Math-Net.Ru

Общероссийский математический портал

Z. Szewczak, A remark on a large deviation theorem for Markov chain with a finite number of states, *Теория вероятн. и ее примен.*, 2005, том 50, выпуск 3, 612–622

DOI: <https://doi.org/10.4213/tvp103>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.81.55.215

26 апреля 2023 г., 15:05:34

REFERENCES

1. Chover J. A law of the iterated logarithm for stable summands. — Proc. Amer. Math. Soc., 1966, v. 17, p. 441–443.
2. Pakshirajan R. P., Vasudeva R. A law of the iterated logarithm for stable summands. — Trans. Amer. Math. Soc., 1977, v. 232, p. 33–42.
3. Vasudeva R., Divanji G. LIL for delayed sums under a non-identically distributed setup. — Теория вероятн. и ее примен., 1992, т. 37, в. 3, с. 534–542.
4. Zinchenko N. M. A modified law of iterated logarithm for stable random variables. — Theory Probab. Math. Statist., 1994, v. 49, p. 69–76.
5. Kiesel R. The law of the iterated logarithm for certain power series and generalized Nörlund methods. — Math. Proc. Cambridge Philos. Soc., 1996, v. 120, № 4, p. 735–753.
6. Jakimovski A., Tietz H. Regularly varying functions and power series methods. — J. Math. Anal. Appl., 1980, v. 73, № 1, p. 65–84.
7. Kiesel R. General Nörlund transforms and power series methods. — Math. Z., 1993, v. 244, № 2, p. 273–286.
8. Bingham N. H., Goldie C. M., Teugels J. L. Regular Variation. Cambridge: Cambridge Univ. Press, 1987, 491 p.
9. Сенета Е. Правильно меняющиеся функции. М.: Наука, 1985, 142 с.
10. Pruitt W. E., Taylor S. J. Sample path properties of processes with stable components. — Z. Wahrscheinlichkeitstheor. Verw. Geb., 1969, v. 12, p. 267–289.

Поступила в редакцию
20.VIII.2002

© 2005 г.

SZEWCZAK Z. S.*

A REMARK ON A LARGE DEVIATION THEOREM FOR MARKOV CHAIN WITH A FINITE NUMBER OF STATES

Рассматривается классическая проблема больших отклонений для сумм случайных величин, заданных на состояниях однородной цепи Маркова с конечным фазовым пространством. Устанавливается точная асимптотика для вероятностей больших отклонений порядка $O(\sqrt{n})$. Доказательство основано на применении нового типа локальной теоремы.

Ключевые слова и фразы: сопряженное распределение, локальная теорема, возмущение спектра, монотонная ε -аппроксимация.

1. Introduction and result. The basic goal of this paper is to fill a gap that so far exists in the classic theory of Markov chains. More precisely, the question is the Cramér large deviations for the sums of random variables defined on the successive states of a Markov chain. As such the simplest chain with a finite number of states is taken. So, the case to be in the focus of our attention is much simpler than that considered, say, in the well-known papers of S. Nagaev [12], [13]. It should be noted that in those papers practically all the typical limit theorems earlier proved for sums of i.i.d. variables found their analogues. The exception is the local theorems of the Shepp type which appeared later on (see [16], [18], [5], [11]). The Markov chain analogues of such local theorems as well as certain refinements of the S. Nagaev results can be found in [17].

* Nicolaus Copernicus University, 87-100 Toruń, Chopina 12/18, Poland; e-mail: zssz@mat.uni.torun.pl

The further progress in the theory concerned, say, the aperiodic Harris recurrent Markov chain does not contribute much to the Cramér large deviations.

Let $\{X_k\}_{k \in \mathbb{N}}$ be a homogeneous Markov chain with a finite number of states. We denote them by $1, \dots, d$. In what follows we assume that the transition matrix $\mathbf{P} = (p_{ij})_{1 \leq i, j \leq d}$ is primitive, that is, there exists $k_0 \geq 1$ such that $\mathbf{P}^{k_0}$ is strictly positive. Then the chain is ergodic (see, e.g., in [9, Theorem 9.5.3]). In particular, it means that there exists the stationary distribution $\pi = (\pi_i)_{1 \leq i \leq d}$.

Each function f defined on $S = \{1, \dots, d\}$ determines a sequence of random variables $\{f(X_k)\}_{k \in \mathbb{N}}$. With no loss in generality one may assume that

$$E_\pi f(X_k) = \sum_{i=1}^d f(i) \pi_i = 0, \quad k = 0, 1, 2, \dots$$

Consider the sums

$$S_n = \sum_{k=1}^n f(X_k), \quad n = 1, 2, \dots$$

Let $p_{ij}^{(n)}$ be the entries of matrix $\mathbf{P}^n$. If the function f is such that

$$\sigma^2 = \sum_{i=1}^d f^2(i) \pi_i + 2 \sum_{n=1}^{\infty} \sum_{j,k=1}^d f(j) f(k) \pi_j p_{jk}^{(n)} > 0,$$

then $E(S_n)^2 = n\sigma^2(1 + o(1))$ as $n \rightarrow \infty$ and, moreover, the central limit theorem takes place, that is,

$$\sup_x \left| P \left\{ \frac{S_n}{\sigma\sqrt{n}} < x \right\} - \Phi(x) \right| = o(1),$$

where $\Phi(x)$ is the distribution function of the standard normal law (see, e.g., [3, Theorem 1, § 16 in Part I]). If $x = x_n \rightarrow \infty$, then $P\{S_n \geq x\sigma\sqrt{n}\}$ is called the probability of the large deviation.

There exists a rich literature devoted to the theory of large deviations both on the case of i.i.d. variables (see, e.g., [14]) and on that of Markov chains. As to the latter the case where $x = x_n \asymp \sqrt{n}$ has not been yet studied. This fact seems incredible. Nevertheless, the author could find no reference related to this case. Here the author aims at completing this gap.

We say that the local theorem is of the *Shepp type* if it concerns the weak convergence of the measures $\sqrt{n}P\{S_n \in A\}$. Theorems of this type proved to be very efficient in the large deviation theory of i.i.d. variables. It turns out that Markov chain context requires a new rather specific form of the Shepp type local theorems.

It is well known that the Cramér large deviation technique is based on the so-called *conjugate distributions*. The first question that immediately arises: *how to carry the notion of the conjugate measure over Markov chains?* It should be emphasized that the transformed measure must inherit the Markov property. Furthermore, the reduction of the conjugate measure to the σ -field generated by $X_0, X_1, \dots, X_n$ should be absolutely continuous with respect to the similar reduction of the underlying measure. Moreover, the corresponding Radon–Nikodým derivative should depend only on the sum $S_n = \sum_{k=1}^n f(X_k)$.

In order to satisfy these restrictions consider the primitive matrix

$$\mathbf{L}(s) = (e^{sf(j)} p_{ij})_{1 \leq i, j \leq d}, \quad s \in \mathbb{R}.$$

By the Perron–Frobenius theorem (see, e.g., [9, IX, § 2] or [6, Theorem 8.2.11]) the maximal eigenvalue of $\mathbf{L}(s)$, frequently called also a spectral radius, is simple while the corresponding left and right eigenvectors are strictly positive. We denote these characteristics, respectively, by $\rho(s)$, $u(s)$, and $v(s)$.

Define the entries of the conjugate transition matrix $\mathbf{P}(s)$ as

$$p_{ij}(s) = \frac{e^{sf(j)} p_{ij} v_j(s)}{\sum_{k=1}^d e^{sf(k)} p_{ik} v_k(s)} = \frac{e^{sf(j)} p_{ij} v_j(s)}{\rho(s) v_i(s)}. \quad (1)$$

Obviously, $\mathbf{P}(s)$ is a stochastic matrix and the mentioned above Radon-Nikodým derivative at the point $(i_0, i_1, \dots, i_n)$ has the following form:

$$(v_{i_0}(s))^{-1} \rho^{-n}(s) \exp \left(s \sum_{k=1}^n f(i_k) \right) v_{i_n}(s)$$

and, therefore, really depends on S_n . However, it depends also on the final state of the considered part of the trajectory. So, we have to establish the weak convergence of the measures $P(S_n \in A, X_n = j | X_0 = i)$, that is, to deal with a very special form of the local limit theorem.

It is this specific form of the conjugate measure and the, related to it, special version of the local limit theorem that are the key points of the present paper.

In order to formulate our result we need some definitions. Let $p_{ij}^{(n)}(s)$ be the entries of matrix $\mathbf{P}^n(s)$, where $\mathbf{P}^n(s)$ is n -th power of $\mathbf{P}(s)$, and let $\pi(s)$ be its stationary vector. In what follows we show that $\pi_i(s) = u_i(s) v_i(s) / (u^T(s) v(s))$. Denote

$$\sigma^2(s) = \sum_{i=1}^d \bar{f}^2(i) \pi_i(s) + 2 \sum_{n=1}^{\infty} \sum_{j,k=1}^d \bar{f}(j) \bar{f}(k) \pi_j(s) p_{jk}^{(n)}(s), \quad (2)$$

where $\bar{f}(\cdot) = f(\cdot) - E_{\pi(s)} f(X_0)$, and

$$M(s) = \frac{d}{ds} \ln \rho(s). \quad (3)$$

Later on in Lemma 4 we will prove that $M(s)$ in $(-\infty, \infty)$ is strictly increasing. Furthermore, it maps $(-\infty, \infty)$ onto (f^-, f^+) , where

$$f^- = \min_{1 \leq j \leq d} f(j), \quad f^+ = \max_{1 \leq j \leq d} f(j).$$

Denote by $s(t)$ a unique solution of the equation $M(s) = t$, $f^- < t < f^+$. From the proof of Lemma 4 it follows that

$$M(0) = E_{\pi} f(X_k) = 0.$$

It means that $\text{sign}(s(t)) = \text{sign}(t)$. Consider the so-called deviation function

$$\alpha(t) = \rho(s(t)) e^{-ts(t)}, \quad t \in (f^-, f^+).$$

From the proof of Lemma 4 (see Section 3) it follows that $\alpha(t)$ strictly increases in $(f^-, 0)$ and strictly decreases in $(0, f^+)$. Obviously, $\alpha(0) = 1$. It is convenient to define $A_{ij}(t)$ as

$$A_{ij}(t) = (\sqrt{2\pi} \sigma(s(t)) s(t))^{-1} a_{ij}(s(t)), \quad t \in (f^-, f^+),$$

where

$$a_{ij}(s) = \frac{v_i(s) u_j(s)}{u^T(s) v(s)}.$$

We need the following condition (*): *There exist $\mu, \nu \in \{1, \dots, d\}$ and $n_0 > 1$ such that the conditional distribution of S_{n_0} given $X_{n_0} = \nu$, $X_0 = \mu$ is not lattice.*

Later on we will restate (*) in terms of the so-called characteristic matrix (see Remark 1).

Theorem 1. *Let $\{X_k\}_{k \in \mathbb{N}}$ be irreducible homogeneous aperiodic Markov chain with a finite phase space. If (*) holds, then*

1) $\sigma^2(s) > 0$ for every $s \in \mathbb{R}$ and the solution $s = s(t)$ of the equation $M(s) = t$ exists for $t \in (f^-, f^+)$;

2) for any $i, j \in S$ and $\varepsilon > 0$

$$\sup_{\varepsilon \leq t \leq f^+ - \varepsilon} \left| \frac{P(\sum_{k=1}^n f(X_k) \geq nt; X_n = j | X_0 = i)}{A_{ij}(t) n^{-1/2} \alpha^n(t)} - 1 \right| = o(1) \quad (4)$$

while

$$\sup_{f^- + \varepsilon \leq t \leq -\varepsilon} \left| \frac{P(\sum_{k=1}^n f(X_k) < nt; X_n = j | X_0 = i)}{A_{ij}(t) n^{-1/2} \alpha^n(t)} - 1 \right| = o(1). \quad (5)$$

The stated Theorem 1 has two obvious corollaries.

Corollary 1. *Under the conditions of Theorem 1 for any $i \in S$ and $\varepsilon > 0$*

$$\sup_{\varepsilon \leq t \leq f + \varepsilon} \left| \frac{P(\sum_{k=1}^n f(X_k) \geq nt \mid X_0 = i)}{A_i(t) n^{-1/2} \alpha^n(t)} - 1 \right| = o(1),$$

where $A_i(t) = \sum_{j=1}^d A_{ij}(t)$.

Corollary 2. *Under the conditions of Theorem 1 for any $\varepsilon > 0$*

$$\sup_{\varepsilon \leq t \leq f + \varepsilon} \left| \frac{P(\sum_{k=1}^n f(X_k) \geq nt)}{A(t) n^{-1/2} \alpha^n(t)} - 1 \right| = o(1),$$

where $A(t) = \sum_{i=1}^d p_i A_i(t)$ and $p = (p_i)_{1 \leq i \leq d}$ is an initial distribution.

As to the case of i.i.d. variables we refer to the overview [14]. It is worth comparing the stated results with those presented in [2], [4, Section 3.1.1], and [15].

The paper is organized as follows. In Section 2 we prove a local limit theorem in the mentioned above specific form. In Section 3 we deal with the properties of conjugate measure determined by (1) and give the proof of Theorem 1. This proof gives also an idea how to carry the statement of Theorem 1 over the case of lattice or even vector-valued f .

The author is very grateful to A. Nagaev for his valuable suggestions and encouragement during the preparation of this manuscript.

2. A local limit theorem. Let

$$\hat{\mathbf{P}}(\zeta) = \mathbf{P}\Phi(\zeta),$$

where $\Phi(\zeta) = \text{diag}(e^{i\zeta f(1)}, \dots, e^{i\zeta f(d)})$, $\zeta \in \mathbf{C}$. If $\zeta = \theta \in \mathbf{R}$, then $\hat{\mathbf{P}}(\theta)$ is referred to as a characteristic matrix for $\{f(X_k)\}_{k \in \mathbf{N}}$. For sufficiently small $|\theta|$ we denote by $\lambda(\theta)$ the spectral radius of $\hat{\mathbf{P}}(\theta)$. Let γ be the second in the absolute value eigenvalue of $\mathbf{P}$. Obviously, $\lambda(0) = 1$ and $|\gamma| < 1$.

Let

$$\|\mathbf{A}\| = \max_{1 \leq j \leq d} \sum_{k=1}^d |a_{jk}|$$

denote the maximal row sum matrix norm. The following proposition is a version of Nagaev's spectral theorem.

Proposition 1. *There exists $\xi = \xi(f, \mathbf{P})$, $\xi > 0$, such that for $|\theta| < \xi$ we have $|\lambda(\theta)| > \kappa$ and*

$$\|\hat{\mathbf{P}}^n(\theta) - \lambda^n(\theta) \hat{\mathbf{P}}_1(\theta)\| \leq O(\kappa^n),$$

where $\hat{\mathbf{P}}_1(\theta)$ is the projection of $\mathbf{P}(\theta)$ on eigenspace associated with $\lambda(\theta)$ satisfying $\hat{\mathbf{P}}_1(\theta) \rightarrow \Pi = \psi \pi^T$ as $\theta \rightarrow 0$, $\psi = (1, \dots, 1)^T$, $\kappa = \frac{1}{3} + \frac{2}{3} |\gamma|$.

The proof of Proposition 1 is only a minor alteration of the argument given in [13, p. 71–72].

Lemma 1. *If condition (*) holds, then for any $0 < a < b < \infty$ there exists $k_0 > 1$ such that for $n > k_0$*

$$\sup_{|\theta| \in [a, b]} \|\hat{\mathbf{P}}^n(\theta)\| < c \chi^n,$$

where $0 < \chi < 1$, $c > 1$.

P r o o f. Let m_0 be such that $\min_{1 \leq j \leq d} p_{j\mu}^{(m_0)} \geq \frac{1}{2} \pi_\mu$ and let

$$\inf_{|\theta| \in [a, b]} (p_{\mu\nu}^{(n_0)} - |p_{\mu\nu}^{(n_0)}(\theta)|) = \eta,$$

where $p_{\mu\nu}^{(n_0)}(\theta)$ is the $\mu\nu$ -element of matrix $\hat{\mathbf{P}}^{n_0}(\theta)$. From (*) it follows that $\eta > 0$.

For any $j \in S$ we have

$$\begin{aligned} \sum_{k=1}^d |p_{jk}^{(m_0+n_0)}(\theta)| &\leq \sum_{k=1}^d \sum_{l=1}^d |p_{jl}^{(m_0)}(\theta)| |p_{lk}^{(n_0)}(\theta)| \leq \sum_{k=1}^d \sum_{l=1}^d p_{jl}^{(m_0)} |p_{lk}^{(n_0)}(\theta)| \\ &= \sum_{k=1}^d \sum_{l=1}^d p_{jl}^{(m_0)} \{p_{lk}^{(n_0)} - (p_{lk}^{(n_0)} - |p_{lk}^{(n_0)}(\theta)|)\} \\ &= 1 - \sum_{l=1}^d p_{jl}^{(m_0)} \sum_{k=1}^d (p_{lk}^{(n_0)} - |p_{lk}^{(n_0)}(\theta)|). \end{aligned}$$

It implies that

$$\begin{aligned} \sum_{k=1}^d |p_{jk}^{(m_0+n_0)}(\theta)| &\leq 1 - p_{j\mu}^{(m_0)} \sum_{k=1}^d (p_{\mu k}^{(n_0)} - |p_{\mu k}^{(n_0)}(\theta)|) \\ &\leq 1 - \min_{1 \leq j \leq d} p_{j\mu}^{(m_0)} \sum_{k=1}^d (p_{\mu k}^{(n_0)} - |p_{\mu k}^{(n_0)}(\theta)|) \\ &\leq 1 - \frac{1}{2} \pi_\mu (p_{\mu\nu}^{(n_0)} - |p_{\mu\nu}^{(n_0)}(\theta)|) \leq 1 - \frac{1}{2} \pi_\mu \eta = \chi_1 \end{aligned}$$

provided $a \leq |\theta| \leq b$. Thus,

$$\sup_{a \leq |\theta| \leq b} \|\widehat{\mathbf{P}}^{(m_0+n_0)}(\theta)\| = \sup_{a \leq |\theta| \leq b} \max_{1 \leq j \leq d} \sum_{k=1}^d |p_{jk}^{(m_0+n_0)}(\theta)| \leq \chi_1 < 1.$$

Hence, for $n > k_0 = m_0 + n_0$

$$\sup_{a \leq |\theta| \leq b} \|\widehat{\mathbf{P}}^n(\theta)\| \leq \sup_{a \leq |\theta| \leq b} \|\widehat{\mathbf{P}}^{(m_0+n_0)}(\theta)\|^{[n/(m_0+n_0)]} \leq \chi_1^{[n/k_0]} \leq (\chi_1^{1/k_0})^{n-1}.$$

It remains to set $\chi = \chi_1^{1/k_0}$ and $c = 1/\chi_1$. Lemma 1 is proved.

Lemma 2. If condition $(*)$ holds, then $E_\pi |S_n| \rightarrow \infty$ as $n \rightarrow \infty$.

P r o o f. Set $\varphi_n(\theta) = E_\pi \exp\{i\theta S_n\}$. It is easily seen that

$$|\varphi_n(\theta)| = \left| \sum_{j=1}^d \sum_{k=1}^d \pi_j p_{jk}^{(n)}(\theta) \right| \leq \sum_{j=1}^d \sum_{k=1}^d \pi_j |p_{jk}^{(n)}(\theta)| \leq \max_{1 \leq j \leq d} \sum_{k=1}^d |p_{jk}^{(n)}(\theta)| = \|\widehat{\mathbf{P}}^n(\theta)\|.$$

Hence, by Proposition 1 for any $0 < a < b < \infty$ and $n > k_0 > 1$

$$\sup_{|\theta| \in [a,b]} |\varphi_n(\theta)| < c \chi^n, \tag{6}$$

where $0 < \chi < 1$, $c > 1$.

By Lemma 3.6.1 in [10] for any random variable Y and for any $K > 0$ we have

$$\int_0^K (P\{Y \geq y\} + P\{Y < -y\}) dy = \frac{1}{\pi} \int_{-\infty}^\infty \frac{1 - \cos K\theta}{\theta^2} (1 - E \exp\{i\theta Y\}) d\theta.$$

Let $\mathcal{N}$ be a $(0,1)$ -normally distributed random variable independent of S_n . By the continuity of $\text{Law}(S_n + \mathcal{N})$

$$\begin{aligned} E|S_n| + E|\mathcal{N}| &\geq E|S_n + \mathcal{N}| \geq \int_0^K \mathbf{P}\{|S_n + \mathcal{N}| > y\} dy \\ &= \frac{1}{\pi} \int_{-\infty}^\infty \frac{1 - \cos K\theta}{\theta^2} (1 - \text{Re } \varphi_n(\theta) e^{-\theta^2/2}) d\theta. \end{aligned}$$

So, taking into account (6), we obtain

$$E|S_n| + \sqrt{\frac{2}{\pi}} \geq \frac{2}{\pi} \int_a^b \frac{1 - \cos K\theta}{\theta^2} (1 - c \chi^n e^{-\theta^2/2}) d\theta.$$

Letting, first, $n \rightarrow \infty$ and, then, $a \rightarrow 0$, $b \rightarrow \infty$ yields

$$\liminf_n E|S_n| + \sqrt{\frac{2}{\pi}} \geq \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos K\theta}{\theta^2} d\theta = K.$$

Since $K > 0$ is arbitrary the lemma is proved.

Set

$$\lambda^{(k)} = \frac{1}{i^k} \frac{d^k}{d\zeta^k} \lambda(\zeta) \Big|_{\zeta=0}.$$

By (2.23) in [12] $\lambda^{(1)} = 0$ while by (2.21) ibidem

$$\sigma^2 = \lambda^{(2)} = E_{\pi} f^2(X_0) + 2 \sum_{n=1}^{\infty} E_{\pi} f(X_0) f(X_n). \quad (7)$$

Lemma 3. *If condition (*) holds, then $\sigma^2 > 0$.*

P r o o f. Assume that $\sigma^2 = 0$. Then from [3, Part I, § 16, Corollary 1] it follows that S_n is bounded in probability. On the other hand by Lemma 2 we have $\sigma_n^2 = E_{\pi} S_n^2 \rightarrow \infty$. It is well known [7, Lemma 18.5.1] that $\{\sigma_n^{-2} S_n^2\}$ is uniformly integrable sequence. So, we have, simultaneously, $\sigma_n^{-2} S_n^2 \rightarrow 0$ in probability and $\sigma_n^{-2} S_n^2 \rightarrow 0$ in L^1 . Since this is impossible we conclude that $\sigma^2 > 0$. Lemma 3 is proved.

R e m a r k 1. As it was mentioned above condition (*) can be restated in terms of $\lambda(\theta)$. More precisely, (*) holds if and only if $|\lambda(\theta)| < 1$ for every $\theta \neq 0$. It follows immediately from Lemma 1 since $\limsup_n \|\hat{\mathbf{P}}^n(\theta)\|^{1/n} = |\lambda(\theta)|$. However, the verification of (*) seems to be much easier.

The following example shows that if (*) is not fulfilled, then even the central limit theorem does not take place.

E x a m p l e 1. Let $d = 4$ and $f(1) = 5$, $f(2) = -3 - \pi$, $f(3) = -2 + \pi$, $f(4) = 0$. Consider

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The straightforward calculations show that $\mathbf{P}^n > 0$ for $n \geq 10$, that is, $\mathbf{P}$ is primitive. But it does not satisfy (*). To see this note that $\det(\lambda(\theta) \mathbf{I} - \hat{\mathbf{P}}(\theta)) = (\lambda(\theta) - 1)(\lambda^3(\theta) + \lambda^2(\theta) + \lambda(\theta) + \frac{1}{2})$ so $\lambda(\theta) \equiv 1$. Thus, $\sigma^2 = 0$ and σ_n^2 is bounded.

Let m denote the Lebesgue measure. It is worth comparing the statement of the following theorem with its i.i.d. counterpart (see, e.g., [11, Lemma 5]).

Theorem 2. *Let condition (*) holds, then for any bounded Borel set A such that $m(\partial A) = 0$*

$$P\left(\sum_{k=1}^n f(X_k) - r \in A; X_n = j \mid X_0 = i\right) = \frac{m(A)}{\sigma\sqrt{2\pi n}} \exp\left\{-\frac{r^2}{2\sigma^2 n}\right\} \pi_j + o\left(\frac{1}{\sqrt{n}}\right)$$

uniformly in $r \in \mathbf{R}$.

P r o o f. Let g be continuous function with a compact support. By [1] g admits a monotone ε -approximation, i.e., for any $\varepsilon > 0$ there exist functions g^+, g^- such that $g^- \leq g \leq g^+$, $g^+ - g^- \leq \varepsilon$, g^+, g^- are absolutely integrable,

$$\int_{-\infty}^{\infty} (g^+(x) - g^-(x)) dx < \varepsilon$$

and

$$\hat{g}^+(\theta) = \int_{-\infty}^{\infty} e^{ix\theta} g^+(x) dx, \quad \hat{g}^-(\theta) = \int_{-\infty}^{\infty} e^{ix\theta} g^-(x) dx$$

are also compactly supported.

Denote

$$E_{ij}g(S_n - r) = \int_{-\infty}^{\infty} g(x - r) dP(S_n < x; X_n = j | X_0 = i).$$

Then

$$E_{ij}^- = E_{ij}g^-(S_n - r) \leq E_{ij}g(S_n - r) \leq E_{ij}g^+(S_n - r) = E_{ij}^+.$$

The estimates of E_{ij}^- and E_{ij}^+ are quite similar. So, we may confine ourself with estimating any of them, say, E_{ij}^+ . Since $\widehat{g}^+(\theta)$ is compactly supported and

$$p_{ij}^{(n)}(\theta) = \int_{-\infty}^{\infty} e^{i\theta x} dP(S_n < x; X_n = j | X_0 = i),$$

there exists $K > 0$ such that

$$E_{ij}^+ = E_{ij}g^+(S_n - r) = \frac{1}{2\pi} \int_{|\theta| < K} e^{-i\theta r} \widehat{g}^+(-\theta) p_{ij}^{(n)}(\theta) d\theta.$$

Represent E_{ij}^+ as $E_{ij}^+ = E'_{ij} + E''_{ij}$, where

$$E'_{ij} = \frac{1}{2\pi} \int_{|\theta| < \xi} e^{-i\theta r} \widehat{g}^+(-\theta) p_{ij}^{(n)}(\theta) d\theta, \quad E''_{ij} = \frac{1}{2\pi} \int_{\xi < |\theta| < K} e^{-i\theta r} \widehat{g}^+(-\theta) p_{ij}^{(n)}(\theta) d\theta.$$

In view of Lemma 1 we have

$$|E''_{ij}| \leq \frac{1}{2\pi} \int_{\xi < |\theta| < K} |\widehat{g}^+(-\theta)| |p_{ij}^{(n)}(\theta)| d\theta = O(\chi^n). \quad (8)$$

In view of (7) we easily obtain that as $|\theta| \rightarrow 0$

$$\lambda(\theta) = 1 - \frac{\sigma^2 \theta^2}{2} + o(\theta^2). \quad (9)$$

By Proposition 1

$$E'_{ij} = \frac{1}{2\pi} \int_{|\theta| < \xi} e^{-i\theta r} \widehat{g}^+(-\theta) \lambda^n(\theta) (\widehat{\mathbf{P}}_1(\theta))_{ij} d\theta + O(\kappa^n), \quad (10)$$

where $(\widehat{\mathbf{P}}_1(\theta))_{ij}$ is the respective element of $\widehat{\mathbf{P}}_1(\theta)$. The following argument is quite routine. First, we represent the integral in (10) as $I_1 + I_2$, where

$$\begin{aligned} I_1 &= \frac{1}{2\pi} \int_{|\theta| < Tn^{-1/2}} e^{-i\theta r} \widehat{g}^+(-\theta) \lambda^n(\theta) (\widehat{\mathbf{P}}_1(\theta))_{ij} d\theta, \\ I_2 &= \frac{1}{2\pi} \int_{Tn^{-1/2} < |\theta| < \xi} e^{-i\theta r} \widehat{g}^+(-\theta) \lambda^n(\theta) (\widehat{\mathbf{P}}_1(\theta))_{ij} d\theta. \end{aligned}$$

It is easily seen that

$$\begin{aligned} I_1 &= \frac{\widehat{g}^+(0) \pi_j}{2\pi} \int_{|\theta| < Tn^{-1/2}} e^{-i\theta r} \lambda^n(\theta) d\theta (1 + o(1)) \\ &= \frac{\widehat{g}^+(0) \pi_j}{2\pi \sigma \sqrt{n}} \int_{|\theta| < T\sigma} e^{-i\theta r / (\sigma \sqrt{n})} \lambda^n\left(\frac{\theta}{\sigma \sqrt{n}}\right) d\theta (1 + o(1)). \end{aligned}$$

By virtue of (9)

$$I_1 = \frac{\widehat{g}^+(0) \pi_j}{\sqrt{n}} \left(\frac{1}{\sigma} \phi\left(\frac{r}{\sigma \sqrt{n}}\right) + \omega(T) \right) (1 + o(1)), \quad (11)$$

where $\omega(T) \rightarrow 0$ as $T \rightarrow \infty$ and $\phi = \Phi'$. Further, from (9) it follows that for $|\theta| \leq \xi$

$$|\lambda(\theta)| \leq 1 - c|\theta|^2.$$

Hence

$$|I_2| \leq c \int_{Tn^{-1/2} \leq |\theta| < \xi} |\lambda(\theta)|^n d\theta \leq c \int_{|\theta| \geq Tn^{-1/2}} e^{-nc\theta^2} d\theta \leq n^{-1/2} \omega(T). \quad (12)$$

Since T is arbitrary, (9)–(12) yield

$$E'_{ij} = \frac{\widehat{g}^+(0) \pi_j}{\sqrt{n}} \frac{1}{\sigma} \phi\left(\frac{r}{\sigma\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right).$$

Combining this relation with (8) we obtain

$$E_{ij}^+ = \frac{\widehat{g}^+(0) \pi_j}{\sqrt{n}} \frac{1}{\sigma} \phi\left(\frac{r}{\sigma\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right).$$

Similarly,

$$E_{ij}^- = \frac{\widehat{g}^-(0) \pi_j}{\sqrt{n}} \frac{1}{\sigma} \phi\left(\frac{r}{\sigma\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right).$$

Both the relations hold uniformly in $r \in \mathbf{R}$. It remains to note that

$$\left| \widehat{g}^+(0) - \int_{-\infty}^{\infty} g(x) dx \right| \leq \varepsilon$$

and to recall that ε is arbitrary. Theorem 2 is proved.

The limit theorems for Markov chains in [13] assume a condition (A) which in our setting looks as follows: *There exists $\emptyset \neq C \subset S$ such that $p_{ij} \neq 0$ for $(i, j) \in C \times C$.* The next example shows that this condition can be violated while Theorem 2 remains valid.

Example 2. Let $d = 5$ and $f(1) = 1$, $f(2) = \pi$, $f(3) = -4$, $f(4) = 2$, $f(5) = -1$. Consider

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

It is easily verified that for $n \geq 7$ we have $\mathbf{P}^n > 0$, that is, $\mathbf{P}$ is primitive. Furthermore,

$$P(f(X_1) + \cdots + f(X_5) = 0; X_5 = 4 \mid X_0 = 1) = \frac{1}{4},$$

$$P(f(X_1) + \cdots + f(X_5) = \pi + 1; X_5 = 4 \mid X_0 = 1) = \frac{1}{16},$$

$$P(f(X_1) + \cdots + f(X_5) = \pi - 1; X_5 = 4 \mid X_0 = 1) = \frac{1}{8}.$$

So $(*)$ holds.

In particular, from Theorem 2 it follows that

$$\sigma\sqrt{2\pi n} \int_{-\infty}^{\infty} g(x) P\left(\sum_{k=1}^n f(X_k) \in dx; X_n = j \mid X_0 = i\right) \rightarrow \pi_j \int_{-\infty}^{\infty} g(x) m(dx) \quad (13)$$

for any continuous function g with a compact support.

3. Proof of Theorem 1. Consider the stochastic matrices $\mathbf{P}$ and $\mathbf{P}(s)$. We remind that $\mathbf{P}(s)$ is defined in (1). Given an initial state i the matrices determine two measures $P_i^{(n)}$ and $P_i^{(n,s)}$ defined on the strings $(i, i_1, \dots, i_n)$. As it was stated in introduction the measures $P_i^{(n)}$ and $P_i^{(n,s)}$ are absolutely continuous with respect to each other. Furthermore,

$$\frac{dP_i^{(n,s)}}{dP_i^{(n)}} = (v_i(s))^{-1} \rho^{-n}(s) \exp\left(s \sum_{k=1}^n f(i_k)\right) v_{i_n}(s).$$

We remind that the stationary distribution of the chain controlled by $\mathbf{P}(s)$ is denoted by $\pi(s)$. According to (1) we have

$$\begin{aligned} \sum_{i=1}^d \frac{u_i(s) v_i(s)}{u^T(s) v(s)} p_{ij}(s) &= \sum_{i=1}^d \frac{u_i(s) v_i(s)}{u^T(s) v(s)} e^{sf(j)} p_{ij} \frac{v_j(s)}{\rho(s) v_i(s)} \\ &= \frac{v_j(s)}{u^T(s) v(s) \rho(s)} \sum_{i=1}^d u_i(s) e^{sf(j)} p_{ij} = \frac{u_j(s) v_j(s)}{u^T(s) v(s)}. \end{aligned}$$

It implies that

$$\pi_j(s) = \frac{u_j(s) v_j(s)}{u^T(s) v(s)} = \sum_{i=1}^d \pi_i(s) p_{ij}(s).$$

Let $\sigma^2(s)$ and $M(s)$ be as in (2) and (3).

Lemma 4. *If condition (*) holds, then $\sigma^2(s) > 0$ for every $s \in \mathbf{R}$ and $M(s)$ maps one-to-one $\mathbf{R}$ onto (f^-, f^+) .*

P r o o f. For any $n \geq 1$ the measures $P_i^{(n)}$ and $P_i^{(n,s)}$ are mutually absolutely continuous. It implies that for both the chains (*) holds. Moreover, in view of Lemma 3, $\sigma^2(s) > 0$ if and only if $\sigma^2 > 0$. According to [9, Exercise 3, § 8, VII] (see also [8, § 2.2, Chapter II])

$$\frac{d}{ds} \rho(s) = u^T(s) \left(\frac{d}{ds} \mathbf{L}(s) \right) v(s), \quad (14)$$

where $\mathbf{L}(s)$ is as in introduction. Obviously,

$$\sum_{i=1}^d \sum_{j=1}^d \frac{u_i(s)}{u^T(s) v(s)} f(j) e^{sf(j)} p_{ij} v_j(s) = \rho(s) \sum_{j=1}^d \pi_j(s) f(j) = \rho(s) E_{\pi(s)} f(X_0),$$

where the expectation $E_{\pi(s)}(\cdot)$ corresponds to $\pi(s)$. Together with (14) it yields

$$M(s) = \frac{d}{ds} \ln \rho(s) = E_{\pi(s)} f(X_0).$$

Similarly, it can be proved that

$$M'(s) = \frac{d^2}{ds^2} \ln \rho(s) = \sigma^2(s)$$

(see, e.g., [9, Exercise 3, § 8, VII] and [8, § 2.2, Chapter II]). Since $\sigma^2(s) > 0$ the function $M(s)$ is strictly increasing.

By Theorem 8.1.22 in [6] we have

$$\min_{1 \leq i \leq d} \left(\sum_{j=1}^d e^{sf(j)} p_{ij} \right)^{1/s} \leq \rho^{1/s}(s) \leq \max_{1 \leq i \leq d} \left(\sum_{j=1}^d e^{sf(j)} p_{ij} \right)^{1/s}$$

whence

$$\lim_{s \rightarrow \infty} \rho^{1/s}(s) = \exp\{\max\{f(j) \mid 1 \leq j \leq d\}\}.$$

Similarly,

$$\lim_{s \rightarrow -\infty} \rho^{1/s}(-s) = \exp\{\min\{f(j) \mid 1 \leq j \leq d\}\}.$$

Thus,

$$\lim_{s \rightarrow \infty} \frac{1}{s} \ln \rho(s) = \max_{1 \leq j \leq d} f(j) \quad \text{and} \quad - \lim_{s \rightarrow -\infty} \frac{1}{s} \ln \rho(s) = \min_{1 \leq j \leq d} f(j).$$

The last two relations together with (3) lead to

$$\lim_{s \rightarrow \infty} \frac{1}{s} \ln \rho(s) = \lim_{s \rightarrow \infty} \frac{1}{s} \int_0^s M(y) dy = \max_{1 \leq j \leq d} f(j).$$

Therefore,

$$\lim_{s \rightarrow \infty} M(s) = \max_{1 \leq j \leq d} f(j) = f^+.$$

Similarly,

$$\lim_{s \rightarrow -\infty} M(s) = \min_{1 \leq j \leq d} f(j) = f^-.$$

The proof is completed.

Lemma 5. For any $i, j \in S$, $n \geq 1$ and $s \in \mathbf{R}_+$ we have

$$\begin{aligned} P_i^{(n)} \left(\sum_{k=1}^n f(X_k) \geq nt; X_n = j \right) \\ = (e^{-st} \rho(s))^n \frac{v_i(s)}{v_j(s)} \int_0^\infty e^{-sx} P_i^{(n,s)} \left(\sum_{k=1}^n \bar{f}(X_k) \in dx; X_n = j \right). \end{aligned} \quad (15)$$

Similarly, for $s \in \mathbf{R}_-$

$$\begin{aligned} P_i^{(n)} \left(\sum_{k=1}^n f(X_k) < nt; X_n = j \right) \\ = (e^{-st} \rho(s))^n \frac{v_i(s)}{v_j(s)} \int_{-\infty}^0 e^{-sx} P_i^{(n,s)} \left(\sum_{k=1}^n \bar{f}(X_k) \in dx; X_n = j \right). \end{aligned} \quad (16)$$

P r o o f. It is easy to see that for any string $i, i_1, i_2, \dots, i_{n-1}, j$ we have

$$p_{ii_1}(s) p_{i_1 i_2}(s) \cdots p_{i_{n-1} j}(s) v_i(s) \rho^n(s) = e^{s(f(i_1) + \cdots + f(i_{n-1}) + f(j))} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} v_j(s).$$

Summing over the corresponding strings we obtain

$$\begin{aligned} \sum_{\{i_1, \dots, i_{n-1} \mid \sum_{k=1}^n f(i_k) \geq nt\}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} &= \frac{v_i(s)}{v_j(s)} e^{-nst} (\rho(s))^n \\ \times \sum_{\{i_1, \dots, i_{n-1} \mid \sum_{k=1}^n f(i_k) \geq nt\}} e^{-s(f(i_1) - t + \cdots + f(j) - t)} &p_{ii_1}(s) p_{i_1 i_2}(s) \cdots p_{i_{n-1} j}(s), \end{aligned}$$

$i_n = j$. The last relation is, obviously, equivalent to (15). The relation (16) is proved similarly.

Lemma 6. Let measures $Q_{ij}^{(n,s)}$ be defined as

$$Q_{ij}^{(n,s)}(A) = \sigma(s) \sqrt{2\pi n} P_i^{(n,s)} \left(\sum_{k=1}^n f(X_k) \in A; X_n = j \right).$$

Then for any continuous function g with a compact support the convergence

$$\int_{-\infty}^{\infty} g(x) Q_{ij}^{(n,s)}(dx) \longrightarrow \pi_j(s) \int_{-\infty}^{\infty} g(x) m(dx)$$

is uniform in s , $|s| \leq s_0 < \infty$.

P r o o f. Let $g(x)$ be a continuous function with compact support. In order to prove Lemma 6 it suffices to recall that all the estimates leading to (13) hold uniformly in s , $|s| \leq s_0$, $s_0 < \infty$.

Having proven Lemma 5 and Lemma 6 we can easily prove Theorem 1. The functions $g_s(x) = e^{-sx}$, $0 < s_- \leq s \leq s_+ < \infty$, are equicontinuous. So by Lemma 6 and Lemma 5 we arrive to (4). The statement (5) is proved similarly. The theorem is proved.

REFERENCES

1. *Bretagnolle J., Dacunha-Castelle D.* Théorèmes limites à distance finie pour les marches aléatoires. — *Ann. Inst. H. Poincaré*, 1968, v. 4, № 1, p. 25–73.
2. *Bryc W.* On large deviations for uniformly strong mixing sequences. — *Stochastic Process. Appl.*, 1992, v. 41, № 2, p. 191–202.
3. *Chung K. L.* Markov Chains with Stationary Transition Probabilities. New York: Springer-Verlag, 1967, 301 p.
4. *Dembo A., Zeitouni O.* Large Deviations Techniques and Applications. Boston: Jones and Bartlett, 1993, 346 p.
5. *Feller W.* On regular variation and local limit theorems. — *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (Berkeley, 1965/66)*. V. II. Part 1. Ed by J. Neyman. Berkeley: Univ. California Press, 1967, p. 373–388.
6. *Хорн Р., Джонсон Ч.* Матричный анализ. М.: Мир, 1989, 656 с.
7. *Ибрагимов И. А., Линник Ю. В.* Независимые и стационарно связанные величины. М.: Наука, 1965, 524 с.
8. *Като Т.* Теория возмущений линейных операторов. М.: Мир, 1972, 740 с.
9. *Ланкастер П.* Теория матриц. М.: Наука, 1982, 269 с.
10. *Лукач Е.* Характеристические функции. М.: Наука, 1979, 423 с.
11. *Нагаев А. В.* Локальная предельная теорема для числа восстановлений. — *Литов. матем. сб.*, 1970, т. 10, № 1, с. 109–119.
12. *Нагаев С. В.* Некоторые предельные теоремы для однородных цепей Маркова. — *Теория вероятн. и ее примен.*, 1957, т. 2, в. 4, с. 389–416.
13. *Нагаев С. В.* Уточнение предельных теорем для однородных цепей Маркова. — *Теория вероятн. и ее примен.*, 1961, т. 6, в. 1, с. 67–86.
14. *Nagaev S. V.* Large deviations of sums of independent random variables. — *Ann. Probab.*, 1979, v. 7, № 5, p. 745–789.
15. *Scheffel P., v. Weizsäcker H.* On risk rates and large deviations in finite Markov chain experiments. — *Math. Methods Statist.*, 1997, v. 6, № 3, p. 293–312.
16. *Shepp L. A.* A local limit theorem. — *Ann. Math. Statist.*, 1964, v. 35, № 1, p. 419–423.
17. *Сираждинов С. Х., Форманов Ш. К.* Предельные теоремы для сумм случайных векторов, связанных в цепь Маркова. Ташкент: Фан, 1979, 171 с.
18. *Stone C. J.* A local limit theorem for nonlattice multidimensional distribution functions. — *Ann. Math. Statist.*, 1965, v. 36, № 2, p. 546–551.

Поступила в редакцию
11.X.2001