LIMITING DISTRIBUTIONS OF FUNCTIONALS OF MARKOV CHAINS

Rajeeva L. KARANDIKAR*

Center for Stochastic Processes, University of North Carolina, Chapel Hill, NC 27514, USA

Vidyadhar G. KULKARNI

Curriculum in Operations Research and Systems Analysis, University of North Carolina, Chapel Hill, NC 27514, USA

Let $\{X_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ be two stochastic processes such that Y_n depends on X_n in a stationary manner, i.e. $P(Y_n \in A | X_n)$ does not depend on n. Sufficient conditions are derived for Y_n to have a limiting distribution. If X_n is a Markov chain with stationary transition probabilities and $Y_n = f(X_n, \ldots, X_{n+k})$ then Y_n depends on X_n is a stationary way. Two situations are considered: (i) $\{X_n, n \geq 0\}$ has a limiting distribution (ii) $\{X_n, n \geq 0\}$ does not have a limiting distribution and exits every finite set with probability 1. Several examples are considered including that of a non-homogeneous Poisson process with periodic rate function where we obtain the limiting distribution of the interevent times.

Markov chains ∗ *limiting distributions* ∗ *periodic nonhomogeneous Poisson processes*

Section 1

Let $\{X_n, n \geq 0\}$ be a discrete time Markov chain with stationary transition probabilities. Consider a process $\{Y_n, n \geq 0\}$ defined by

$$Y_n = f(X_n, X_{n+1}, \ldots, X_{n+k}). \tag{1.1}$$

It should be emphasized that f and k do not depend upon n. In this paper we address the following question: under what conditions on the Markov chain $\{X_n, n \geq 0\}$ and the function f will $\{Y_n, n \geq 0\}$ have a limiting distribution?

As a first example of the above situation, suppose $\{X_n, n \geq 0\}$ is a random walk i.e. $X_0 = 0$ and $X_n = Z_1 + Z_2 + \cdots + Z_n$ where $\{Z_n, n \geq 0\}$ is a sequence of i.i.d. random variables. Let $Y_n = f(X_n, X_{n+1}) = X_{n+1} - X_n$. Here $\{X_n, n \geq 0\}$ itself does not possess a limiting distribution (except in the trivial case where $Z_n = 0$ w.p. 1 for all $n \geq 0$), but $\{Y_n, n \geq 0\}$ does have a limiting distribution (in fact it is a sequence of i.i.d. r.v.)

As a second example consider a nonhomogeneous Poisson process with rate function $\lambda(t)$. Suppose that $\lambda(t)$ is a periodic function of t. Let X_n be the nth event

occurrence time and let $Y_n = f(X_n, X_{n+1}) = X_{n+1} - X_n$ be the nth interevent time. Now, $\{X_n, n \geq 0\}$ is a transient Markov chain, but due to the periodic nature of $\lambda(t)$, one expects $\{Y_n, n \geq 0\}$ to have a limiting distribution. This example is treated in detail in Example 2 of Section 3.

Though we have stated the problem for a Markov chain $\{X_n, n \geq 0\}$ and its functional process $\{Y_n, n \geq 0\}$, the theory that we develop in the next section, in fact, does not use the Markov property of $\{X_n, n \geq 0\}$ or the functional dependence of Y_n on $X_n, \ldots, X_{n+k}$. The general structure that we assume is as follows:

Let $\{X_n, n \geq 0\}$ be a sequence of random variables and $\{Y_n, n \geq 0\}$ be another sequence of random variables with the property that

$$P(Y_n \in A \mid X_n) \quad \text{does not depend on } n. \tag{1.2}$$

We derive a sufficient condition under which $\{Y_n, n \geq 0\}$ has a limiting distribution. Notice that if $\{X_n, n \geq 0\}$ is a Markov chain with stationary transition probabilities and Y_n is defined by (1.1) then condition (1.2) is automatically satisfied.

In Theorem 1 in the next section we state a sufficient condition for the existence of limiting distribution of $\{Y_n, n \geq 0\}$. We also show with an example that the condition is not necessary. In the general setting of the theorem is seems difficult to derive a useful necessary condition. Even though the condition stated in Theorem 1 is only sufficient, it is nonetheless a powerful tool to unify several cases as is shown by the material in Sections 3 and 4.

In Section 3 we consider stochastic processes $\{X_n, n \geq 0\}$ having a limiting distribution. From Theorem 1 we obtain Proposition 1 which gives the limiting distribution of $\{Y_n, n \geq 0\}$. In Section 4 we consider countable state space stochastic processes $\{X_n, n \geq 0\}$ which do not possess a limiting distribution and have the property that $\{X_n, n \geq 0\}$ exits every finite set with probability 1. In Proposition 2 we state a sufficient condition for $\{Y_n, n \geq 0\}$ to have a limiting distribution in this case. Several examples are given to illustrate both the propositions. Although the results are derived for general stochastic processes, the examples deal with Markov chains $\{X_n, n \geq 0\}$. This is purely for the sake of computational ease.

Limit theorems have been studied in the literature for the case when $\{X_n, n \geq 0\}$ is a Markov chain and $Y_n = f(X_n)$. These limit theorems deal with the partial sums $\sum_{i=1}^{n} Y_i$. (See [3].) We are not aware of any theorems for $\{Y_n, n \geq 0\}$ itself. Another specific problem that has been addressed in the literature is: If $\{X_n, n \geq 0\}$ is a Markov chain, under what conditions is $Y_n = f(X_n)$ also a Markov chain? (See [4].) In this paper we are interested in the limiting behavior and not the Markov nature of $\{Y_n, n \geq 0\}$.

Section 2

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(E, \mathcal{E})$ be a measurable space. Let $\{X_n, n \geq 0\}$ be a sequence of $(E, \mathcal{E})$ valued random variables on $(\Omega, \mathcal{F}, P)$. Let S be

a complete separable metric space and let $\mathscr{B}(S)$ be its Borel σ-field and $\mathscr{M}(S)$ be the space of probability measures on $(S, \mathscr{B}(S))$. Let $\{Y_n, n \geq 0\}$ be a sequence of $(S, \mathscr{B}(S))$ valued random variables on $(\Omega, \mathscr{F}, P)$ such that

$$P(Y_n \in A \mid \sigma(X_n)) = p_{X_n}(A) \tag{2.1}$$

where $p.(\cdot)$ is a mapping from $E \times \mathscr{B}(S) \to \mathbb{R}$ such that

$$\text{for all } x \in E, p_x(\cdot) \in \mathscr{M}(S), \tag{2.2}$$

$$\text{for all } A \in \mathscr{B}(S), p.(A) \text{ is a measurable function on } (E, \mathscr{E}). \tag{2.3}$$

Here, $\sigma(X_n)$ denotes the smallest σ-field on Ω with respect to which X_n is measurable.

Let $S^* = \mathscr{M}(S)$ be equipped with the topology of weak convergence. (See [1, 6]). Recall that $\mu_n \to \mu$ in S^* iff for all bounded continuous functions f on S $\int f \, d\mu_n \to \int f \, d\mu$. S^* itself is a complete separable metric space under this topology. (See [6].) Let $\mathscr{B}(S^*)$ and $\mathscr{M}(S^*)$ denote the Borel σ-field on S^* and the space of probability measures on $(S^*, \mathscr{B}(S^*))$ respectively. $\mathscr{M}(S^*)$ is also equipped with the topology of weak convergence.

Using (2.2) and (2.3) it can be shown that $x \to p_x$ is a measurable mapping from $(E, \mathscr{E})$ into $(S^*, \mathscr{B}(S^*))$ and hence p_{X_n} is a $(S^*, \mathscr{B}(S^*))$ valued random variable. Let Γ_n denote the distribution of p_{X_n}, i.e., for $B \in \mathscr{B}(S^*)$,

$$\Gamma_n(B) = P(p_{X_n}(\cdot) \in B). \tag{2.4}$$

With these notations we have the following

Theorem 1. *Suppose*

$$\Gamma_n \to \Gamma \quad (say) \quad weakly \ in \ \mathscr{M}(S^*), \tag{2.5}$$

then Y_n converges in distribution to a measure $\nu \in S^*$ given by

$$\nu(A) = \int_{S^*} \mu(A) \, d\Gamma(\mu) \quad (A \in \mathscr{B}(S)). \tag{2.6}$$

Proof. Let f be a bounded continuous function on S and let $F: S^* \to \mathbb{R}$ be defined by

$$F(\mu) = \int f \, d\mu. \tag{2.7}$$

Then, by the definition of weak convergence, it follows that F is a bounded continuous function. Thus, from (2.5), we get

$$\int F(\mu) \, d\Gamma_n(\mu) \to \int F(\mu) \, d\Gamma(\mu). \tag{2.8}$$

Now

$$\int F(\mu) \, d\Gamma_n(\mu) = E[F(p_{X_n}(\cdot))] = E\left[\int f(z) p_{X_n}(dz)\right]. \tag{2.9}$$

From (2.1), (2.2) and (2.3) it follows that

$$E(f(Y_n)|\sigma(X_n)) = \int f(z) p_{X_n}(\mathrm{d}z)$$

and hence by (2.9)

$$\int F(\mu) \, \mathrm{d}\Gamma_n(\mu) = E[E[f(Y_n)|\sigma(X_n)]] = E[f(Y_n)]. \qquad (2.10)$$

Also,

$$\int F(\mu) \, \mathrm{d}\Gamma(\mu) = \int_{S^*} \left\{ \int_S f(z) \, \mathrm{d}\mu(z) \right\} \mathrm{d}\Gamma(\mu) = \int_S f(z) \, \mathrm{d}\nu(z). \qquad (2.11)$$

The last equality in (2.11) follows from the definition of ν for a simple function f and then, by the usual arguments for a general function f.

Thus from (2.8), (2.10) and (2.11) we have, for all bounded continuous functions f on S,

$$E(f(Y_n)) \to \int f(z) \, \mathrm{d}\nu(z) \qquad (2.12)$$

and hence, Y_n converges in distribution to ν. $\square$

Remark. The condition (2.5) is not necessary as is shown by the following example. Let $E = \{0, \frac{1}{2}, 1\}$ and let $\{X_n, n \geq 0\}$ be such that, for $n \geq 0$,

$$P(X_{2n} = 0) = P(X_{2n} = 1) = \tfrac{1}{2}, \qquad P(X_{2n+1} = \tfrac{1}{2}) = 1.$$

Let $S = \{0, 1\}$ and define $p_x(\cdot)$ as follows:

$$p_x(\{0\}) = x = 1 - p_x(\{1\}) \quad (x \in E).$$

Let $\{Y_n, n \geq 0\}$ be S valued random variables satisfying (2.1). Then, for all $n \geq 0$,

$$P(Y_n = 0) = EX_n = \tfrac{1}{2}$$

and hence, trivially, Y_n converges in distribution. However, it is easy to see that Γ_n does not converge weakly.

The above theorem provides a general framework to study the problems mentioned in the introduction. In later sections, we discuss several special problems and obtain sufficient conditions for convergence in distribution of $\{Y_n, n \geq 0\}$. In each of these cases, the results could be proved by alternate techniques, but the above theorem, whose proof is simple, provides a unified view of the problem.

Section 3

In this section we restrict ourselves to the class of stochastic processes $\{X_n, n \geq 0\}$ which possess limiting distributions, i.e. X_n converges in distribution to some

measure π as $n \to \infty$. In the framework of Section 2, suppose that E is a metric space and $\mathscr{E}$ is its Borel σ-field. The following is an easy consequence of Theorem 1:

Proposition 2. *Suppose that*

$$x \to p_x \text{ is a continuous function from } E \text{ into } S^* \text{ and} \tag{3.1}$$

$$x_n \text{ converges in distribution to some measure } \pi \text{ on } (E, \mathscr{E}). \tag{3.2}$$

Then Y_n converges in distribution to a measure ν on $(S, \mathscr{B}(S))$ defined by

$$\nu(A) = \int_E p_x(A) \, d\pi(x) \quad (A \in \mathscr{B}(S)). \tag{3.3}$$

Proof. Equations (3.1) and (3.2) imply that p_{X_n} converges in distribution to $\Gamma = \pi \circ (p.)^{-1}$, which is the same as $\Gamma_n \to \Gamma$ weakly. Hence Y_n converges in distribution to ν given by, for $A \in \mathscr{B}(S)$,

$$\nu(A) = \int_{S^*} \mu(A) \, d\Gamma(\mu) = \int_{S^*} \mu(A) \, d\pi \circ (p.)^{-1}(\mu) = \int_E p_x(A) \, d\pi(x). \quad \square$$

We now give two examples illustrating the above result. In the remaining paper, all finite or countable sets will be equipped with the discrete topology. Our first example is that of a positive recurrent Markov chain.

Example 1. Let $E = \{0, 1, 2, \ldots\}$ and let $\{X_n, n \geqslant 0\}$ be an E-valued Markov chain with stationary transition probabilities given by

$$P(X_1 = k + 1 \mid X_0 = k) = \beta \quad (k \geqslant 1),$$

$$P(X_1 = k - 1 \mid X_0 = k) = \alpha = 1 - \beta \quad (k \geqslant 1), \qquad P(X_1 = 1 \mid X_0 = 0) = 1.$$

Let $S = \{-1, 1\}$ and $Y_n = X_{n+1} - X_n$. Then (2.1) is satisfied with

$$p_k(\{1\}) = \beta \quad (k \geqslant 1), \qquad p_k(\{-1\}) = \alpha \quad (k \geqslant 1), \qquad p_0(\{1\}) = 1.$$

Since the sets E and S are equipped with discrete topology, conditions (2.3) and (3.1) are trivially satisfied. Now suppose $\alpha > \beta > 0$. Then $\{X_n, n \geqslant 0\}$ is positive recurrent and its stationary distribution is given by

$$\pi_0 = \frac{\alpha - \beta}{2\alpha}, \qquad \pi_k = \frac{\alpha - \beta}{2\alpha\beta} \left(\frac{\beta}{\alpha}\right)^k \quad (k \geqslant 1).$$

Hence $\{X_n, n \geqslant 0\}$ converges in distribution to π. Hence, by Proposition 2, Y_n converges in distribution to ν where

$$\nu(\{1\}) = \lim_{n \to \infty} P(Y_n = 1) = \sum_{k=0}^{\infty} p_k(\{1\}) \pi_k = \tfrac{1}{2}$$

and

$$\nu(\{-1\}) = \lim_{n \to \infty} P(Y_n = -1) = \tfrac{1}{2}.$$

Thus, Y_n converges in distribution to $(\tfrac{1}{2}, \tfrac{1}{2})$. Note that Y_n, the increment of the $\{X_n, n \ge 0\}$ process, has a stationary distribution independent of α!

Our next example is that of a nonhomogeneous Poisson process.

Example 2. Let $\{N(t), t \ge 0\}$ be a nonhomogeneous Poisson process with strictly positive rate function $\lambda(t)$. Assume that $\lambda(t)$ is periodic with period τ, i.e.,

$$\lambda(t + \tau) = \lambda(t) \quad \text{for all } t \ge 0.$$

Let $Z_0 = 0$ and Z_n be the time that the nth event occurs, i.e.

$$Z_{n+1} = \inf\{t \ge Z_n : N(t) > N(Z_n)\}, \quad n \ge 0.$$

and let $Y_n = Z_{n+1} - Z_n$ be the nth interevent time. Now, for $0 \le v_1 \le v_2 \le \cdots \le v_n \le u$,

$$P(Z_{n+1} > u \mid Z_1 = v_1, \ldots, Z_n = v_n)$$

$$= P(N(u) = n \mid N(s): 0 \le s \le v_n, N(v_n) = n) = \exp\left(-\int_{v_n}^{u} \lambda(s)\, ds\right).$$

Hence $\{Z_n, n \ge 0\}$ is a Markov chain. As Z_n increases with n, it is a transient Markov chain.

Let $[x]$ denote the largest integer $\le x$ and define

$$X_n = Z_n - [Z_n / \tau]\tau.$$

Also, for $t \ge 0$, define

$$\Lambda(t) = \int_0^t \lambda(s)\, ds.$$

Now, for $0 \le y < \infty$,

$$P(Y_n > y \mid Z_n) = P(Z_{n+1} > Z_n + y \mid Z_n)$$

$$= \exp(-(\Lambda(Z_n + y) - \Lambda(Z_n))) = \exp(-(\Lambda(X_n + y) - \Lambda(X_n))).$$

For $0 \le x < \tau$, let p_x denote the probability measure on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ given by

$$p_x((y, \infty)) = \exp(-(\Lambda(x + y) - \Lambda(x))). \tag{3.4}$$

Then for $A \in \mathcal{B}(\mathbb{R}^+)$, we have

$$P(Y_n \in A \mid Z_n) = P(Y_n \in A \mid X_n) = p_{X_n}(A).$$

Let $E = [0, \tau)$. Equip E with a topology which makes the mapping $x \to \exp(i2\pi x/\tau)$ a homeomorphism from E onto the unit circle in the complex plane. Thus

$$x_n \to x \quad \text{iff} \quad \exp(i2\pi x_n/\tau) \to \exp(i2\pi x/\tau).$$

Under this topology E is a complete separable metric space and it can be easily checked that the mapping $x \to p_x$ in (3.4) from $E \to \mathcal{M}(\mathbb{R})$ is continuous.

We shall now show that $\{X_n, n \geq 0\}$ converges in distribution to some measure on E. In fact, it can be shown that $\{X_n, n \geq 0\}$ is a Markov chain on E with stationary transition probabilities given by

$$P(X_{n+1} > y \mid X_n = x) = e^{\Lambda(x)}(e^{-\Lambda(y)} - B)/(1 - B) \qquad \text{if } 0 \leq x < y < \tau,$$

$$= 1 - B e^{\Lambda(x)}(1 - e^{-\Lambda(y)})/(1 - B) \qquad \text{if } 0 \leq y \leq x < \tau. \qquad (3.5)$$

where $B = \exp(-\Lambda(\tau))$.

Let $f(\cdot, x)$ be the conditional density of X_{n+1} given $X_n = x$. Then from (3.5) we get

$$f(y, x) = \begin{cases} \lambda(y) \exp(\Lambda(x) - \Lambda(y))/(1 - B) & \text{if } 0 \leq x < y < \tau, \\ B\lambda(y) \exp(\Lambda(x) - \Lambda(y))/(1 - B) & \text{if } 0 \leq y \leq x < \tau. \end{cases} \qquad (3.6)$$

By a slight modification of the arguments in Example (b) of VII.7 of Feller Vol. 2 [2], one can show that $\{X_n, n \geq 0\}$ has a limiting distribution. Let $g(\cdot)$ be the limiting density of X_n. Then $g(\cdot)$ satisfies

$$g(y) = \int_0^\tau f(y, x)g(x) \, dx.$$

Thus

$$g(y) = \frac{\lambda(y) \exp(-\Lambda(y))}{1 - B} \left[\int_0^y e^{\Lambda(x)}g(x) \, dx + B \int_y^\tau e^{\Lambda(x)}g(x) \, dx \right].$$

Now let $h(y) = g(y) \exp(\Lambda(y))$. Then

$$h(y) = \frac{\lambda(y)}{1 - B} \left[\int_0^y h(x) \, dx + B \int_y^\tau h(x) \, dx \right].$$

Differentiating the above equation we get

$$h'(y) = \left(\frac{\lambda'(y)}{\lambda(y)} + \lambda(y) \right) h(y)$$

which has a unique solution

$$h(y) = \lambda(y) \exp(\Lambda(y)).$$

Hence we get $g(y) = C\lambda(y)$, the constant of integration C is found by using

$$\int_0^\tau g(y) \, dy = 1.$$

Thus $g(y) = \lambda(y)/\Lambda(\tau)$ $(0 \leq y < \tau)$ is the limiting density of X_n.

Now, using Proposition 2, we get the following: $\{Y_n, n \geq 0\}$ has a limiting distribution given by

$$\lim_{n \to \infty} P(Y_n > y) = \int_0^\tau \exp(-(\Lambda(x + y) - \Lambda(x)))\lambda(x) \, dx/\Lambda(\tau). \qquad (3.7)$$

Section 4

In this section we consider the case in which $\{X_n, n \geq 0\}$ does not converge in distribution, but $\{Y_n, n \geq 0\}$ does. First we give a general result and then consider some specific examples.

Suppose that E is a countable set and that

$$\lim_{n \to \infty} P(X_n \in B) = 0 \tag{4.1}$$

for all finite subsets B of E. Furthermore, suppose that there exists a partition $\{E_j, j \in J\}$ of E, where either $J = \{1, 2, \ldots, k\}$ or $J = \{1, 2, 3, \ldots\}$, and each E_j is a countable set such that

$$\lim_{n \to \infty} P(X_n \in E_j) = \alpha_j \quad \text{exists} \tag{4.2}$$

and

$$\sum_{j \in J} \alpha_j = 1. \tag{4.3}$$

Now, let $\{x_{ij}\}$, $i \geq 0$ be an enumeration of E_j and suppose that

$$\lim_{i \to \infty} p_{x_{ij}} = \mu_j \tag{4.4}$$

exists. (Recall that $\mathcal{M}(S)$ is equipped with the topology of weak convergence.) With this structure we get the following

Proposition 3. *Let* (4.1)–(4.4) *hold. Then* $\{Y_n, n \geq 0\}$ *converges in distribution to* ν *given by*

$$\nu(A) = \sum_{j \in J} \mu_j(A)\alpha_j \quad (A \in \mathcal{B}(S)). \tag{4.5}$$

Proof. The desired conclusion will follow from Theorem 1 if we show that

$$\Gamma_n \to \Gamma \quad \text{in } \mathcal{M}(S^*) \tag{4.6}$$

where Γ is given by

$$\Gamma(B) = \sum_{j \in J} 1_B(\mu_j)\alpha_j \quad (B \in \mathcal{B}(S^*)).$$

Let f be a bounded continuous function on S^*. Then

$$\left| \int f \, d\Gamma_n - \int f \, d\Gamma \right| = \left| E(f(p_{X_n})) - \sum_{j \in J} f(\mu_j)\alpha_j \right|$$

$$\leq \sum_{j \in J} \left| Ef(p_{X_n}) 1_{\{X_n \in E_j\}} - f(\mu_j)\alpha_j \right|$$

$$\leq \sum_{j \in J} \left| E(f(p_{X_n}) - f(\mu_j)) 1_{\{X_n \in E_j\}} \right| + \sum_{j \in J} |f(\mu_j)| \left| P(X_n \in E_j) - \alpha_j \right|$$

$$= \sum_{j \in J} U_j^n + \sum_{j \in J} V_j^n \text{ (say).} \tag{4.7}$$

Let $\alpha_j^n = P(X_n \in E_j)$. Then

$$
\begin{aligned}
\sum_{j \in J} |\alpha_j^n - \alpha_j| &= \sum_{j \in J} \alpha_j^n + \sum_{j \in J} \alpha_j - 2 \sum_{j \in J} \alpha_j^n \wedge \alpha_j \\
&= 2 - 2 \sum_{j \in J} \alpha_j^n \wedge \alpha_j \to 2 - 2 \sum_{j \in J} \alpha_j = 0.
\end{aligned}
\tag{4.9}
$$

(By dominated convergence theorem.) Since f is bounded, (4.9) implies that

$$
\lim_{n \to \infty} \sum_{j \in J} V_j^n = 0.
\tag{4.10}
$$

Equation (4.9) also implies that, given $\varepsilon_0 > 0$, there exists a finite subset J_0 of J such that,

$$
\sup_n \sum_{j \in J_1} \alpha_j^n < \varepsilon_0
\tag{4.11}
$$

where $J_1 = J \backslash J_0$. Equation (4.11) is actually the assertion that L^1-convergence implies uniform integrability.

Next we show that, for $j \in J$

$$
\lim_{n \to \infty} U_j^n = 0.
\tag{4.12}
$$

Recall that $p_{x_{ij}} \to \mu_j$ as $i \to \infty$ and f is continuous on S^*, so that $f(p_{x_{ij}}) \to f(\mu_j)$. Thus, given $\varepsilon > 0$, there is an i_0 such that, for all $i > i_0$,

$$
|f(p_{x_{ij}}) - f(\mu_j)| < \varepsilon/2.
\tag{4.13}
$$

Let $B = \{x_{ij} : 1 \le i \le i_0\}$. Then in view of (4.1), there is an n_0 such that for $n \ge n_0$,

$$
P\{X_n \in B\} \le \varepsilon/4K
\tag{4.14}
$$

where K is such that $|f| \le K$. Then for $n \ge n_0$, we have

$$
\begin{aligned}
U_j^n &= E\{|f(p_{X_n}) - f(\mu_j)| 1_{\{X_n \in E_j\}}\} \\
&= E\{|f(p_{X_n}) - f(\mu_j)| \{1_{\{X_n \in B\}} + 1_{\{X_n \in E_j \backslash B\}}\}\} \\
&\le 2KP(X_n \in B) + \frac{\varepsilon}{2} P(X_n \in E_j \backslash B) \\
&\le \varepsilon/2 + \varepsilon/2 \quad \text{(by (4.13) and (4.14))} \\
&= \varepsilon.
\end{aligned}
$$

This proves (4.12). Now, for a given $\varepsilon > 0$, define $\varepsilon_0 = \varepsilon/2K$. For this ε_0 get J_0 and J_1 as above, so that (4.11) holds. Then

$$
\sum_{j \in J} U_j^n \le \sum_{j \in J_0} U_j^n + \varepsilon_0.
\tag{4.15}
$$

Since J_0 is finite, we have, using (4.12) and (4.15)

$$
\limsup_{n \to \infty} \sum_{j \in J} U_j^n \le \varepsilon.
\tag{4.16}
$$

Since $\varepsilon > 0$ is arbitrary, this shows that

$$\lim_{n \to \infty} \sum_{j \in J} U_j^n = 0.$$

In view of (4.8), (4.9) and (4.10) we get

$$\lim_{n \to \infty} \left| \int f \, d\Gamma_n - \int f \, d\Gamma \right| = 0$$

which implies $\Gamma_n \to \Gamma$ weakly, and this completes the proof. $\square$

We now take two examples illustrating the above proposition.

Example 3. Consider the same Markov chain $\{X_n, n \geq 0\}$ from Example 1, and suppose $\alpha < \beta$. Then $\{X_n, n \geq 0\}$ is a transient Markov chain and hence (4.1) holds.

Since $\lim_{k \to \infty} p_k = \mu_1$ where $\mu_1(\{1\}) = \beta$ and $\mu_1(\{-1\}) = \alpha$, the conditions of Proposition 3 are satisfied if we take $E_1 = E$ and $x_{i1} = i$ ($i \geq 0$). Then $Y_n = X_{n+1} - X_n$ converges in distribution to μ_1, i.e.

$$\lim_{n \to \infty} P(Y_n = 1) = \beta, \qquad \lim_{n \to \infty} P(Y_n = -1) = \alpha.$$

Example 4. Let $E = \{(j, i): j, i \text{ integers}, 0 \leq j \leq i < \infty\}$. Let $\{X_n\}$ be a Markov chain with state space E and stationary transition probabilities given by

$$P\{X_{n+1} = (1, i+1) \mid X_n = (0, i)\} = 1, \quad i \geq 0.$$

$$P\{X_{n+1} = (j+1, i+1) \mid X_n = (j, i)\} = \beta, \quad 0 < j \leq i < \infty.$$

$$P\{X_{n+1} = (j-1, i) \mid X_n = (j, i)\} = \alpha, \quad 0 < j \leq i < \infty.$$

Where $\alpha + \beta = 1$, $0 < \beta < \alpha < 1$. For this chain all states are inessential, as $(j, i) \to (j+1, i+1)$ but $(j+1, i+1) \nrightarrow (j, i)$. Hence $\{X_n, n \geq 0\}$ is transient and (4.1) is satisfied.

Let $\{Y_n, n \geq 0\}$ be a real valued process such that the conditional distribution of Y_n given $X_n = (j, i)$ is normal with mean $j + 1/(i+1)$ and variance $1/(j+1) + 2^{-i}$. Thus

$$p_{(j,i)} = N\left(j + \frac{1}{i+1}, \frac{1}{j+1} + 2^{-i}\right)$$

where $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Now let $J = \{0, 1, 2, 3, \ldots\}$ and

$$E_j = \{(j, i): i \geq j\}.$$

Also let $x_{ij} = (j, j+i)$ be an enumeration of E_j. Then from the properties of the normal distribution it follows that

$$\lim_{i \to \infty} p_{x_{ij}} = N\left(j, \frac{1}{j+1}\right).$$

Now let $Z_n = j$ if $X_n \in E_j$. It is easily seen that Z_n itself is a Markov chain with stationary transition probabilities as described in Example 1. Thus

$$\lim_{n \to \infty} P(X_n \in E_j) = \lim_{n \to \infty} P(Z_n = j) = \pi_j$$

where

$$\pi_0 = \frac{\alpha - \beta}{2\alpha}, \qquad \pi_j = \frac{\alpha - \beta}{2\alpha\beta} \left(\frac{\beta}{\alpha}\right)^j \quad (j \geq 1).$$

Thus $\sum_{j=0}^{\infty} \pi_j = 1$ and hence (4.2) and (4.3) hold. Hence, by Proposition 3, $\{Y_n, n \geq 0\}$ converges in distribution to a measure ν given by

$$\nu = \frac{\alpha - \beta}{2\alpha} N(0, 1) + \sum_{j=1}^{\infty} \frac{\alpha - \beta}{2\alpha\beta} \left(\frac{\beta}{\alpha}\right)^j N\left(j, \frac{1}{j+1}\right).$$

Section 5

The results in Sections 2, 3 and 4 can be extended to continuous time stochastic processes in a straightforward manner. Extension of the results in Section 4 to more general state spaces seems possible but presents many technical difficulties. The results in Section 4 also suggest a relationship between our approach and the boundary theory for Markov chains, but at this stage we have not been able to make it precise. It should be mentioned that the result in Section 3 about the periodic nonhomogeneous Poisson processes is of interest in itself.

The problem considered in this paper is of interest in the theory of partially observable processes. In this context one can think of $\{X_n, n \geq 0\}$ as the core process and $\{Y_n, n \geq 0\}$ as the observation process. (See [5]). For example X_n may represent the 'state' of the internal components of a machine at time n while Y_n may represent its 'performance', which may be the only observable quantity about the machine. The limiting behavior of the observation process $\{Y_n, n \geq 0\}$ is obviously of importance in the design of the machine.

References

[1] P. Billingsley, Convergence of Probability Measures (Wiley, New York, 1968).
[2] W. Feller, An Introduction to Probability Theory and Its Applications, Vol II (Wiley Eastern Ltd., New Delhi, 1966).
[3] D. Freedman, Markov Chains (Holden Day, San Francisco, 1971).
[4] J.G. Kemeny and J.L. Snell, Finite Markov Chains (D. Van Nostrand Company, Princeton, NJ, 1960).
[5] G.E. Monahan, A survey of partially observable Markov decision processes: Theory, models and algorithms, Management Science 28 (1982) pp 1-16.
[6] K.R. Parthasarathy, Probability Measures on Metric Spaces (Academic Press, New York, 1967).