Z. Wahrscheinlichkeitstheorie verw. Gebiete
35, 57–63 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On the Rate of Convergence in the Central Limit Theorem for Markov-Chains

D. Landers and L. Rogge

Mathematisches Institut der Universität Köln, Weyertal 86-90, D-5000 Köln 41
Fachbereich Mathematik der Universität Konstanz, Postfach 7733, D-7750 Konstanz
Federal Republic of Germany

1. Introduction

All results concerning the accuracy of the normal approximation for sums of not necessarily independent random variables assume Doeblin's condition or the condition of φ-mixing (see e.g. [1, 3, 5, 7, 9]). Both assumptions mean in some sense that the random variables are "asymptotically independent", and they are rarely fulfilled for Markov-chains.

Using Doeblin's condition or the condition of φ-mixing the rate of convergence to the normal distribution obtained in some of the papers cited above is of order $n^{-1/2}$. The authors do not know of any results on the accuracy of the normal approximation holding without such conditions. In this paper we prove under weak moment conditions that for Markov-chains the normal approximation is of order $n^{-\alpha}$ for each $\alpha < 1/4$.

2. Notations

Let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, where $\mathbb{N}$ is the set of natural numbers.

Let $(\Omega, \mathscr{A}, P)$ be a probability space and X_n, $n \in \mathbb{N}_0$, be a positive recurrent irreducible Markov-chain on $(\Omega, \mathscr{A}, P)$ with countable state space I. For each $\nu \in \mathbb{N}$, $i \in E$, denote by $\tau_\nu^{(i)}$ the time of the ν-th entrance into the state i, by $r_\nu^{(i)} = \tau_{\nu+1}^{(i)} - \tau_\nu^{(i)}$ the ν-th return time of the state i and by $l_n^{(i)}$ the number of entrances into the state i up to time n.

If $\varphi: I \to \mathbb{R}$ define

$$\varphi_\nu^{(i)}(\omega) = \sum_{k=\tau_\nu^{(i)}(\omega)}^{\tau_{\nu+1}^{(i)}(\omega)-1} \varphi(X_k(\omega)), \ \omega \in \Omega.$$

It is well known that the functions $\varphi_\nu^{(i)}$, $\nu \in \mathbb{N}$, are independent and identically distributed.

For $\varphi\equiv 1$ we have $\varphi_\nu^{(i)}=r_\nu^{(i)}$. For $\varphi=1_{\{i\}}$ we have $\varphi_\nu^{(i)}\equiv 1$. Let $\pi_i=\frac{1}{\mu_i}$, where μ_i is the mean recurrence time of the state i. Since X_n, $n\in\mathbb{N}_0$, is positive recurrent we have $\pi_i>0$.

If $|\varphi|_\nu^{(i)}$ is P-integrable then $\sum_{j\in I}\pi_j\,\varphi(j)$ is absolutely convergent and

$$P[\varphi_\nu^{(i)}]=\frac{1}{\pi_i}S(\varphi)$$

where $S(\varphi)=\sum_{j\in I}\pi_j\,\varphi(j)$.

$L_r(P)$ is the space of all $\mathscr{A}$-measurable functions $f\colon \Omega\to\mathbb{R}$ with $P[|f|^r]<\infty$, $r\geqq 1$.

Φ denotes the distribution function of the normal distribution with mean 0 and variance 1.

If $a_n, b_n\in\mathbb{R}$, $n\in\mathbb{N}$, we write

$$a_n=O(b_n) \text{ iff } \sup_{n\in\mathbb{N}}\left|\frac{a_n}{b_n}\right|<\infty.$$

3. The Results

The following Theorem is an essential tool for the proof of our main Theorem 3. It seems to the authors that Theorem 1 is of some interest of its own. It shows that for the number $I_n^{(i)}$ of entrances into the state i up to time n the approximation by the normal distribution is of order $n^{-1/2}$. So it sharpens the result of our Theorem 3 for a special case, namely for $\varphi=1_{\{i\}}$. In this case we have $I_n^{(i)}(\omega)=\sum_{\nu=0}^{n}\varphi(X_\nu(\omega))$ and $\operatorname{var}(\pi_i\,r_\nu^{(i)})=\operatorname{var}(\varphi_\nu^{(i)}-S(\varphi)\,r_\nu^{(i)})$.

Theorem 1. *Let X_n, $n\geqq 0$, be a positive recurrent, irreducible Markov-chain and let the state $i\in I$ be fixed. Assume that*

(a) $\tau_1^{(i)}\in L_1(P)$, $r_\nu^{(i)}\in L_3(P)$,

(b) $\sigma_i^2=\operatorname{var}(\pi_i\cdot r_\nu^{(i)})>0$.

Then

$$\sup_{t\in\mathbb{R}}\left|P\left\{\omega\colon \frac{I_n^{(i)}(\omega)-n\,\pi_i}{\sigma_i\sqrt{n\,\pi_i}}<t\right\}-\Phi(t)\right|=O(n^{-1/2}).$$

Proof. Since $1-\Phi(t)\leqq\frac{1}{t}e^{-\frac{t^2}{2}}$ for $t>0$ we have $\Phi(-n^{1/4})=1-\Phi(n^{1/4})=O(n^{-1/2})$. Hence it suffices to prove

$$\sup_{|t|\leqq n^{1/4}}\left|P\left\{\omega\colon \frac{I_n^{(i)}(\omega)-n\,\pi_i}{\sigma_i\sqrt{n\,\pi_i}}<t\right\}-\Phi(t)\right|=O(n^{-1/2}).$$

This is equivalent to

$$\sup_{|t|\leqq n^{1/4}}\left|P\left\{\omega\colon \frac{I_n^{(i)}(\omega)-n\,\pi_i}{\sigma_i\sqrt{n\,\pi_i}}\geqq t\right\}-\Phi(-t)\right|=O(n^{-1/2}). \tag{1}$$

Let $k_n(t)=\langle t\,\sigma_i\sqrt{n\,\pi_i}+n\,\pi_i\rangle$, where $\langle x\rangle=\min\{k\in\mathbb{N}\colon x\leqq k\}$. Then there exist $n_0\in\mathbb{N}$ and $c_1, c_2>0$, $c_1<c_2$ such that

$$n\geqq n_0,\quad |t|\leqq n^{1/4}\quad \text{imply}\quad k_n(t)\geqq 2,\quad c_1\leqq\frac{k_n(t)-1}{n}\leqq c_2. \tag{2}$$

We have for $n\geqq n_0$

$$P\left\{\omega\colon \frac{\mathrm{I}_n^{(i)}(\omega)-n\,\pi_i}{\sigma_i\sqrt{n\,\pi_i}}\geqq t\right\}=P\{\omega\colon \mathrm{I}_n^{(i)}(\omega)\geqq k_n(t)\}$$
$$=P\{\omega\colon \tau_1^{(i)}(\omega)+r_1^{(i)}(\omega)+\cdots+r_{k_n(t)-1}^{(i)}(\omega)\leqq n\}$$
$$=P\left\{\omega\colon \frac{\tau_1^{(i)}(\omega)+r_1^{(i)}(\omega)+\cdots+r_{k_n(t)-1}^{(i)}(\omega)-(k_n(t)-1)\,\mu_i}{\sigma_i\,\mu_i\sqrt{k_n(t)-1}}\leqq a_n(t)\right\}$$

where

$$a_n(t)=\frac{1}{\sigma_i\,\mu_i\sqrt{k_n(t)-1}}\,(n-(k_n(t)-1)\,\mu_i)\quad\text{and}\quad \mu_i=\frac{1}{\pi_i}.$$

Since $(\tau_1^{(i)}, r_1^{(i)}, r_2^{(i)}, \ldots)$ is a sequence of independent functions and $r_\nu^{(i)}\in L_3(P)$, $\nu\in\mathbb{N}$, are identically distributed with mean μ_i and variance $(\mu_i\,\sigma_i)^2$ and since $\tau_1^{(i)}$ is integrable a little modification of the Theorem of Berry-Esseen yields together with (2) that for $n\geqq n_0$

$$\sup_{|t|\leqq n^{1/4}}\left|P\left\{\omega\colon \frac{\tau_1^{(i)}(\omega)+r_1^{(i)}(\omega)+\cdots+r_{k_n(t)-1}^{(i)}(\omega)-(k_n(t)-1)\,\mu_i}{\mu_i\,\sigma_i\sqrt{k_n(t)-1}}\leqq a_n(t)\right\}\right.$$
$$\left.-\Phi(a_n(t))\right|=O(n^{-1/2}).$$

Therefore it suffices to prove

$$\sup_{|t|\leqq n^{1/4}}|\Phi(a_n(t))-\Phi(-t)|=O(n^{-1/2}). \tag{3}$$

Since $\pi_i\,\mu_i=1$ we have for $n\geqq n_0$ that

$$(k_n(t)-1)\,\mu_i=n+t\,\sigma_i\sqrt{n\,\mu_i}+r_n(t)$$

with $|r_n(t)|\leqq\mu_i$. Hence

$$a_n(t)=-\frac{t\sqrt{n}}{\sqrt{(k_n(t)-1)\,\mu_i}}+O(n^{-1/2})$$
$$=-\frac{t}{\sqrt{1+b_n(t)}}+O(n^{-1/2}) \tag{4}$$

with $b_n(t)=\frac{1}{\sqrt{n}}\,(t\,\sigma_i\sqrt{\mu_i}+r_n'(t))$ where $|r_n'(t)|\leqq\mu_i/\sqrt{n}$. There exists $n_1\in\mathbb{N}$ such that

$$n\geqq n_1,\quad |t|\leqq n^{1/4}\quad\text{implies}\quad \tfrac{1}{2}\leqq 1+b_n(t)\leqq\tfrac{3}{2}. \tag{5}$$

To prove (3) it suffices to show according to (4) that

$$\sup_{|t|\leqq n^{1/4}}\left|\Phi\left(\frac{t}{\sqrt{1+b_n(t)}}\right)-\Phi(t)\right|=O(n^{-1/2}). \tag{6}$$

Using (5) we have for all $n \geqq n_1$, $|t| \leqq n^{1/4}$ that

$$\left|\Phi\left(\frac{t}{\sqrt{1+b_n(t)}}\right)-\Phi(t)\right| \leqq |t|\left|\frac{1}{\sqrt{1+b_n(t)}}-1\right| \cdot \max\left(e^{-\frac{t^2}{2}}, e^{-\frac{1}{2}\frac{t^2}{1+b_n(t)}}\right)$$

$$\leqq |t|\left|\frac{1}{\sqrt{1+b_n(t)}}-1\right| \cdot e^{-\frac{1}{3}t^2}.$$

Since $\left|\frac{1}{\sqrt{1+b_n(t)}}-1\right| \leqq 2|b_n(t)|$ we obtain for $n \geqq n_1$ and $|t| \leqq n^{\frac{1}{4}}$ that

$$\left|\Phi\left(\frac{t}{\sqrt{1+b_n(t)}}\right)-\Phi(t)\right| \leqq 2|t||b_n(t)| \cdot e^{-\frac{1}{3}t^2} \leqq \frac{1}{\sqrt{n}} 2|t| e^{-\frac{1}{3}t^2}\left(|t|\sigma_i\sqrt{\mu_i}+\mu_i\right)$$

which proves (6).

Remark 2. Under the assumptions of Theorem 1 we have

$$P\{\omega : |l_n^{(i)}(\omega)-n\pi_i| > d(n \cdot \log n)^{1/2}\} = O(n^{-1/2})$$

where $d = \sigma_i \sqrt{\pi_i}$.

Proof. According to Theorem 1 we have

$$P\{\omega : |l_n^{(i)}(\omega)-n\pi_i| > d(n \cdot \log n)^{1/2}\} = P\left\{\omega : \left|\frac{l_n^{(i)}(\omega)-n\pi_i}{\sigma_i\sqrt{n\pi_i}}\right| > (\log n)^{1/2}\right\}$$

$$= 2(1-\Phi((\log n)^{1/2})) + O(n^{-1/2})$$

$$= O(n^{-1/2}).$$

Theorem 3. *Let X_n, $n \geqq 0$, be a positive recurrent, irreducible Markov chain, starting at $X_0 \equiv i$.*

Let φ be a real-valued function defined on the state space I. Assume that

(a) $|\varphi|_\nu^{(i)}, r_\nu^{(i)} \in L_3(P)$,

(b) $\sigma_i^2 = \operatorname{var}(\varphi_\nu^{(i)} - S(\varphi) r_\nu^{(i)}) > 0$.

Then

$$\sup_{t\in\mathbb{R}}\left|P\left\{\omega : \frac{1}{\sigma_i\sqrt{n\pi_i}}\left(\sum_{\nu=0}^{n}\varphi(X_\nu(\omega))-(n+1)S(\varphi)\right)<t\right\}-\Phi(t)\right|$$
$$= O(n^{-1/4}(\log n)^{1/4}).$$

Proof. In the following we omit the upper state index i, if no ambiguity can arise.

For each $n \geqq 0$ let $\rho_n(\omega)$ be the time of the first entrance of the state i after time n, i.e.

$$\rho_n(\omega) = \min\{\nu > n : X_\nu(\omega) = i\}.$$

Let $\psi_\nu = \varphi_\nu - S(\varphi) r_\nu$. Then $P[\psi_\nu] = 0$ (see e.g. [2], p. 96).

We have the following dissection formula

$$\sum_{\nu=0}^{n} \varphi \circ X_\nu - (n+1) S(\varphi) = Y_n + V_n + Z_n \tag{1}$$

where

$$Y_n(\omega)=\sum\{\psi_\nu(\omega)\colon 1\leqq \nu\leqq \mathrm{l}_n(\omega)\}, \tag{2}$$

$$V_n(\omega)=-\sum\{\varphi\circ X_\nu(\omega)\colon n<\nu<\rho_n(\omega)\}, \tag{3}$$

$$Z_n(\omega)=-S(\varphi)(\rho_n(\omega)-(n+1)). \tag{4}$$

Since $\psi_\nu\in L_3(P)$, $\nu\in\mathbb{N}$, are independent and identically distributed and $\sigma_i^2=\operatorname{var}\psi_\nu>0$ we obtain from Lemma 5 and Remark 2 that

$$\sup_{t\in\mathbb{R}}\left|P\left\{\omega\colon \frac{Y_n(\omega)}{\sigma_i\sqrt{n\pi_i}}\leqq t\right\}-\Phi(t)\right|=O(n^{-1/4}(\log n)^{1/4}).$$

According to (1), (5) and Lemma 6 it suffices to prove that for each $c>0$

$$P\left\{\omega\colon \frac{|V_n(\omega)|}{\sqrt{n}}>c\cdot\frac{(\log n)^{1/4}}{n^{1/4}}\right\}=O(n^{-1/4}(\log n)^{1/4}) \tag{6}$$

and

$$P\left\{\omega\colon \frac{|Z_n(\omega)|}{\sqrt{n}}>c\cdot\frac{(\log n)^{1/4}}{n^{1/4}}\right\}=O(n^{-1/4}(\log n)^{1/4}). \tag{7}$$

If we replace in (3) the function φ by the constant function $-S(\varphi)$ we see that the "new" $V_n(\omega)$ leads to $Z_n(\omega)$ of (4). Therefore it suffices to prove only (6).
According to Remark 2 there exists a constant $d>0$ such that

$$P\{\omega\colon |\mathrm{l}_n(\omega)-n\pi_i|\geqq d(n\cdot\log n)^{1/2}\}=O(n^{-1/2}). \tag{8}$$

Let $K=\{\nu\in\mathbb{N}\colon |\nu-n\pi_i|<d(n\cdot\log n)^{1/2}\}$. Since

$$|V_n(\omega)|\leqq\sum_{\nu=1}^{n+1}|\varphi|_\nu(\omega)\,1_{\{\omega\colon \mathrm{l}_n(\omega)=\nu\}}$$

we obtain, using (8), that

$$\begin{aligned}
&P\left\{\omega\colon \frac{|V_n(\omega)|}{\sqrt{n}}>c\frac{(\log n)^{1/4}}{n^{1/4}}\right\}\\
&\leqq P\{\omega\colon |\mathrm{l}_n(\omega)-n\pi_i|\geqq d(n\log n)^{1/2}\}\\
&\quad+P\{\omega\colon |V_n(\omega)|>c\,n^{1/4}(\log n)^{1/4};\ |\mathrm{l}_n(\omega)-n\pi_i|<d(n\log n)^{1/2}\}\\
&\leqq O(n^{-1/2})+\sum_{\nu\in K}P\{\omega\colon |\varphi|_\nu(\omega)>c\,n^{1/4}(\log n)^{1/4}\}\\
&\leqq O(n^{-1/2})+(2d(n\log n)^{1/2}+1)\,e\,n^{-3/4}(\log n)^{-3/4} \qquad (*)\\
&=O(n^{-1/2})+O(n^{-1/4})=O(n^{-1/4})
\end{aligned}$$

where $(*)$ follows from the Markov-inequality with some appropriate constant $e>0$. This proves (6) and hence the proof is finished.

Remark 4. The assertion of Theorem 3 also holds true for a general initial distribution of X_0, if we assume additionally that

$$\tau_1^{(i)} \in L_1(P) \quad \text{and} \quad \sum_{\nu=0}^{\tau_1^{(i)}-1} |\varphi \circ X_\nu| \in L_1(P).$$

Proof. Use the notations of Theorem 3. Then we have the following dissection formula

$$\sum_{\nu=0}^{n} \varphi \circ X_\nu - (n+1) S(\varphi) = R_n + Y_n + V_n + Z_n$$

where $|R_n| \leqq R = \sum_{\nu=0}^{\tau_1^{(i)}-1} |\varphi \circ X_\nu| \in L_1(P)$. Since by the Markov inequality

$$P\left\{\left|\frac{R_n}{\sqrt{n}}\right| > n^{-1/4}\right\} = O(n^{-1/4})$$

the assertion follows with Lemma 6 by applying the same methods as in the proof of Theorem 3.

For the sake of completeness we cite the following results:

Lemma 5. *Let X_n, $n \in \mathbb{N}$, be a sequence of independent and identically distributed random variables with $P(X_n) = 0$, $P(X_n^2) = 1$ and $P(|X_n|^3) < \infty$. Let $\mathfrak{l}_n: \Omega \to \mathbb{N}$ be $\mathscr{A}$-measurable and assume that for some constants $\alpha, d > 0$*

$$P\left\{\omega: \left|\frac{\mathfrak{l}_n(\omega)}{n} - \alpha\right| > d n^{-1/2} (\log n)^{1/2}\right\} = O(n^{-1/2}).$$

Then

$$A_n = \sup_{t \in \mathbb{R}} \left| P\left\{\omega: \frac{\sum_{\nu=1}^{\mathfrak{l}_n(\omega)} X_\nu(\omega)}{\sqrt{n\alpha}} \leqq t\right\} - \Phi(t)\right| = O(n^{-1/4} (\log n)^{1/4}).$$

Proof. See [4], Theorem 1.

A counterexample given in [4] shows that under the assumptions of Lemma 5 $A_n = O(n^{-1/4})$ does not hold true in general.

The result of Sreehari [8], applied under the assumptions of Lemma 5, leads only to $A_n = O(n^{-1/6})$.

The following Lemma is well known.

Lemma 6. *Let $Y_n, Z_n: \Omega \to \mathbb{R}$ be $\mathscr{A}$-measurable and $\varepsilon_n > 0$, $n \in \mathbb{N}$, be a sequence with $\varepsilon_n \to 0$. Assume that*

$$\sup_{t \in \mathbb{R}} |P(Y_n \leqq t) - \Phi(t)| = O(\varepsilon_n)$$

and

$$P(|Z_n| > \varepsilon_n) = O(\varepsilon_n).$$

Then

$$\sup_{t \in \mathbb{R}} |P(Y_n + Z_n \leqq t) - \Phi(t)| = O(\varepsilon_n).$$

References

1. Aleskevicus, G.J.: Some limit theorems for sums of random variables defined on a homogeneous regular Markov chain (In Russian). Litovsk. Mat. Sb. **6**, 297–311 (1966)
2. Chung, K.L.: Markov chains with stationary transition probabilities. New York: Springer 1967
3. Iosifescu, M., Theodorescu, R.: Random processes and learning. New York: Springer 1969
4. Landers, D., Rogge, L.: The exact approximation order in the central limit theorem for random summation. (Submitted for publication)
5. Philipp, W.: Mixing sequences of random variables and probabilistic number theory. Mem. Amer. Math. Soc. **114**, 1–102 (1971)
6. Prakasa Rao, B.L.S.: On the rate of convergence in the random central limit theorem for martingales. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **22**, 1255–1260 (1974)
7. Prakasa Rao, B.L.S.: Remark on the rate of convergence in the random central limit theorem for mixing sequences. Z. Wahrscheinlichkeitstheorie verw. Gebiete **31**, 157–160 (1975)
8. Sreehari, M.: An estimate of the residual in the central limit theorem for sums of random numbers of random variables. University of Baroda: Unpublished manuscript
9. Stein, C.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proc. 6th Berkeley Sympos. Math. Statist. Probab. Univ. Calif. **2**, 583–602 (1972)

Received June 10, 1975