(X_{n+s})|}{a_{n+s}} = \int \frac{|f_{n+s}(X_{n+s})|}{a_{n+s}} p(d\omega)$$

$$+ \int \frac{|f_{n+s}(X_{n+s})|}{\varepsilon a_{n+s}} \varepsilon p(d\omega)$$

$$\leq \int \frac{|f_{n+s}(X_{n+s})|}{\varepsilon a_{n+s}} \varepsilon p(d\omega)$$

$$+ \int \frac{|f_{n+s}(X_{n+s})|}{\varepsilon^2 a_{n+s}} \varepsilon^2 p(d\omega)$$

$$\leq \varepsilon + \varepsilon e^{1-s} E|f_{n+s}(X_{n+s})|'/a_{n+s}$$

(45)

so, by (40), we get

$$\lim_{n} E|f_{n+s}(X_{n+s})|'/a_{n+s} = 0 .$$

(46)

By (45) and (46), we have

$$\lim_{n} \sup \frac{|E_{f_{n+s}(X_{n+s})}|}{a_{n+s}} \leq \varepsilon .$$

Therefore, by the arbitrariness of $\varepsilon > 0$, we have

$$\lim_{n} \frac{|E_{f_{n+s}(X_{n+s})}|}{a_{n+s}} = 0$$

for all $s \geq 1$. By (39) we get

$$\lim_{n} \frac{|E_{f_{n+s}(X_{n+s})}|}{a_n} = 0$$

(47)

for $s \geq 1$. Now, from (44) and (47) we conclude that (43) is valid for all $N \geq 1$.

Going back to (42), by (43) we have
This completes the proof of Theorem 2. \(\square\)

**Corollary 3.** Under the conditions of Theorem 2, if

\[
\lim \sup_n \lim \sup_{m=0}^{N} \frac{1}{a_n} \sum_{m=0}^{n} \sup_{i,j} \left| \sum_{k=1}^{N} \left( \frac{1}{N} \sum_{k=1}^{N} \left( p^{(m,m+k)}(i,j) \right) - p^{(m+k)}(j) f_{m+k}(j) \right) \right| = 0 ,
\]

then

\[
\lim_{n} \frac{1}{a_n} \sum_{m=0}^{n} (f_{m}(X_m) - Ef_{m}(X_m)) = 0 \quad a.s.
\]

**Proof.** Since

\[
\left| \frac{1}{N} \sum_{k=1}^{N} \left( E(f_{m+k}(X_{m+k}) \mid X_m) - Ef_{m+k}(X_{m+k}) \right) \right|
\]

\[
- \left| \frac{1}{N} \sum_{j} \sum_{k=1}^{N} \left( p^{(m,m+k)}(X_m,j) - p^{(m+k)}(j) f_{m+k}(j) \right) \right|
\]

\[
\leq \sup_i \left| \sum_{j} \frac{1}{N} \sum_{k=1}^{N} \left( p^{(m,m+k)}(i,j) - p^{(m+k)}(j) f_{m+k}(j) \right) \right|
\]

we have
By (48) and the inequality above we have

$$\lim_{n \to \infty} \sup_{N} \left( \sum_{m=0}^{n} \sum_{k=1}^{N} \left( \frac{1}{N} \sum_{k=1}^{N} \left( E\left(f_{m+k}(X_{m+k}) \mid X_{m}\right) - E_{f_{m+k}}(X_{m+k})\right) \right) \right) = 0 \quad \text{a.s.}$$

It is obvious that

$$\lim_{n \to \infty} \sup_{N} \left( \sum_{m=0}^{n} \sum_{k=1}^{N} \left( \frac{1}{N} \sum_{k=1}^{N} \left( E\left(f_{m+k}(X_{m+k}) \mid X_{m}\right) - E_{f_{m+k}}(X_{m+k})\right) \right) \right) = 0 \quad \text{a.s.}$$

This completes the proof of the corollary. \(\square\)

**Theorem 3.** Let \(\{X_n, n \geq 0\}\) be a countable nonhomogeneous Markov chain and let \(f\) be a real function such that \(|f(x)| \leq M\) for all \(x \in \mathbb{R}\). If

$$\sup_{m, i} \sum_{j} \left| \frac{1}{N} \sum_{k=1}^{N} p^{(m, m+k)}(i, j) - p(j) \right| \to 0 \quad (N \to \infty),$$

then

$$\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) \to \sum_{j} p(j) f(j) \quad \text{a.s.}$$
Proof. Since $|f(X_n)| \leq M (n \geq 0)$, we have

$$\sum_n \frac{E|f(X_n)|^2}{(n+1)^2} \leq \sum_n \frac{M^2}{(n+1)^2} < \infty.$$ 

We conclude, by Theorem 1, that

$$\frac{1}{n} \sum_{m=0}^{n-1} (f(X_m) - E(f(X_m) \mid X_{m-k})) \to 0 \quad \text{a.s.},$$

or, equivalently,

$$\left(\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) - \sum_j f(j)p(j)\right)$$

$$- \left(\frac{1}{n} \sum_{m=0}^{n-1} E(f(X_m) \mid X_{m-k}) - \sum_j f(j)p(j)\right) \to 0 \quad \text{a.s.} \quad (51)$$

Imitating the proof of Theorem 2, we can get

$$\left(\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) - \sum_j f(j)p(j)\right)$$

$$- \left(\frac{1}{n} \sum_{m=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k}) \mid X_m) - \sum_j f(j)p(j)\right) \to 0 \quad \text{a.s.} \quad (52)$$

Since

$$\left|\frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k}) \mid X_m) - \sum_j f(j)p(j)\right|$$

$$= \left|\frac{1}{N} \sum_{k=1}^{N} \sum_j (p^{(m,m+k)}(X_m, j) - p(j))f(j)\right|$$

$$\leq \left|\sum_j \left(\frac{1}{N} \sum_{k=1}^{N} p^{(m,m+k)}(X_m, j) - p(j)\right)f(j)\right|$$

$$\leq \sum_j \left|\frac{1}{N} \sum_{k=1}^{N} p^{(m,m+k)}(X_m, j) - p(j)\right| M$$

$$\leq M \sup_{m,d} \sum_j \left|\frac{1}{N} \sum_{k=1}^{N} p^{(m,m+k)}(i, j) - p(j)\right|,$$

we have

$$\left|\frac{1}{n} \sum_{m=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k}) \mid X_m) - \sum_j f(j)p(j)\right|$$
\[
M \sup_{m,i} \sum_j \left| \frac{1}{N} \sum_{k=1}^{N} p^{(m,m+k)}(i,j) - p(j) \right|
\]

Hence, by (50) and (53), as \( N \to \infty \),
\[
\frac{1}{n} \sum_{m=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k} \mid X_m) - \sum_j f(j)p(j) \to 0 \quad \text{a.s.}
\]
uniformly for all \( n \geq 1 \). Noticing that
\[
\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) - \sum_j f(j)p(j)
\]
\[
= \left( \frac{1}{n} \sum_{m=0}^{n-1} f(X_m) - \sum_j f(j)p(j) \right)\]
\[
- \left( \frac{1}{n} \sum_{m=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k}) \mid X_m) - \sum_j f(j)p(j) \right)
\]
\[
+ \left( \frac{1}{n} \sum_{m=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} E(f(X_{m+k}) \mid X_m) - \sum_j f(j)p(j) \right),
\]
by (52) and (54) we have
\[
\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) \to \sum_j f(j)p(j) \quad \text{a.s.}
\]
This completes the proof of Theorem 3. \( \square \)

Acknowledgment

The authors would like to thank Professor P. Jagers and the referee for their valuable remarks which improved the presentation of the paper.

References