Stochastic Processes and their Applications 3 (1975) 55–64

ON THE CONVERGENCE OF MOMENTS IN STATIONARY MARKOV CHAINS*

P.J. HOLEWIJN
Wiskundig Seminarium, Vrije Universiteit, Amsterdam, The Netherlands

and

A. HORDIJK
Mathematisch Centrum, Amsterdam, The Netherlands
Wiskundig Seminarium, Vrije Universiteit, Amsterdam, The Netherlands

Received 26 June 1973
Revised 2 May 1974

Necessary and sufficient conditions are given for the convergence of the first moment of functionals of Markov chains.

stationary Markov chain   convergence of moments

1. Introduction

Let $\{x(t)\}_{t=0}^{\infty}$ be an irreducible positive recurrent Markov chain with countable state space and with standard semigroup $\{P(t)\}_{t=0}^{\infty}$ of transi tion probabilities. If $\pi(j) := \lim_{t\to\infty} p_{ij}(t)$ and f is a functional of the chain, is it true that

$$\lim_{t\to\infty} \mathbf{E}\{f(x(t))\} = \sum_j \pi(j) f(j)? \tag{1.1}$$

Here **E** stands for expectation.

If $\Sigma_j \pi(j)\,|f(j)| < \infty$ and the chain is not started too badly, it is reasonable to suppose relation (1.1) to be correct. In applications it is common practice to take (1.1) for granted and to take for example for the

* Mathematical Centre Report SW 28/74.

mean "in the long run" the mean with respect to the stationary distribution $\{\pi(j)\}$. However, it turns out to be difficult to find an explicit discussion of the problem in the literature. As far as we know the only place where a proof of (1.1) (for chains with discrete time parameter) is presented is a paper by Kesten and Runnenburg [3, Theorem 4.3].

In this paper we give a very simple proof of (1.1) and consider some additional and related questions, yielding a simplified proof of a well-known theorem due to Chung (cf. [2, Theorem 3, p. 93]).

For the terminology we refer to Chung [2].

2. Results

2.1. Theorem. *Consider an irreducible positive recurrent Markov chain with one-step transition matrix* (p_{ij}), $i, j = 1, 2, \dots$. *Let*

$$\pi(j) := \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} p_{ij}^{(k)},$$

with $p_{ij}^{(k)}$ *the k-step transition probabilities. Let* $x_0 \equiv i_0$ *and let* $f \geqslant 0$ *be a functional of the chain with* $\Sigma_{j=1}\, \pi(j) f(j) < \infty$. *Then*

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \mathrm{E}\{f(x_k)\} = \sum_{j=1}^{\infty} \pi(j) f(j)$$

for each $i_0 = 1, 2, \dots$, *and if the chain is aperiodic,*

$$\lim_{n\to\infty} \mathrm{E}\{f(x_n)\} = \sum_{j=1}^{\infty} \pi(j) f(j)$$

for each $i_0 = 1, 2, \dots$.

Proof. Fix i_0. Since $\Sigma_{i=1}^{\infty}\, \pi(i) p_{ij}^{(n)} = \pi(j)$ for each $n = 1, 2, \dots$, we have

$$\pi(i_0) p_{i_0 j}^{(n)} \leqslant \pi(j),$$

and hence

$$p_{i_0 j}^{(n)} \leqslant \frac{\pi(j)}{\pi(i_0)}, \qquad n^{-1} \sum_{k=1}^{n} p_{i_0 j}^{(k)} \leqslant \frac{\pi(j)}{\pi(i_0)}.$$

Both assertions follow now by the bounded convergence theorem. □

2.2. Remarks

(1) Theorem 2.1 remains true if we start the chain with a distribution with finite support.

(2) If $f \geqslant 0$ and $\Sigma_{j=1}^{\infty} \pi(j) f(j) = \infty$, it follows from Fatou's lemma that the assertions are true if the chain is started with arbitrary initial distribution.

(3) Theorem 2.1 is also correct for arbitrary f, provided that

$$\sum_{j=1}^{\infty} \pi(j)\, |f(j)| < \infty .$$

(4) The proof of Theorem 2.1 applies verbatim in case the chain has continuous time parameter and is standard (in this case the phenomenon of periodicity does not occur).

In the next theorem we consider relation (1.1) in case the chain is null- or non-recurrent and obtain at the same time an alternative proof of the aperiodic part of Theorem 2.1. We write ${}_{i_0}p_{ij}^{(n)}$ for

$$\mathbf{P}[x_n = j,\ x_\nu \neq i_0,\ 1 \leqslant \nu < n \mid x_0 = i].$$

2.3. Theorem. *Let the Markov chain* $\{x_n\}_{n=0}^{\infty}$ *with discrete time parameter n be irreducible and null- or non-recurrent with* $x_0 \equiv i_0$. *Let* τ *stand for the return time to state* i_0, *i.e.*

$$\tau := \inf\{n > 0\colon x_n = i_0\} .$$

If $f \geqslant 0$ *is a functional of the chain with* $\mathbf{E}\{\Sigma_{n=0}^{\tau} f(x_n)\} < \infty$, *then*

$$\lim_{n\to\infty} \mathbf{E}\{f(x_n)\} = 0.$$

Proof. We apply the "last exit decomposition" of state i_0. We find (with $p_{i_0 i_0}^{(0)} := 1$):

$$p_{i_0 j}^{(n)} = \sum_{\nu=0}^{n-1} p_{i_0 i_0}^{(\nu)}\ {}_{i_0}p_{i_0 j}^{(n-\nu)} \qquad \text{for all } n \geqslant 1, j \geqslant 1.$$

It follows that

$$\sum_{j=1}^{\infty} p_{i_0 j}^{(n)} f(j) = \sum_{\nu=0}^{n-1} p_{i_0 i_0}^{(\nu)} \left(\sum_{j=1}^{\infty} {}_{i_0}p_{i_0 j}^{(n-\nu)} f(j) \right)$$

$$= \sum_{l=1}^{n} \left(\sum_{j=1}^{\infty} {}_{i_0}p_{i_0 j}^{(l)} f(j) \right) p_{i_0 i_0}^{(n-l)} . \tag{2.1}$$

We write

$$a_l := \sum_{j=1}^{\infty} {}_{i_0}p_{i_0 j}^{(l)} f(j), \qquad b_l := p_{i_0 i_0}^{(l)}, \qquad l \geqslant 1,$$

so that by hypothesis

$$\sum_{l=1}^{\infty} a_l = \sum_{l=1}^{\infty} \left(\sum_{j=1}^{\infty} {}_{i_0}p_{i_0 j}^{(l)} f(j) \right) < \infty . \tag{2.2}$$

According to (2.1) we have $\mathrm{E}\{f(x_n)\} < \infty$ for all $n \geqslant 1$. As $\lim_{n\to\infty} b_n = 0$, clearly

$$\lim_{n\to\infty} \mathrm{E}\{f(x_n)\} = \lim_{n\to\infty} \sum_{l=1}^{n} a_l b_{n-l} = 0. \square$$

2.4. Remark. If the conditions of Theorem 2.1 apply, we obtain

$$\sum_{l=1}^{\infty} a_l = \sum_{j=1}^{\infty} \frac{\pi(j)}{\pi(i_0)} f(j),$$

since then

$$\sum_{l=1}^{\infty} {}_{i_0}p_{i_0 j}^{(l)} = \frac{\pi(j)}{\pi(i_0)}$$

(cf. [2, p. 51]), and $\lim_{n\to\infty} b_n = \pi(i_0)$ if the chain is assumed to be aperiodic. We then find

$$\lim_{n\to\infty} \mathrm{E}\{f(x_n)\} = \sum_{j=1}^{\infty} \pi(j) f(j),$$

an alternative proof of the aperiodic part of Theorem 2.1.

In general, Theorem 2.1 fails to be true if the chain is started arbitrarily. This will be illustrated by the following example.

2.5. Example. Consider the Markov chain on the non-negative integers $\{0, 1, 2, \ldots\}$ with

$$\begin{aligned} p_{0j} &:= q^j p && \text{for } j = 0, 1, 2, \ldots, \\ p_{i,i-1} &:= 1 && \text{for } i = 1, 2, \ldots, \\ p_{ij} &:= 0 && \text{otherwise,} \end{aligned}$$

for $0 < p < 1$ and $q = 1-p$. A simple calculation shows that

$$\begin{aligned} p_{ij}^{(n)} &= q^j p && \text{for } j = 0, 1, 2, \ldots, i = 0, \ldots, n-1, \\ p_{n-1+k,k-1}^{(n)} &= 1, && \text{for } k = 1, 2, \ldots, \\ p_{ij}^{(n)} &= 0 && \text{otherwise.} \end{aligned}$$

This chain is irreducible, aperiodic and positive recurrent with

$$\lim_{n\to\infty} p_{ij}^{(n)} = q^j p, \qquad j = 0, 1, 2, \ldots .$$

If $\Sigma_{j=0}^{\infty} q^j\, |f(j)| < \infty$, we have with initial distribution $\{p(n)\}_{n=0}^{\infty}$ that

$$\sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} p(i)\, p_{ij}^{(n)}\right) f(j) = (p(0)+\ldots+p(n-1))\sum_{j=0}^{\infty} p\, q^j f(j) + \sum_{j=0}^{\infty} p(n+j) f(j),$$

and we have convergence to $\Sigma_{j=0}^{\infty} p q^j f(j)$ as $n \to \infty$ if and only if

$$\lim_{n\to\infty} \sum_{j=0}^{\infty} p(n+j) f(j) = 0.$$

If now

$$\begin{aligned} &f(j) := j^2 && \text{for } j \geqslant 0, \\ &p(0) := 0, && p(j) := 6\pi^{-2} j^{-2} \quad \text{for } j \geqslant 1, \end{aligned}$$

we have

$$\mathbf{E}\{f(\boldsymbol{x}_n)\} = \sum_{j=0}^{\infty} p(n+j) f(j) = \infty \quad \text{for all } n \geqslant 0.$$

In the next theorem we give necessary and sufficient conditions for the convergence of moments in a stationary chain.

2.6. Theorem. *Consider an irreducible, aperiodic, positive recurrent chain with initial distribution* $\{p(i)\}$ *and stationary distribution* $\{\pi(j)\}$. *Let* $f \geqslant 0$ *be a functional of the chain with* $\Sigma_j \, \pi(j) f(j) < \infty$. *Let*

$$p^0(j) := p(j), \qquad p^n(j) := \sum_i p(i)\, p_{ij}^{(n)}, \qquad n \geqslant 1.$$

Let τ *stand for the entrance time in state* i_0, *i.e.*

$$\tau := \inf\{n > 0 \colon x_n = i_0\}.$$

Then in order that

$$\lim_{n\to\infty} \sum_j p^n(j) f(j) = \sum_j \pi(j) f(j),$$

it is necessary and sufficient that

$$\lim_{n\to\infty} \mathrm{E}\{f(x_n)\, \chi_{\{\tau > n\}}\} = 0, \tag{2.3}$$

where χ *stands for the indicator function.*

Proof. With the "last exit decomposition" of state i_0 we find

$$\sum_j p^n(j) f(j) = \sum_{k=1}^{n-1} p^k(i_0) \Big(\sum_j {}_{i_0}p_{i_0 j}^{(n-k)} f(j)\Big) + \sum_i p(i) \Big(\sum_j {}_{i_0}p_{ij}^{(n)} f(j)\Big).$$

As in the proof of Theorem 2.3, it follows that the first term on the right-hand side tends to $\Sigma_j \, \pi(j) f(j)$ as $n \to \infty$. Since

$$\mathrm{E}\{f(x_n)\, \chi_{\{\tau > n\}}\} = \sum_i p(i) \Big(\sum_j {}_{i_0}p_{ij}^{(n)} f(j)\Big),$$

the assertion follows. □

2.7. Remark. One easily verifies that condition (2.3) is satisfied if a constant c and an $n_0 \geqslant 0$ exist such that

$$p^{n_0}(i)/\pi(i) \leqslant c \qquad \text{for all } i.$$

For, with $n > n_0$,

$$\sum_i p(i)\left(\sum_j {}_{i_0}p_{ij}^{(n)} f(j)\right) = \sum_{i \neq i_0} p^{n_0}(i)\left(\sum_j {}_{i_0}p_{ij}^{(n-n_0)} f(j)\right)$$

$$\leqslant c \sum_{i \neq i_0} \pi(i) \left(\sum_j {}_{i_0}p_{ij}^{(n-n_0)} f(j)\right)$$

$$= c \sum_{i \neq i_0} \pi(i_0)\left(\sum_{l=1}^{\infty} {}_{i_0}p_{i_0 i}^{(l)}\right) \left(\sum_j {}_{i_0}p_{ij}^{(n-n_0)} f(j)\right)$$

$$= c\,\pi(i_0) \sum_{l=n-n_0+1}^{\infty} \left(\sum_j {}_{i_0}p_{i_0 j}^{(l)} f(j)\right), \qquad (2.4)$$

since

$$\frac{\pi(i)}{\pi(i_0)} = \sum_{l=1}^{\infty} {}_{i_0}p_{i_0 i}^{(l)} .$$

The right-hand term in (2.4) tends to zero as $n \to \infty$. It should be noted, however, that this result is also immediate from the bounded convergence theorem. For $p^{n_0}(j) \leqslant c\,\pi(j)$ for all j implies by iteration that $p^n(j) \leqslant c\,\pi(j)$ for all j and $n \geqslant n_0$, and hence

$$\lim_{n\to\infty} \sum_j p^n(j) f(j) = \sum_j \pi(j) f(j).$$

As an application we now show that a well-known theorem due to Chung [2, Theorem 3, p. 93]) is an immediate consequence of Theorem 2.1 and a strong law of large numbers. To be complete, we first state this strong law in the next lemma.

2.7. Lemma (cf. [2, Theorem 2, p. 92]). *Let the chain* $\{x_n\}_{n=0}^{\infty}$ *with* $x_0 \equiv i_0$ *be irreducible and positive recurrent with stationary distribution* $\{\pi(j)\}$, *so that*

$$\pi(j) := \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} p_{i_0 j}^{(k)} ,$$

and let f be a functional of the chain with $\Sigma_j \pi(j)\, |f(j)| < \infty$. *Then*

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} f(x_i) = \sum_j \pi(j) f(j) \qquad a.s. \qquad (2.5)$$

Proof. We may restrict to the case $f \geqslant 0$. Define $\tau_0 \equiv 0$ and

$$\tau_n := \inf\{k\colon x_k = i_0 \text{ for the } n^{\text{th}} \text{ time}\}, \quad n \geqslant 1,$$

$$\tau_{l(n)} \leqslant n < \tau_{l(n)+1}, \qquad n \geqslant 0,$$

i.e. $\tau_{l(n)}$ is the time of last visit to state i_0 before or at time n $(n \geqslant 0)$ and $l(n)$ is the number of visits to i_0 after time 0 and before or at time n. We now have

$$\sum_{k=0}^{l(n)-1} \sum_{\tau_k}^{\tau_{k+1}-1} f(x_i) \leqslant \sum_{i=1}^{n} f(x_i) \leqslant \sum_{k=0}^{l(n)} \sum_{\tau_k}^{\tau_{k+1}-1} f(x_i), \qquad n \geqslant 1. \quad (2.6)$$

The random variables

$$y_k := \sum_{\tau_k}^{\tau_{k+1}-1} f(x_i), \qquad k \geqslant 0,$$

are independent and identically distributed and have finite mean

$$\mathrm{E}\{y_1\} = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j),$$

since

$$\mathrm{E}\{y_1\} = \int \left(\sum_{i=0}^{\tau_1-1} f(x_i)\right) \mathrm{d}P = \sum_{k=1}^{\infty} \int_{\{\tau_1=k\}} \left(\sum_{i=0}^{k-1} f(x_i)\right) \mathrm{d}P$$

$$= f(i_0) + \sum_{k=1}^{\infty} \sum_{i=1}^{k-1} \sum_{j \neq i_0} {}_{i_0}p_{i_0 j}^{(i)} \, {}_{i_0}p_{j i_0}^{(k-i)} f(j)$$

$$= f(i_0) + \sum_{j \neq i_0} \sum_{i=1}^{\infty} \sum_{k=i+1}^{\infty} {}_{i_0}p_{i_0 j}^{(i)} \, {}_{i_0}p_{j i_0}^{(k-i)} f(j)$$

$$= \sum_j \sum_{i=1}^{\infty} {}_{i_0}p_{i_0 j}^{(i)} f(j) = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j),$$

since

$$\sum_{l=1}^{\infty} {}_{i_0}p_{j i_0}^{(l)} = 1, \qquad \sum_{i=1}^{\infty} {}_{i_0}p_{i_0 j}^{(i)} = \frac{\pi(j)}{\pi(i_0)}.$$

We further have (recurrence) that $l(n) \to \infty$ $(n \to \infty)$ a.s., so that (Kolmogorov's law of large numbers)

$$\lim_{n\to\infty} \frac{1}{l(n)} \sum_{k=0}^{l(n)} y_k = \lim_{n\to\infty} \frac{1}{l(n)} \sum_{k=0}^{l(n)-1} y_k \quad \text{a.s.,}$$

$$= \mathrm{E}\{y_1\} \quad \text{a.s.}$$

and with (2.6) we conclude

$$\lim_{n\to\infty} \frac{1}{l(n)} \sum_{i=1}^{n} f(x_i) = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j) \quad \text{a.s.}$$

If we choose, in particular, $f \equiv 1$ it follows that

$$\frac{n}{l(n)} \to \frac{1}{\pi(i_0)} \quad \text{a.s.} \quad \text{as } n \to \infty,$$

which proves (2.5). □

2.8. Theorem. *Let the chain* $\{x_n\}_{n=0}^{\infty}$ *with* $x_0 \equiv i_0$ *be irreducible and positive recurrent with stationary distribution* $\{\pi(j)\}$, *and let f be a functional of the chain with* $\Sigma_j\, \pi(j)\, |f(j)| < \infty$. *Then*

$$\lim_{n\to\infty} \left| \mathrm{E}\; n^{-1} \sum_{i=1}^{n} f(x_i) - \sum_j \pi(j) f(j) \right| = 0.$$

Proof. From Theorem 2.1 we know that

$$\lim_{n\to\infty} \mathrm{E} \left\{ n^{-1} \sum_{i=1}^{n} |f(x_i)| \right\} = \sum_j \pi(j)\, |f(j)|. \tag{2.7}$$

(2.7) and (2.5) imply that the sequence $\{n^{-1} \Sigma_{i=1}^{n} |f(x_i)|\}$, and a fortiori the sequence $\{n^{-1} \Sigma_{i=1}^{n} f(x_i)\}$, is uniformly integrable. This fact and (2.5) finally imply the assertion (cf. [1, pp. 91, 94]). □

Acknowledgment

The authors are indebted to Professor J.Th. Runnenberg (University of Amsterdam) for valuable comments and for the simple Example 2.5. Our original example was more complicated.

References

[1] L. Breiman, Probability (Addison-Wesley, Reading Mass., 1968).
[2] K.L. Chung, Markov Chains with Stationary Transition Probabilities (Springer, Berlin, 1960).
[3] H. Kesten and J.Th. Runnenburg, Priority in waiting line problems I, Proc. Koninkl. Nederl. Akad. Wetensch. (A) 60 (1957) 312–324.