PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 9, Fasc. 2 (1988), p. 77–83

ON THE CONVERGENCE OF WEIGHTED AVERAGES OF RANDOM VARIABLES ARISING FROM A FINITE MARKOV CHAIN

BY

DAVID B. WOLFSON (MONTRÉAL)

Abstract. Jamison et al. [6] discussed the convergence of weighted sums of independent random variables to a degenerate random variable. In this note one of their results is extended (with the same condition on the weights) to the sequence of holding times of a Markov Renewal process. A similar growth condition on the weights ensures the convergence of these weighted sums (suitably normalized) to the Normal law.

1. Introduction. Let $\{(J_n, X_n), n = 1, 2, \ldots\}$ be a two-dimensional Markov process defined as follows:

(1.1) $X_0 = 0$ a.s.

(1.2) $\{J_n\}_{n=1}^{\infty}$ forms an m-state ($m < \infty$) ergodic Markov chain with initial distribution vector, $\boldsymbol{\alpha}$, say.

(1.3) $$\mathrm{P}(X_n \leqslant x, J_n = j \mid X_0, J_0, \ldots, X_{n-1}, J_{n-1}) = \mathrm{P}(X_n \leqslant x, J_n = j \mid J_{n-1}) = P_{J_{n-1},j} H_j(x) \text{ a.s.,}$$

where $P_{ij} = \mathrm{P}(J_n = j \mid J_{n-1} = i)$ and $H_j(\cdot)$ is a proper distribution function.

Such a process, called a *J–X process*, has been extensively studied in the context of Markov renewal processes and semi-Markov processes. The book [2], for instance, contains most of the basic properties of these processes.

The random variables $\{X_n\}$ are *conditionally independent* when $m > 1$, and *independent* when $m = 1$. It is shown in this note that the obvious modifications of the sufficient conditions for the convergence of weighted sums of independent random variables yield convergence in this more general setup. Jamison et al. [6] discuss the situation where X_n's are independent.

2. A sufficient condition for convergence in probability.

THEOREM 2.1. *Let $\{(J_n, X_n), n = 0, 1, \ldots\}$ be an m-state ($m < \infty$) ergodic J–X process. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of positive weights such that $\max_{1 \leqslant k \leqslant n} a_k/A_n \to 0$ as $n \to \infty$, where $A_n = \sum_{k=1}^{n} a_k$. If*

$$\lim_{T \to \infty} TP[|X_j| \geqslant T | J_{j-1} = i] = 0,$$
$$\lim_{T \to \infty} \int_{|x| < T} x\, dP_\pi(X_j \leqslant x | J_{j-1} = i) = \mu_i < \infty \tag{2.1}$$

for each i, then

$$A_n^{-1} \sum_{k=1}^{n} a_k X_k \xrightarrow{P} \sum_{i=1}^{m} \pi_i \mu_i.$$

Here $P_\pi(\cdot)$ denotes the conditional probability on the J–X process with $\alpha = \pi$.

Before proceeding directly with the proof we shall need two results. First, we assume that the initial distribution of the underlying Markov chain is π. With this assumption it is easily seen (see e.g. [1]) that the sequence $\{X_k\}$ is stationary.

Next we provide the following lemma, which is a simple extension of a similar result established by O'Brien [1].

LEMMA 2.1. *Let α and β be two initial distributions for the underlying Markov chain $\{J_n\}$. Let $\max_{1 \leqslant k \leqslant n} a_k/A_n \to 0$ as $n \to \infty$ and suppose that $P_\alpha[A_n^{-1}(a_1 X_1 + \ldots + a_n X_n) + B_n \leqslant x]$ converges weakly to a proper distribution $F(x)$. Then $P_\beta[A_n^{-1}(a_1 X_1 + \ldots + a_n X_n) + B_n \leqslant x]$ also converges weakly to $F(x)$.*

Proof. Let $S_n^N = a_{N+1} X_{N+1} + \ldots + a_n X_n$ and $S_n = S_n^0$. Let N be sufficiently large that $\|\beta P^N - \alpha P^N\| < \varepsilon$. Let A be any event which depends on $a_{N+1} X_{N+1}, a_{N+2} X_{N+2}, \ldots$ Then

$$|P_\beta(A) - P_\alpha(A)| \leqslant \sum_i P(A | J_N = i) |P_\beta(J_N = i) - P_\alpha(J_N = i)|$$
$$\leqslant \sum_i |(\beta P^N)_i - (\alpha P^N)_i| = \|\beta P^N - \alpha P^N\| < \varepsilon.$$

Let x be any continuity point of F. Let $\delta > 0$ be sufficiently small so that $F(x+\delta) - F(x-2\delta) < \varepsilon$. Let $n > N$ be sufficiently large that $P_\beta(|A_n^{-1} S_N| > \delta) < \varepsilon$ and $P_\alpha(|A_n^{-1} S_N| > \delta) < \varepsilon$. This is possible since $A_n \to \infty$. One then proceeds step by step, in exactly the same way as O'Brien to conclude the result.

Proof of Theorem 2.1. We shall establish the result with $\alpha = \pi$ and then use Lemma 2.1 to complete the proof for general α.

Let

$$X_{nk} = \begin{cases} X_k & \text{if } |X_k| \leqslant A_n/a_k, \\ 0 & \text{otherwise.} \end{cases}$$

Let $S_n = \sum_{l=1}^{n} a_k X_k$ and $S_{nn} = \sum_{k=1}^{n} a_k X_{nk}$. Then, for all n sufficiently large, since $\max_{1 \leqslant k \leqslant n} a_k/A_n \to 0$, we have, for any $\varepsilon > 0$,

$$\mathrm{P}(S_{nn} \neq S_n) \leqslant \sum_{k=1}^{n} P(X_{nk} \neq X_k) = \sum_{k=1}^{n} P\left[|X_1| > \frac{A_n}{a_n}\right]$$

$$= \sum_{k=1}^{n} \sum_{k=1}^{m} P\left[|X_1| > \frac{A_n}{a_n} \middle| J_0 = i\right] P_\pi(J_0 = i) < \varepsilon,$$

using assumption (2.1).

Hence, since the two sequences of random variables are equivalent, we consider S_{nn} instead of S_n.

Now,

$$\mathrm{E}(A_n^{-1} S_{nn}) = A_n^{-1} \sum_{i=1}^{m} \sum_{k=1}^{n} a_k \int_{|x| < A_n/a_k} x dP(X_1 \leqslant x \mid J_0 = i) P(J_0 = i)$$

$$\to \sum_{i=1}^{m} \pi_i \mu_i \text{ as } n \to \infty,$$

by using the second hypothesis of assumption (2.1).

The proof will be complete by showing that, for n sufficiently large, $\mathrm{Var}(S_{nn}/A_n) < \varepsilon$. First, we have

$$T^{-1} \int_{|x| < T} x^2 dP(X_1 \leqslant x) \tag{2.2}$$

$$= T^{-1} \{-T^2 P(|X_1| \geqslant T) + 2 \int_{0 \leqslant x < T} x dP(|X_1| \geqslant x)\} \to 0 \quad \text{as } T \to \infty.$$

Now,

$$\mathrm{Var}(S_{nn}/A_n) = A_n^{-2} \sum_{k=1}^{n} a_k^2 \mathrm{Var}(X_{nk}) + A_n^{-2} \sum_{k \neq j} a_j a_k \mathrm{Cov}(X_{nj}, X_{nk}).$$

Call the first term on the right L_n and the second term M_n. Then

$$L_n \leqslant A_n^{-2} \sum_{k=1}^{n} a_k^2 \int_{|x| < A_n/a_k} x^2 dP(X_1 \leqslant x)$$

$$= A_n^{-1} \sum_{k=1}^{n} a_k \frac{a_k}{A_n} \int_{|x| < A_n/a_k} x dP(X_1 \leqslant x) < \varepsilon$$

for n large enough, using (2.2).

On the other hand,

$$|M_n| = \Big|A_n^{-2} \sum_{k \neq j} a_j a_k [\mathrm{E}(X_{nj} X_{nk}) - \mathrm{E}(X_{nj}) \mathrm{E}(X_{nk})]\Big|$$

$$= \Big|A_n^{-2} \sum_{k \neq j} a_j a_k \sum_{r,s=1}^{m} \mathrm{E}(X_{nj} X_{nk} \mid J_{j-1} = r, J_{k-1} = s) P(J_{j-1} = r, J_{k-1} = s) -$$

$$- \mathrm{E}(X_{nj} \mid J_{j-1} = r) \mathrm{E}(X_{nk} \mid J_{k-1} = s) \mathrm{P}(J_{j-1} = r) \mathrm{P}(J_{k-1} = s)\Big|$$

$$\leqslant \sum_{r,s=1}^{m} A_n^{-2} \sum_{k \neq j} |\mathrm{E}(X_{nj} X_{nk} \mid J_{j-1} = r, J_{k-1} = s| \, |\mathrm{P}(J_{j-1} = r, J_{k-1} = s)$$

$$- \mathrm{P}(J_{j-1} = r) \mathrm{P}(J_{k-1} = s)| \, a_j a_k$$

$$\leqslant 2 \sum_{r,s=1}^{m} A_n^{-2} \sum_{k<j} |\mathrm{E}(X_{nj} X_{nk} \mid J_{j-1} = r, J_{k-1} = s| \, |\mathrm{P}(J_{k-1} = s \mid J_{j-1} = r)$$

$$- \mathrm{P}(J_{k-1} = s)| \, a_j a_k.$$

Consider, first

$$A_n^{-2} \sum_{j<k} a_j a_k |\mathrm{E}(X_{nj} X_{nk} \mid J_{j-1} = r, J_{k-1} = s)| \, |\mathrm{P}(J_{k-1} = s) - \pi_s| = 0,$$

since $\mathrm{P}(J_{k-1} = s) = \pi_s$.

All that remains

$$= A_n^{-2} \sum_{j<k} a_j a_k |\mathrm{E}(X_{nj} \mid J_{j-1} = r) \mathrm{E}(X_{nk} \mid J_{k-1} = s)| \, |\mathrm{P}(J_{k-1} = s \mid J_{j-1} = r) - \pi_s|$$

$$= A_n^{-2} \sum_{j<k} a_j a_k |\mathrm{E}(X_{nj} \mid J_{j-1} = r) \mathrm{E}(X_{nk} \mid J_{k-1} = s)| \, |\mathrm{P}(J_{k-j} = s \mid J_0 = r) - \pi_s|$$

$$= A_n^{-2} \sum_{k=1}^{n} |\mathrm{P}(J_k = s \mid J_0 = r) - \pi_s| \sum_{l=1}^{n-k} a_l a_{l+k} |\mathrm{E}(X_{nl} \mid J_{l-1} = r) \mathrm{E}(X_{nl+k} \mid J_{l+k-1}$$

$$= s|$$

$$\leqslant A_n^{-2} \max_{1 \leqslant l \leqslant n} \mathrm{a}_l \Big\{ \max_{1 \leqslant l \leqslant n} \Big| \int_{|x| \leqslant A_n/a_l} x d\mathrm{P}(X_l \leqslant x \mid J_{l-1} = r) \Big| \Big\}^2 \sum_{k=1}^{n} |\mathrm{P}(J_k = s \mid J_0 = r)$$

$$- \pi_s| \sum_{l=1}^{n} a_l = U_n V_n W_n,$$

where $U_n = A_n^{-1} \max_{1 \leqslant l \leqslant n} \mathrm{a}_l \to 0$ by hypothesis, as $n \to \infty$,

$$V_n = \Big\{ \max_{1 \leqslant l \leqslant n} \Big| \int_{|x| \leqslant A_n/a_l} x dP(X_l \leqslant x \mid J_{l-1} = r \Big| \Big\}^2$$

has a finite limit as $n \to \infty$, using condition (2.1) and

$$W_n = \sum_{k=1}^{n} |P(J_k = s \mid J_0 = r) - \pi_s|$$

has a finite limit as $n \to \infty$ because of the exponential rate of convergence of the distribution of J_n to the stationary distribution π_s (see e.g. [4]).

Hence, since $\mathrm{Var}(S_{nn}/A_n) \to 0$ as $n \to \infty$, a simple Chebyshev's inequality allows us to conclude that

$$|A_n^{-1} S_{nn} - \sum_{i=1}^{m} \pi_i \mu_i| \leqslant A_n^{-1} S_{nn} - \mathrm{E}(A_n^{-1} S_{nn})| + |A_n^{-1} S_{nn} - \pi_i \mu_i|$$

converges to 0 in probability as $n \to \infty$. The proof is completed by using Lemma 2.1 to allow us to proceed to an arbitrary initial distribution. We have immediately

COROLLARY. *Let $\{(J_n, X_n), n = 0, 1, \ldots\}$ be an m-state ($m < \infty$) ergodic J–X process. Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of positive weights such that*

$$\max_{1 \leqslant k \leqslant n} a_k/A_n \to 0 \quad \textit{as} \quad n \to \infty, \quad \textit{where} \quad A_n = \sum_{k=1}^{n} a_k.$$

If

$$\int x dP_\pi(X_j \leqslant x \mid J_{j-1} = i) = \mu_i < \infty \quad \textit{for each } i, \tag{2.3}$$

then

$$A_n^{-1} \sum_{k=1}^{n} a_k X_k \xrightarrow{\mathrm{P}} \sum_{i=1}^{m} \pi_i \mu_i.$$

Proof. (2.3) implies condition (2.1)

3. Convergence of weighted sums to the Normal law. In this section we provide a sufficient condition on the weights $\{a_{k,n}\}$, for the convergence of the suitably normalized sums of the holding times of a J–X process, to the Normal law. This condition is similar to that placed on the weights in Theorem 2.1. Again, the result is the obvious generalization of the result for independent random variables.

THEOREM 3.1. *Let $X_1, X_2, \ldots$ be the holding times of an m-state homogeneous J–X process as defined in Section 1. Suppose that $\mathrm{E}(X_i) = 0$ and that*

$$\int_{-\infty}^{\infty} x^2 dG_i(x) = \sigma_i^2 < \infty, \quad \textit{where} \quad G_i(x) = \mathrm{P}(X_n \leqslant x \mid J_{n-1} = i).$$

Consider the double array

$$\begin{array}{lllll} a_{11} X_1 & & & & \\ a_{21} X_1 & a_{22} X_2 & & & \\ a_{31} X_1 & a_{32} X_2 & a_{33} X_3 & & \\ \ldots & \ldots & \ldots & & \\ a_{k1} X_1 & a_{k2} X_2 & a_{k3} X_3 & \ldots & a_{kn} X_n. \end{array}$$

Then, provided the sequence of weights $\{a_{kn}\}$ *satisfies the condition*

$$\max_{1\leqslant k\leqslant n}\left(a_{kn}^2/\sum_{k=1}^{n} a_{kn}^2\right)\to 0 \text{ as } n\to\infty,$$

$$\mathrm{P}\left[\frac{a_{k1}X_1+a_{k2}X_2+\ldots+a_{kn}X_n}{s_n}\leqslant x\right]\xrightarrow{\mathrm{D}}\varphi(x) \text{ as } m\to\infty,$$

where, in the notation of Gyires [4],

$$s_n^2=\sum_{k=1}^{n}\sum_{i=1}^{m} D_i\sigma_i^2 a_{kn}^2/(D_1+D_2+\ldots+D_n).$$

Proof. The random variables, $X_{nk}=a_{kn}X_k$, for $n=1, 2, \ldots$ and $k=1, 2, \ldots, n$, form the holding times of an n-step J–X process [3] with

$$\mathrm{P}(X_{nk}\leqslant x\,|_n J_{k-1}=i)=G_i(x/a_{nk}). \tag{3.1}$$

The notation on the left-hand side of (3.1) has the obvious meaning of referring to the n-th n-step Markov chain. Also, we have

$$\mathrm{Var}(X_{nk}\,|_n J_{k-1}=i)=a_{nk}^2\sigma_i^2.$$

It now follows [5] that it is sufficient to show that the Lindeberg condition

$$K_n=\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x|>\varepsilon s_n} x^2\,dG_i\left[\frac{x}{a_{kn}}\right]\to 0 \text{ as } n\to\infty \text{ for each } i=1, 2, \ldots, m, \tag{3.2}$$

is satisfied for the sequence of non-homogeneous n-step processes. By making the transformation $y=x/a_{nk}$ in expression (3.2), we may write

$$K_n=\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{C_{nk}} a_{kn}^2 y^2\,dG_i(y),$$

where

$$C_{nk}=\left\{y\colon |y|>\varepsilon\sqrt{\sum_{k=1}^{n} a_{kn}^2\sum_{i=1}^{m} D_i\sigma_i^2/(D_1+\ldots+D_m)}\Big/a_{kn}\right\}.$$

Hence

$$K_n\leqslant\frac{1}{B^2\sum_{k=1}^{n} a_{kn}^2}\sum_{k=1}^{n} a_{kn}^2\int_A y^2\,dG_i(y)=\frac{1}{B^2}\int_A y^2\,dG_i(y),$$

where

$$B^2=\sum_{i=1}^{m} D_i\sigma_i^2/(D_1+\ldots+D_m),\qquad A=\{y\colon |y|>\varepsilon B[\sum_{k=1}^{n} a_{kn}^2]^{1/2}/\max_{1\leqslant k\leqslant n} a_{kn}\}.$$

Finally, because $\alpha_i^2<\infty$ for each i, $K_n\to 0$ as $n\to\infty$ and the proof of Theorem 3.1 is complete.

REFERENCES

[1] G. L. O'Brien, *Limit theorems for sums of chain dependent processes*, J. Appl. Probab. 11 (1974), p. 582–587.

[2] E. Çinlar, *Introduction to stochastic processes*, Prentice-Hall, New Jersey 1975.

[3] R. L. Dobrushin, *Central limit theorem for nonstationary Markov chains*, Th. Probab. Appl. 1.4 (1956), p. 329–383.

[4] J. L. Doob, *Stochastic processes*, John Wiley, New York 1967.

[5] B. Gyires, *Eine Verallgemeinerung des zentralen Grentzwertsatzes*, Acta Math. Acad. Sci. Hung. 13 (1962).

[6] B. Jamison, S. Orey and W. Pruitt, *Convergence of weighted averages of independent random variables*, Zeitschr. Wahr. 4 (1965), p. 40–44.

Department of Mathematics and Statistics
McGill University
805 Sherbrooke Street, West
Montreal, Quebec
Canada H3A 2K6

Received on 13. 8. 1986