

**stichting
mathematisch
centrum**



AFDELING MATHEMATISCHE STATISTIEK

SW 28/74

MARCH

P.J. HOLEWIJN & A. HORDIJK
ON THE CONVERGENCE OF MOMENTS
IN STATIONARY MARKOV CHAINS

Prepublication

BIBLIOTHEEK MATHEMATISCH CENTRUM
AMSTERDAM

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

On the convergence of moments in stationary Markov chains^{†)}

P.J. Holeyijn^{*)} and A. Hordijk^{**)}

Abstract

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

1. Introduction

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

$$(1) \quad \lim_{t \rightarrow \infty} E f(\underline{x}(t)) = \sum_j \pi(j) f(j) ?$$

Here E stands for expectation.

If $\sum_j \pi(j) |f(j)| < \infty$ and if the chain is not started too badly, it is reasonable to suppose relation (1) to be correct. In applications it is common practice to take (1) for granted and to take for example for the 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

^{†)} This paper is not for review; it is meant for publication in a journal.

^{*)} Free University at Amsterdam, Netherlands.

^{**)} Mathematical Centre, Amsterdam and Free University at Amsterdam, Netherlands.

only place where a proof of (1) (for chains with discrete time-parameter) is presented is a paper by H. Kesten and J.Th. Runnenburg: *Priority in Waiting Line Problems I* (theorem 4.3) that appeared in the Proc. Kon. Nederl. Akad. Wetensch. in 1957 ([3]).

In this paper we give a very simple proof of (1) and consider some additional and related questions, yielding a simplified proof of a well-known theorem due to Chung (cf. [1], theorem 3, p.93).

For terminology we refer to Chung ([1]).

2.

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

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

with $p_{ij}^{(k)}$ the k -step transition probabilities. Let $\underline{x}_0 \equiv i_0$ and $f \geq 0$ be a functional of the chain with $\sum_{j=1}^{\infty} \pi(j)f(j) < \infty$. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n E f(\underline{x}_k) = \sum_{j=1}^{\infty} \pi(j)f(j)$$

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

$$\lim_{n \rightarrow \infty} E f(\underline{x}_n) = \sum_{j=1}^{\infty} \pi(j)f(j)$$

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

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

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

and hence

$$p_{i_0 j}^{(n)} \leq \frac{\pi(j)}{\pi(i_0)}$$

and

$$\frac{1}{n} \sum_{k=1}^n p_{i_0 j}^{(k)} \leq \frac{\pi(j)}{\pi(i_0)}.$$

Both assertions follow now by the bounded convergence theorem. $\square$

Remarks

- 1) Theorem 1 remains true if we start the chain with a distribution with finite support.
- 2) If $f \geq 0$ and $\sum_{j=1}^{\infty} \pi(j) f(j) = \infty$ it follows from Fatou's lemma that the assertions are true if the chain is started with arbitrary initial distribution.
- 3) Theorem 1 is also correct for arbitrary f provided that $\sum_{j=1}^{\infty} \pi(j) |f(j)| < \infty$.
- 4) The proof of theorem 1 applies verbally 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) in case the chain is null- or non-recurrent and obtain at the same time an alternative proof of the aperiodic part of theorem 1. We write $p_{i_0 j}^{(n)}$ for $P(\underline{x}_n = j, \underline{x}_v \neq i_0, 1 \leq v < n \mid \underline{x}_0 = i)$.

Theorem 2. Let the Markov chain $\{\underline{x}_n\}_{n=0}^{\infty}$ with discrete time parameter n be irreducible and null- or non-recurrent with $\underline{x}_0 \equiv i_0$. Let τ stand for the return time to state i_0 , i.e. $\tau := \inf \{n > 0: \underline{x}_n = i_0\}$. If $f \geq 0$ is a functional of the chain with $E(\sum_{n=0}^{\tau} f(\underline{x}_n)) < \infty$, then

$$\lim_{n \rightarrow \infty} E f(\underline{x}_n) = 0.$$

Proof. We apply the "last exit decomposition" of state i_0 . We find

(with $p_{i_0 i_0}^{(0)} := 1$):

$$p_{i_0 j}^{(n)} = \sum_{v=0}^{n-1} p_{i_0 i_0}^{(v)} p_{i_0 j}^{(n-v)}, \text{ for all } n \geq 1 \text{ and } j \geq 1.$$

It follows that

$$\begin{aligned} \sum_{j=1}^{\infty} p_{i_0 j}^{(n)} f(j) &= \sum_{v=0}^{n-1} p_{i_0 i_0}^{(v)} \left(\sum_{j=1}^{\infty} p_{i_0 j}^{(n-v)} f(j) \right) = \\ (2) \quad &= \sum_{l=1}^n \left(\sum_{j=1}^{\infty} p_{i_0 j}^{(l)} f(j) \right) p_{i_0 i_0}^{(n-l)}. \end{aligned}$$

We write $a_l := \sum_{j=1}^{\infty} p_{i_0 j}^{(l)} f(j)$ and $b_l := p_{i_0 i_0}^{(l)}$, $l \geq 1$, so that by hypothesis

$$(3) \quad \sum_{l=1}^{\infty} a_l = \sum_{l=1}^{\infty} \left(\sum_{j=1}^{\infty} p_{i_0 j}^{(l)} f(j) \right) < \infty.$$

According to (2) we have $E f(\underline{x}_n) < \infty$ for all $n \geq 1$. As $\lim_{n \rightarrow \infty} b_n = 0$, clearly

$$\lim_{n \rightarrow \infty} E f(\underline{x}_n) = \lim_{n \rightarrow \infty} \sum_{l=1}^n a_l b_{n-l} = 0. \quad \square$$

Remark. If the conditions of theorem 1 apply we obtain $\sum_{l=1}^{\infty} a_l =$
 $= \sum_{j=1}^{\infty} \frac{\pi(j)}{\pi(i_0)} f(j)$, since then $\sum_{l=1}^{\infty} i_0 p_{i_0 j}^{(1)} = \frac{\pi(j)}{\pi(i_0)}$ (cf. [1], p.51), and
 $\lim_{n \rightarrow \infty} b_n = \pi(i_0)$, if the chain is assumed to be aperiodic. We then find
 $\lim_{n \rightarrow \infty} E f(\underline{x}_n) = \sum_{j=1}^{\infty} \pi(j) f(j)$, an alternative proof of the aperiodic part
of theorem 1.

In general theorem 1 fails to be true if the chain is started arbitrarily. This will be illustrated by the following example.

Consider the Markov chain on the nonnegative integers $\{0, 1, 2, \dots\}$
with $p_{0j} := q^j p$ for $j=0, 1, 2, \dots$ and, $p_{i, i-1} := 1$ for $i=1, 2, \dots$ and $p_{ij} := 0$
otherwise, for $0 < p < 1$ and $q = 1-p$. A simple calculation shows that
 $p_{ij}^{(n)} = q^j p$ for $j=0, 1, 2, \dots$, $i=0, \dots, n-1$, $p_{n-1+k, k-1}^{(n)} = 1$ for $k=1, 2, \dots$,
and $p_{ij}^{(n)} = 0$ otherwise. This chain is irreducible, aperiodic and positive
recurrent with $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = q^j p$, $j=0, 1, 2, \dots$.

If $\sum_{j=0}^{\infty} q^j |f(j)| < \infty$ we have with initial distribution $\{p(n)\}_{n=0}^{\infty}$
that $\sum_{j=0}^{\infty} (\sum_{i=0}^{\infty} p(i) p_{ij}^{(n)}) f(j) = (p(0) + \dots + p(n-1)) \sum_{j=0}^{\infty} p q^j f(j) +$
 $+ \sum_{j=0}^{\infty} p(n+j) f(j)$ and we have convergence to $\sum_{j=0}^{\infty} p q^j f(j)$ as $n \rightarrow \infty$ if and
only if $\lim_{n \rightarrow \infty} \sum_{j=0}^{\infty} p(n+j) f(j) = 0$. If now $f(j) := j^2$ for $j \geq 0$,
 $p(j) := \frac{6}{\pi^2} \frac{1}{j^2}$ for $j \geq 1$ and $p(0) := 0$, we have $E f(\underline{x}_n) = \sum_{j=0}^{\infty} p(n+j) f(j) = \infty$
for all $n \geq 0$.

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

Theorem 3. Consider an irreducible, aperiodic, positive recurrent chain
with initial distribution $\{p(i)\}$ and stationary distribution $\{\pi(j)\}$. Let

$f \geq 0$ be a functional of the chain with $\sum_j \pi(j)f(j) < \infty$. Let $p^n(j) := \sum_i p(i)p_{ij}^{(n)}$, $n \geq 1$ (and $p^0(j) := p(j)$). Let τ stand for the entrance time in state i_0 , i.e. $\tau := \inf \{n > 0: \underline{x}_n = i_0\}$. Then in order that

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

it is necessary and sufficient that

$$(4) \quad \lim_{n \rightarrow \infty} E(f(\underline{x}_n) \chi_{\{\tau > n\}}) = 0,$$

where χ stands for the indicatorfunction.

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

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

As in the proof of theorem 2 it follows that the first term on the right-hand side tends to $\sum_j \pi(j)f(j)$ as $n \rightarrow \infty$. Since

$$E(f(\underline{x}_n) \chi_{\{\tau > n\}}) = \sum_i p(i) \left(\sum_j p_{i_0 ij}^{(n)} f(j) \right),$$

the assertion follows. $\square$

Remark. One easily verifies that condition (4) is satisfied if a constant c and an $n_0 \geq 0$ exist such that

$$\frac{p_{i_0(i)}^{n_0}}{\pi(i)} \leq c \quad \text{for all } i.$$

For with $n > n_0$

$$(5) \quad \sum_i p(i) \left(\sum_j p_{i_0 ij}^{(n)} f(j) \right) = \sum_{i \neq i_0} p_{i_0(i)}^{n_0} \left(\sum_j p_{i_0 ij}^{(n-n_0)} f(j) \right) \leq$$

$$\begin{aligned}
&\leq c \sum_{i \neq i_0} \pi(i) \left(\sum_j p_{i_0 i j}^{(n-n_0)} f(j) \right) = \\
&= c \sum_{i \neq i_0} \pi(i_0) \left(\sum_{l=1}^{\infty} p_{i_0 i_0 i}^{(l)} \right) \left(\sum_j p_{i_0 i j}^{(n-n_0)} f(j) \right) = \\
&= c \pi(i_0) \sum_{l=n-n_0+1}^{\infty} \left(\sum_j p_{i_0 i_0 j}^{(l)} f(j) \right),
\end{aligned}$$

since

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

The right-hand term in (5) tends to zero as $n \rightarrow \infty$. It should be noted, however, that this result is also immediate from the bounded convergence theorem. For, $p^{n_0}(j) \leq c \pi(j)$ for all j , implies by iteration that $p^n(j) \leq c \pi(j)$ for all j and $n \geq n_0$ and hence

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

As an application we now show that a well-known theorem due to Chung ([1], theorem 3, p.93) is an immediate consequence of theorem 1 and a strong law of large numbers. To be complete we first state this strong law in the next lemma.

Lemma (cf. [1], theorem 2, p.92). Let the chain $\{\underline{x}_n\}_{n=0}^{\infty}$ with $\underline{x}_0 \equiv i_0$ be irreducible and positive-recurrent with stationary distribution $\{\pi(j)\}$ ($\pi(j) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{i_0 j}^{(k)}$) and let f be a functional of the chain with $\sum_j \pi(j) |f(j)| < \infty$. Then

$$(6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(\underline{x}_i) = \sum_j \pi(j) f(j) \text{ a.s.}$$

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

$$\tau_n := \inf \{k : \underline{x}_k = i_0 \text{ for the } n\text{-th time}\}, \quad n \geq 1,$$

$$\tau_{\underline{l}(n)} \leq n < \tau_{\underline{l}(n)+1}, \quad n \geq 0,$$

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

$$(7) \quad \begin{aligned} \sum_{k=0}^{\underline{l}(n)-1} \sum_{\tau_k}^{\tau_{k+1}-1} f(\underline{x}_i) &\leq \sum_{i=1}^n f(\underline{x}_i) \leq \\ &\leq \sum_{k=0}^{\underline{l}(n)} \sum_{\tau_k}^{\tau_{k+1}-1} f(\underline{x}_i), \quad n \geq 1. \end{aligned}$$

The random variables

$$Y_k := \sum_{\tau_k}^{\tau_{k+1}-1} f(\underline{x}_i), \quad k \geq 0,$$

are independent and identically distributed and have finite mean

$$E(Y_1) = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j),$$

since

$$\begin{aligned} E(Y_1) &= \int \left(\sum_{i=0}^{\tau_1-1} f(\underline{x}_i) \right) dP = \sum_{k=1}^{\infty} \int_{\{\tau_1=k\}} \left(\sum_{i=0}^{k-1} f(\underline{x}_i) \right) dP = \\ &= f(i_0) + \sum_{k=1}^{\infty} \sum_{i=1}^{k-1} \sum_{j \neq i_0} p_{i_0 i_0}^{(i)} p_{i_0 j}^{(k-i)} f(j) = \end{aligned}$$

$$\begin{aligned}
&= f(i_0) + \sum_{j \neq i_0} \sum_{i=1}^{\infty} \sum_{k=i+1}^{\infty} p_{i_0 i_0 j}^{(i)} p_{i_0 j i_0}^{(k-i)} f(j) = \\
&= \sum_j \sum_{i=1}^{\infty} p_{i_0 i_0 j}^{(i)} f(j) = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j),
\end{aligned}$$

since

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

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

$$\lim_{n \rightarrow \infty} \frac{1}{\underline{l}(n)} \sum_{k=0}^{\underline{l}(n)} y_k \stackrel{\text{a.s.}}{=} \lim_{n \rightarrow \infty} \frac{1}{\underline{l}(n)} \sum_{k=0}^{\underline{l}(n)-1} y_k \stackrel{\text{a.s.}}{=} E(y_1),$$

and with (7) we conclude

$$\lim_{n \rightarrow \infty} \frac{1}{\underline{l}(n)} \sum_{i=1}^n f(\underline{x}_i) = \sum_j \frac{\pi(j)}{\pi(i_0)} f(j) \text{ a.s.}$$

If we choose in particular $f \equiv 1$ it follows that $\frac{n}{\underline{l}(n)} \rightarrow \frac{1}{\pi(i_0)}$ ($n \rightarrow \infty$) a.s., which proves that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(\underline{x}_i) = \sum_j \pi(j) f(j) \text{ a.s. } \square$$

Theorem 4. Let the chain $\{\underline{x}_n\}_{n=0}^{\infty}$ with $\underline{x}_0 \equiv i_0$ be irreducible and positive-recurrent with stationary distribution $\{\pi(j)\}$ and let f be a functional of the chain with $\sum_j \pi(j) |f(j)| < \infty$. Then

$$\lim_{n \rightarrow \infty} E \left| \frac{1}{n} \sum_{i=1}^n f(\underline{x}_i) - \sum_j \pi(j) f(j) \right| = 0.$$

Proof. From theorem 1 we know that

$$(8) \quad \lim_{n \rightarrow \infty} E \left(\frac{1}{n} \sum_{i=1}^n |f(\underline{x}_i)| \right) = \sum_j \pi(j) |f(j)|.$$

(8) and (6) imply that the sequence $\{\frac{1}{n} \sum_{i=1}^n |f(\underline{x}_i)|\}$, and *a fortiori* the sequence $\{\frac{1}{n} \sum_{i=1}^n f(\underline{x}_i)\}$, is uniformly integrable. This last fact and (6) finally imply the assertion (cf. [2], pp.91 and 94). $\square$

Acknowledgement

The authors are indebted to professor J.Th. Runnenburg (University of Amsterdam) for valuable comments and for the simple example after theorem 2. Our original example was more complicated.

References

- [1] Chung, K.L., *Markov chains with stationary transition probabilities*, Springer-Verlag, Berlin, 1960.
- [2] Breiman, L., *Probability*, Addison-Wesley Publ. Coy., 1968.
- [3] Kesten, H. & Runnenburg, J.Th., *Priority in waiting line problems I*, Proc.Kon.Nederl.Akad.Wetensch. series A, 60 (1957) 312-324.