Statistica Sinica **13**(2003), 227-242

POISSON STYLE CONVERGENCE THEOREMS FOR ADDITIVE PROCESSES DEFINED ON MARKOV CHAINS

Y. H. Wang and Lingqi Tang

Tunghai University and University of California

Abstract: Given a Markov chain $\{X_i : i \geq 0\}$ with finite state space and irreducible primitive stationary transition matrix $\mathbf{P}$, at time n, corresponding to each possible one-step transition $j \to k$, we associate random variables W_n with distribution F_{jk}, depending only on states j and/or k. Given $\{X_i\}$, W_1, $W_2, \ldots,$ are conditionally independent and need not be integer-valued, nor positive. Define the cumulative sum $Y_n = W_1 + \cdots + W_n$, with $Y_0 = 0$. It is proved under certain conditions that the limiting distribution of $\{Y_n\}$ is in the class of compound Poisson type distributions. Some applications of the theorem are illustrated.

Key words and phrases: Compound Poisson distribution, convergence theorem, interarrival time, limiting distribution, Markov chain, Markov renewal process, sum of random variables.

1. Introduction

Let $\{X_n : n \geq 0\}$ be a Markov chain with state space $S = \{0, \ldots, s\}$, $s < \infty$, stationary transition matrix $\mathbf{P} = (p_{ij})_{i,j \in S}$ and initial probability $\mathbf{p} = (p_0, \ldots, p_s)$. It is assumed that $\mathbf{P}$ is irreducible and primitive. ($\mathbf{P}$ is said to be *primitive* if 1 is a simple eigenvalue, see Cox and Miller (1965).)

At time n, corresponding to each one-step transition from state j to state k, we associate a random variable W_n with distribution function F_{jk}. The random variables W_n are in general not independent but, conditional on a given realization of the Markov chain $\{X_n\}$ they are mutually independent. That is

$$\Pr(W_n \leq w \mid X_0 = x_0, \ldots, X_{n-1} = j, X_n = k) = F_{jk}(w),$$

$$\Pr(\cap_{i=1}^{n} \{W_i \leq w_i\} \mid \cap_{i=0}^{n} \{X_i = x_i\}) = \Pi_{i=0}^{n-1} F_{x_i x_{i+1}}(w_{i+1}).$$

The W_n are not necessarily discrete, nor positive. A simple example of a continuous case is the Markov renewal process constructed from the $(s+1)$-state Markov chain by letting W_n be the nth interarrival time, i.e., $F_{jk}(w) = 1 - e^{-\lambda_j w}$, $\lambda_j > 0$, for $w \geq 0$. In this example, W_n depends only on X_{n-1}.

Define the cumulative sum

$$Y_n = W_1 + \cdots + W_n, \quad n \geq 1, \tag{1}$$

with $Y_0 = 0$. The vector process $\{X_n, Y_n\}$ is a two-dimensional Markov process in discrete time. The first sequence $\{X_n\}$ has discrete state space S while the second sequence $\{Y_n\}$ is an additive process which may be discrete or continuous. For detailed discussion of this additive process, see Cox and Miller (1965), pp.352-353, and Pyke (1961).

In contrast to the cumulative sums $\{Y_n\}$, we define the usual partial sum S_n of the Markov chain $\{X_n\}$ as

$$S_n = X_0 + X_1 + \cdots + X_n, \quad n \geq 1. \tag{2}$$

Note that the partial sum S_n has $n+1$ summands while the cumulative sum Y_n has only n.

The sequence of random variables $\{W_n\}$ and its cumulative sums $\{Y_n\}$ so constructed cover a wide spectrum of stochastic processes associated with the Markov chain $\{X_n\}$. Following are some examples.

(a). For a fixed state $\xi \in S$, take $W_n = 1$ if the state ξ is occupied after the nth transition, and $= 0$ otherwise. The cumulative sum Y_n is the number of times state ξ is occupied among the times $1, \ldots, n$. If $\{X_n\}$ is a Markov Bernoulli chain with $\xi = 1$, then $Y_n = S_n - X_0$. More generally, for any finite $s \geq 1$, if we take $W_n = X_n$ for $n \geq 1$, then

$$Y_n = S_n - X_0. \tag{3}$$

In this case $F_{jk}(w) = 1$ if $w \geq k$ and is 0, otherwise. Therefore the sequence of partial sums $\{S_n\}$ can be considered as a special case of the cumulative sums $\{Y_n\}$. From this point of view, many results on Poisson/compound Poisson sequences $\{S_n\}$, such as those obtained by Hsiau (1997), Koopman (1950), Lin and Wang (1994), Pedler (1971, 1978), Wang (1981), Wang and Yang (1995), Wang and Tang (1997), etc., are special cases of the main result of this paper.

(b). Let $W_n = 1$ if the nth transition is "$j \to k$" and $= 0$ otherwise. Then Y_n is the number of times the transition "$j \to k$" occurs in the first n transitions of $\{X_n\}$. If we let $j = k = 1$ in the Markov Bernoulli sequence with "1" denoting "Head", then Y_n is the number of runs of two consecutive heads observed in a series of tossing a "Markovian" coin. For example, $Y_{14} = 5$ for the sequence "01111001110010". (There are many types of "runs". For a detailed analysis of different types of "runs" in Markov Bernoulli trials, see Wang and Liang (1993) and Wang and Ji (1995).)

(c). More generally, let $\{W_n : n \geq 0\}$ be an arbitrary i.i.d. sequence of random variables with $W_0 = 0$, then the additive process $\{Y_n\}$ is a random walk in which the dependence of steps is controlled by the transition matrix $\mathbf{P}$.

Central limit theorems and related results for the additive process $\{Y_n\}$ have been studied by many authors, such as Volkov (1958), Gabriel (1959), Miller (1961, 1962a, 1962b), Kelson and Wishart (1964) and Mathews (1970).

In this paper, we investigate the Poisson style convergence theorems of the additive sequence $\{Y_n\}$. We show that, under certain conditions, the limiting distribution of $\{Y_n\}$ is in the class of the compound Poisson distributions. Different problems, such as the occupation times, the number of transitions, etc., considered in Pedler (1978) for the two-state Markov chain, can be consolidated as a single problem through our approach.

2. Preliminaries

Let

$$\phi_{jk}(t) = \int_{-\infty}^{\infty} e^{itx} F_{jk}(dx) = E\left(e^{itW} \mid X_0 = j, X_1 = k\right),$$

be the characteristic function (ch.f.) of F_{jk}. Denote by $F_{jk}^{(n)}$ the distribution function of Y_n conditional on the n-stage transition from state j at time 0 to state k at time n, then the ch.f. of $F_{jk}^{(n)}$ is

$$\phi_{jk}^{(n)}(t) = \int_{-\infty}^{\infty} e^{itx} F_{jk}^{(n)}(dx) = E(e^{itY_n} \mid X_0 = j, X_n = k).$$

Denote the n-step transition probabilities by $\mathbf{P}^{(n)} = (p_{jk}^{(n)})$. Define the matrices $\mathbf{P}(t) = (p_{jk}\phi_{jk}(t))_{j,k\in S}$ and $\mathbf{P}^{(n)}(t) = (p_{jk}^{(n)}\phi_{jk}^{(n)}(t))_{j,k\in S}$. The matrix $\mathbf{P}^{(n)}(t)$ is not the usual n-step transition stochastic matrix, but by conditioning on X_{n-1} and using mathematical induction, it can be shown that it satisfies the improtant property $\mathbf{P}^{(n)}(t) = \mathbf{P}^n(t)$. (See Cox and Miller (1965), p.136, or Miller (1962a) for details.) The unconditional ch.f. of Y_n is thus

$$G_n(t) = E(e^{itY_n}) = \mathbf{p}\mathbf{P}^n(t)\mathbf{1}, \quad t \in R, \tag{4}$$

where $\mathbf{1}$ is the unit column vector in R^{s+1} and $\mathbf{p} = (p_0, \ldots, p_s)$ is the initial probability distribution of X_0. We consider some simple examples of (4).

For the two-state case, if we take $W_j = X_j$ and $Y_n = S_n - X_0$, then Y_n is the occupation times of state 1 and

$$\mathbf{P}(t) = \begin{pmatrix} p_{00} & (1-p_{00})e^{it} \\ p_{10} & (1-p_{10})e^{it} \end{pmatrix}. \tag{5}$$

If we take $W_j = 1$ when "$1 \to 1$" transition occurs at time j and $= 0$ otherwise, then Y_n is the total number of "runs" of type "$1 \to 1$" and

$$\mathbf{P}(t) = \begin{pmatrix} p_{00} & (1-p_{00}) \\ p_{10} & (1-p_{10})e^{it} \end{pmatrix}. \tag{6}$$

The ch.f. $\mathbf{P}(t)$ in (5) and (6) can be found in probability generating function (p.g.f.) form in Koopman (1950), Edwards (1960) and Pedler (1971).

For the general finite-state case, take $W_n = X_n$, so that

$$\mathbf{P}(t) = \left(p_{jk}e^{ikt}\right)_{j,k\in S}. \tag{7}$$

With $Y_n = S_n - X_0$, the unconditional ch.f. of Y_n is (4) with $\mathbf{P}(t)$ as given in (7). If we take $X_0 = 0$ *a.s.*, the unconditional ch.f. of Y_n is

$$G_n(t) = (p_{00}, p_{01}e^{it}, \ldots, p_{0s}e^{ist})\mathbf{P}(t)^{n-1}\mathbf{1}. \tag{8}$$

Since $(p_{00}, \ldots, p_{0s})$ is the unconditional probability of X_1, it can be considered as the "initial" probability of the Markov chain $\{X_i : i \geq 1\}$. The ch.f. (8) in p.g.f. form was first obtained by Lin and Wang (1994), equation (1.5), and later independently derived by Hsiau (1997), Lemma 3.2. The p.g.f. in Koopman (1950) is a special case of (8) for the 2×2 case.

It can be shown that (8) can also be written as $G_n(t) = (1, e^{it}, \ldots, e^{ist})\times \mathbf{Q}^{n-1}(t)\mathbf{q}$, where $\mathbf{Q}(t) = (p_{kj}e^{ikt})$ and $\mathbf{q} = (p_{00}, \ldots, p_{0s})^T$. (See Lin and Wang (1994), Theorem 1.1.)

The arguments which lead to (8) can be used to show that the ch.f. of S_n, with initial probability $\mathbf{p}$, is

$$G_n(t) = \mathbf{p}(\mathbf{t})\mathbf{P}(t)^n\mathbf{1}, \tag{9}$$

where $\mathbf{p}(t) = (p_0, p_1e^{it}, \ldots, p_se^{ist})$.

3. The Main Results

3.1. The limit conditions

To state our limit conditions and results, we need to introduce a triangular array of random variables X_{ni}, $i = 0, \ldots, n$, $n = 1, \ldots,$ defined on the same state space S, where the nth row $X_{n0}, \ldots, X_{nn}$ forms a Markov chain with transition matrix $\mathbf{P}_n = (p_{njk})$, j, $k \in S$, and initial probability $\mathbf{p}_n = (p_{n0}, \ldots, p_{ns})$. Let W_{ni}, $i = 1, \ldots, n$, be the associated random variables with $F_{njk}(w) = \Pr(W_{ni} \leq w \mid X_{n,i-1} = j,\ X_{ni} = k)$, $Y_{ni} = W_{n1} + \cdots + W_{ni}$, $i = 1, \ldots, n$, $(Y_{n0} \equiv 0)$, be the associated additive process. The ch.f.'s $\phi_{njk}(t)$ and $\phi^{(n)}_{njk}(t) = E(e^{itY_{nn}} \mid X_{n0} = j, X_{nn} = k)$ and the matrices $\mathbf{P}_n(t) = (p_{njk}\phi_{njk}(t))_{j,k\in S}$ and $\mathbf{P}^{(n)}_n(t) = (p^{(n)}_{njk}\phi^{(n)}_{njk}(t))_{j,k\in S}$ are as defined in the previous section and we take $G_{nj}(t) = E(e^{itY_{nj}})$.

Let $0 \leq c < s$, $C = \{0, \ldots, c\}$ and $S = C \cup C^c$. The following limit conditions, as $n \to \infty$, are needed for the main results.

$$(\mathbf{C}-\mathrm{i}) \qquad \mathbf{P}_n \to \mathbf{A} = \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{T} & \mathbf{Q} \end{pmatrix},$$

where $\mathbf{R} = (r_{jk})$ is a $(c+1) \times (c+1)$ irreducible primitive stochastic matrix, $\mathbf{T} =(t_{jk}) \neq \mathbf{0}$ is a $(s-c) \times (c+1)$ matrix and $\mathbf{Q} =(q_{jk})$ is a $(s-c) \times (s-c)$ substochastic matrix such that the row sums of $\mathbf{Q}$ are all less than 1.

$$(\mathbf{C}-\text{ii}) \qquad np_{njk} \to \lambda_{jk}, \qquad \text{for } j \in C \text{ and } k \in C^c.$$

$$(\mathbf{C}-\text{iii}) \qquad \mathbf{p}_n \to \boldsymbol{\rho} = (\rho_0, \ldots, \rho_s), \qquad \textstyle\sum_{j=0}^{s} \rho_j = 1, \rho_j \geq 0.$$

$$(\mathbf{C}-\text{iv}) \qquad \begin{cases} n(\phi_{njk}(t)-1) \to 0 & \text{for } j,k \in C, \\ \phi_{njk}(t) \to \varphi_{jk}(t) & \text{for } j \in C^c \text{ or } k \in C^c, \end{cases}$$

where $\varphi_{jk}(t)$ are ch.f.'s.

Condition ($\mathbf{C}$−i) implies that, even though the transition matrix $\mathbf{P}_n$ is irreducible, it is asymptotically reducible. Its limit has subspace C as an ergodic class. Evidently, the cumulative sums $\{Y_{nn}\}$ would diverge with probability one if $\mathbf{P}_n$ were not asymptotically reducible.

Condition ($\mathbf{C}$−ii) implies that the rate of convergence of p_{njk} to 0 is of order $O(1/n)$. It is suitable for our purpose but, in view of Prohorov (1953), Kolmogorov (1956) and Deheuvels and Pfeifer (1986), we believe this condition could be weakened to "$\max_{jk} p_{njk} \to 0$, for $j \in C$ and $k \in C^c$, as $n \to \infty$".

Condition ($\mathbf{C}$−iv) implies that, for j, $k \in C$, F_{jk} converges to $\delta(0)$, the degenerate distribution at the origin and the convergence rate is of order $o(1/n)$. We denote by χ_0 a random variable having distribution $\boldsymbol{\rho}$.

In the sequel, we refer to ($\mathbf{C}$−i)−($\mathbf{C}$−iv) as the limit conditions. The boldfaced "$\mathbf{0}$" sometimes refers to a matrix of zeros and sometimes denotes a vector of zeros, but "$\mathbf{1}$" always denotes the unit vector and "$\mathbf{I}$" always denotes the identity matrix.

Parallel to $\mathbf{P}_n(t)$, define

$$\mathbf{A}(t) = \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{T}(t) & \mathbf{Q}(t) \end{pmatrix}, \tag{10}$$

where $\mathbf{T}(t) = (t_{jk}\varphi_{jk}(t))$ and $\mathbf{Q}(t) = (q_{jk}\varphi_{jk}(t))$.

Evidently under the limit conditions, we have, for each $t \in R$,

$$\mathbf{P}_n(t) \to \mathbf{A}(t), \quad n \to \infty. \tag{11}$$

3.2. Some lemmas

We first state the well-known Perron-Frobenius theorem for irreducible nonnegative square matrices. (See Cox and Miller (1965), Graham (1987) and Iosifescu (1980). We give it in the form adequate for our purpose, not in its full generality.) For two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimension, we write $\mathbf{A} \geq \mathbf{B}$ if every element of $\mathbf{A} - \mathbf{B}$ is nonnegative.

Theorem 1. (Perron-Frobenius Theorem)

(a) *Let* $\mathbf{A} \geq 0$ *be an irreducible primitive finite square matrix, and let* α_1 *be its maximal positive eigenvalue. Then* α_1 *is simple and there exist strictly positive left and right eigenvectors corresponding to it. In particular, if* $\mathbf{A}$ *is stochastic, then* $\alpha_1 = 1$ *is the simple maximal eigenvalue.*

(b) *Let* $\mathbf{B} =(b_{jk})$ *be a complex-valued matrix of the same dimension as* $\mathbf{A}$*, and let* $\mathbf{B}^* = (|b_{jk}|)$($\mathbf{B}$ *in modulus*)*. Let* β *denote the maximal* (*in modulus*) *eigenvalues of* $\mathbf{B}$*. If* $\mathbf{B}^* \leq \mathbf{A}$*, then* $|\beta| \leq \alpha_1$*. Moreover, if* $|\beta| = \alpha_1$ *and* $\mathbf{B}^* \leq \mathbf{A}$*, then* $\mathbf{B}^* = \mathbf{A}$*.*

Lemma 2. *For all real* t*,* $\mathbf{A}(t)$ *has a simple maximal eigenvalue* $\alpha_0(t) = 1$ *such that if* $\alpha(t)$ *is any other eigenvalue of* $\mathbf{A}(t)$*, then* $|\alpha(t)| < \alpha_0(t) = 1$*.*

Proof. By (10), to find the eigenvalues of $\mathbf{A}(t)$ it is sufficient to find those of $\mathbf{R}$ and $\mathbf{Q}(t)$. The stochastic matrix $\mathbf{R}$ is primitive and irreducible, therefore it has 1 as the simple maximal eigenvalue. The row sums of the matrix $\mathbf{Q}$ are all less than 1, therefore all of its eigenvalues are strictly less than 1. Furthermore, $\mathbf{Q}^*(t) \leq \mathbf{Q}$, ($\mathbf{Q}(t)$ in modulus) and hence any eigenvalue of $\mathbf{Q}(t)$ is strictly less than 1 in modulus.

Lemma 3. *Let* $\mathbf{\Pi} = (\pi_0, \ldots, \pi_c)$ *be the stationary distribution of* $\mathbf{R}$*, i.e.,* $\mathbf{\Pi}$ *satisfies* $\mathbf{\Pi R} = \mathbf{\Pi}$ *with* $\sum_{j=0}^{c} \pi_j = 1$ *and* $\pi_j \geq 0$*. Under the limit conditions, we have*

$$\lim_{n\to\infty} \mathbf{A}^n(t) = \mathbf{x}(t)\mathbf{y}(t), \tag{12}$$

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{1} \\ (\mathbf{I} - \mathbf{Q}(t))^{-1}\, \mathbf{T}(t)\mathbf{1} \end{pmatrix}, \tag{13}$$

and $\mathbf{y}(t) = (\mathbf{\Pi}, \mathbf{0})$ *are right* (*column*) *and left* (*row*) *eigenvectors, respectively, of the square matrix* $\mathbf{A}(t)$ *corresponding to the eigenvalue* $\alpha_0(t) = 1$*.*

Proof. It is readily seen that $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are the right (column) and left (row) eigenvectors of $\mathbf{A}(t)$ corresponding to the eigenvalue $\alpha_0(t) = 1$.

Since $\alpha_0(t) = 1$ is the simple maximal eigenvalue, the matrix $\mathbf{A}(t)$ can be written in Jordan canonical form

$$\mathbf{A}(t) = \mathbf{U}(t) \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(t) \end{pmatrix} \mathbf{U}^{-1}(t),$$

where $\mathbf{J}(t)$ is a Jordan matrix whose diagonal elements are all less than 1 in modulus, $\mathbf{x}(t)$ is the first column of $\mathbf{U}(t)$ and $\mathbf{y}(t)$ is the first row of $\mathbf{U}^{-1}(t)$.

Since

$$\mathbf{A}^n(t) = \mathbf{U}(t) \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}^n(t) \end{pmatrix} \mathbf{U}^{-1}(t)$$

and $\lim_{n\to\infty} \mathbf{J}^n(t) = \mathbf{0}$, we have

$$\lim_{n\to\infty} \mathbf{A}^n(t) = \mathbf{U}(t) \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^{-1}(t) = \mathbf{x}(t)\mathbf{y}(t).$$

Lemma 4. *Under the limit conditions, for each $t \in R$ there exists an eigenvalue $\beta_{n0}(t)$ of $\mathbf{P}_n(t)$ such that $\beta_{n0}(t) \to 1$ and*

$$\lim_{n\to\infty} \mathbf{P}_n^n(t) = \lim_{n\to\infty} \beta_{n0}^n(t)\mathbf{x}(t)\mathbf{y}(t), \quad t \in R, \tag{14}$$

where $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are as defined in Lemma 3.

Proof. The limit conditions lead to (11) which in turn implies that the eigenvalues $\beta_{ni}(t)$ of $\mathbf{P}_n(t)$ converge to those of $\mathbf{A}(t)$. Therefore for large n, we can take $\beta_{n0}(t)$ to be the simple maximal eigenvalue of $\mathbf{P}_n(t)$ such that $|\beta_{ni}(t)| < |\beta_{n0}(t)| \le 1$, for $i = 1, \dots, s$. (The last inequality follows from the fact that $\mathbf{P}_n^*(t) \le \mathbf{P}_n(0) = \mathbf{P}_n$ and $\mathbf{P}_n$ has 1 as the simple maximal eigenvalue.) Hence, in Jordan canonical form,

$$\mathbf{P}_n(t) = \mathbf{H}_n(t) \begin{pmatrix} \beta_{n0}(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_n(t) \end{pmatrix} \mathbf{H}_n^{-1}(t), \tag{15}$$

with $\lim_{n\to\infty} \mathbf{J}_n^n(t) \to \mathbf{0}$, the first column of $\mathbf{H}_n(t)$ is the right eigenvector of $\mathbf{P}_n(t)$ corresponding to $\beta_{n0}(t)$ which converges to $\mathbf{x}(t)$, and the first row of $\mathbf{H}_n^{-1}(t)$ is the left eigenvector of $\mathbf{P}_n(t)$ corresponding to $\beta_{n0}(t)$ which converges to $\mathbf{y}(t)$.

The limit (14) thus follows from the Jordan canonical form (15).

Lemma 5. *For an irreducible primitive stochastic matrix $\mathbf{R}$ with eigenvalues $\{1, \alpha_1, \dots, \alpha_c\}$, we have*

$$\frac{adj(\mathbf{R} - \mathbf{I})}{\Pi_{j=1}^c (\alpha_j - 1)} = \mathbf{1\Pi},$$

where $\mathbf{\Pi}$ is the stationary distribution of $\mathbf{R}$ and $adj(\mathbf{R} - \mathbf{I})$ denotes the adjoint of the square matrix $\mathbf{R} - \mathbf{I}$.

Proof. This lemma follows from Theorem 1.8 in Iosifescu (1980), by the fact that 1 is the simple maximal eigenvalue of $\mathbf{R}$ such that $\mathbf{1}$ and $\mathbf{\Pi}$ are the right (column) and left (row) eigenvectors corresponding to this eigenvalue.

Lemma 6. *Under the limit conditions,*

$$\lim_{n\to\infty} \beta_{n0}^n(t) = G(t), \quad t \in R, \tag{16}$$

where $G(t)$ is defined by

$$\ln G(t) = -\sum_{k=c+1}^{s} \sum_{j=0}^{c} \lambda_{jk}\pi_j \left(1 - \varphi_{jk}(t)x_k(t)\right), \tag{17}$$

and $(\mathbf{I}-\mathbf{Q}(t))^{-1}\mathbf{T}(t)\mathbf{1} = (x_{c+1}(t),\dots,x_s(t))^T$ *as in* (13).

Proof. Denote by $\{\beta_{n0}(t),\dots,\beta_{ns}(t)\}$ the eigenvalues of $\mathbf{P}_n(t)$, $\{1,\alpha_1,\dots,\alpha_c\}$ those of $\mathbf{R}$ and $\{\alpha_{c+1}(t),\dots,\alpha_s(t)\}$ those of $\mathbf{Q}(t)$. Then, by (11),

$$\begin{aligned}\lim_{n\to\infty}\det(\mathbf{P}_n(t)-z\mathbf{I}) &= \lim_{n\to\infty}\Pi_{j=0}^{s}(\beta_{nj}(t)-z)\\ &= (1-z)\Pi_{j=1}^{c}(\alpha_j-z)\det(\mathbf{Q}(t)-z\mathbf{I}). \qquad (18)\end{aligned}$$

With $\det(\mathbf{P}_n(t)-z\mathbf{I}) = \Pi_{j=0}^{s}(\beta_{nj}(t)-z)$ and by letting $z=1$, we can write

$$\beta_{n0}(t) = 1+\frac{\det(\mathbf{P}_n(t)-\mathbf{I})}{\Pi_{j=1}^{s}(\beta_{nj}(t)-1)} = 1+\frac{C_n(t)}{n},$$

$$C_n(t) = \frac{n\det(\mathbf{P}_n(t)-\mathbf{I})}{\Pi_{j=1}^{s}(\beta_{nj}(t)-1)}. \qquad (19)$$

On the other hand, noting that, by (18), $\lim_{n\to\infty}\Pi_{j=0}^{s}(\beta_{nj}(t)-z) = (1-z)\lim_{n\to\infty}\Pi_{j=1}^{s}(\beta_{nj}(t)-z)$, and letting $z=1$, we have

$$\lim_{n\to\infty}\Pi_{j=1}^{s}(\beta_{nj}(t)-1) = \Pi_{j=1}^{c}(\alpha_j-1)\det(\mathbf{Q}(t)-\mathbf{I}). \qquad (20)$$

We write $\widetilde{p}_{njk}(t)$ for $p_{njk}\phi_{njk}(t)$. Denote

$$\mathbf{P}_n(t) = (\widetilde{p}_{nij}(t)) = \begin{pmatrix}\mathbf{P}_{n11}(t) & \mathbf{P}_{n12}(t)\\ \mathbf{P}_{n21}(t) & \mathbf{P}_{n22}(t)\end{pmatrix},$$

where $\mathbf{P}_{n11}(t)$ and $\mathbf{P}_{n22}(t)$ are of dimension $(c+1)\times(c+1)$ and $(s-c)\times(s-c)$, respectively.

By adding columns $2,\dots,c+1$ to the first one of the matrix $\mathbf{P}_n(t)-\mathbf{I}$, we have

$$\det(\mathbf{P}_n(t)-\mathbf{I}) = \det\begin{pmatrix}\mathbf{B}_{n11}(t) & \mathbf{P}_{n12}(t)\\ \mathbf{B}_{n21}(t) & \mathbf{P}_{n22}(t)-\mathbf{I}\end{pmatrix}, \qquad (21)$$

where each column of $\mathbf{B}_{n11}(t)$ is one from $\mathbf{P}_{n11}(t)-\mathbf{I}$ except the first one, which is $(\sum_{k=0}^{c}\widetilde{p}_{njk}(t)-1)^T_{j=0,\dots,c}$. Similarly, each column of $\mathbf{B}_{n21}(t)$ is one from $\mathbf{P}_{n21}(t)$ except the first one which is $(\sum_{k=0}^{c}\widetilde{p}_{njk}(t))^T_{j=c+1,\dots,s}$.

(Note that in obtaining the form in (21), we add to the first the rest of the columns in $\mathbf{P}_n(t)-\mathbf{I}$. We could have chosen *any* one among the first $(c+1)$ columns in $\mathbf{P}_n(t)-\mathbf{I}$ and add to that one the rest of the columns and the proof of this lemma will not be affected.)

By $\lim_{n\to\infty}\det(\mathbf{P}_{n22}(t)-\mathbf{I})/\det(\mathbf{Q}(t)-\mathbf{I}) = 1$, combining (18), (19) and (20) yields

$$\lim_{n\to\infty}C_n(t) = \lim_{n\to\infty}\frac{n\det(\mathbf{B}_{n11}(t)-\mathbf{P}_{n12}(t)\mathbf{D}_n(t))}{\Pi_{j=1}^{c}(\alpha_j-1)}, \qquad (22)$$

where $\mathbf{D}_n(t) = (d_{nij}(t)) = (\mathbf{P}_{n22}(t) - \mathbf{I})^{-1}\mathbf{B}_{n21}(t)$. (Note that $\mathbf{D}_n(t)$ is a $(s-c) \times (c+1)$ matrix.) Therefore, to prove (16) we need to find the limit of the numerator in (22).

We first note that the first column of $\mathbf{D}_n(t)$ converges to $-(\mathbf{I} - \mathbf{Q}(t))^{-1}\mathbf{T}(t)\mathbf{1} = (-x_{c+1}(t), \ldots, -x_s(t))^T$ of $\mathbf{x}(t)$ as in (13). And the jth element, $j \in C$, of the first column of the matrix $\mathbf{B}_{n11}(t) - \mathbf{P}_{n12}(t)\mathbf{D}_n(t)$ in the numerator of (22) is

$$\begin{aligned} &\sum_{k=0}^{c} \tilde{p}_{njk}(t) - 1 - \sum_{k=c+1}^{s} \tilde{p}_{njk}(t) d_{nk1}(t) \\ &= \sum_{k=0}^{c} p_{njk}(\phi_{njk}(t) - 1) - \sum_{k=c+1}^{s} p_{njk}(1 + \phi_{njk}(t) d_{nk1}(t)). \end{aligned} \tag{23}$$

Multiplying (23) by n and letting $n \to \infty$, from the limit conditions, we get the limit $-\sum_{k=c+1}^{s} \lambda_{jk}(1 - \varphi_{jk}(t)x_k(t))$.

On the other hand, the jth element, $j \in C$, of the kth column $(1 \le k \le c)$ of the matrix $\mathbf{B}_{n11}(t) - \mathbf{P}_{n12}(t)\mathbf{D}_n(t)$ is

$$\tilde{p}_{njk}(t) - \delta_{jk} - \sum_{\ell=c+1}^{s} \tilde{p}_{nj\ell}(t) d_{n(\ell-c)j}(t). \tag{24}$$

Since $p_{njk} \to r_{jk}$ and $\phi_{njk}(t) \to 1$ for $j,\ k \in C$, and $p_{nj\ell} \to 0$ for $\ell \in C^c$, (24) converges to r_{jk} for all $j \in C$.

Thus

$$\lim_{n\to\infty} C_n(t) = \frac{-\det \begin{pmatrix} \sum_{j=c+1}^{s} \lambda_{0j}\,(1 - \varphi_{0j}(t)x_j(t))\,, r_{01}, & \ldots, r_{0c} \\ \sum_{j=c+1}^{s} \lambda_{1j}\,(1 - \varphi_{1j}(t)x_j(t))\,, r_{11} - 1, & \ldots, r_{1c} \\ \vdots \qquad\qquad\qquad \vdots & \vdots \quad \vdots \\ \sum_{j=c+1}^{s} \lambda_{cj}\,(1 - \varphi_{cj}(t)x_j(t))\,, r_{c1}, & \ldots, r_{cc} - 1 \end{pmatrix}}{\Pi_{j=1}^{c}\,(\alpha_j - 1)}.$$

By Lemma 5 and expansion of this determinant by the first column, we complete the proof of Lemma 6.

3.3. The main theorem

We are now ready to state and prove our main theorem.

Theorem 7. *Under the limit conditions, the limiting ch.f. of $Y_{nn} = W_{n1} + \cdots + W_{nn}$ is*

$$\lim_{n\to\infty} G_{nn}(t) = \boldsymbol{\rho}\mathbf{x}(t)G(t), \quad t \in R, \tag{25}$$

where $G(t)$ is defined by (17), *$\boldsymbol{\rho} = (\rho_0, \ldots, \rho_s)$ is the limit of the initial probability $\mathbf{p}_n$, and $\mathbf{x}(t)$ is the column vector defined by* (13).

Proof. By (4), Lemma 4 and Lemma 6, we have $\lim_{n\to\infty} G_{nn}(t) = \lim_{n\to\infty} \mathbf{p}_n \times \mathbf{P}_n^n(t)\mathbf{1} = \lim_{n\to\infty} \beta_{n0}^n(t)\mathbf{p}_n\mathbf{x}(t)\mathbf{y}(t)\mathbf{1} = G(t)\boldsymbol{\rho}\mathbf{x}(t)\mathbf{y}(t)\mathbf{1}$. Since $\mathbf{y}(t)= (\mathbf{\Pi}, \mathbf{0})$, $\mathbf{y}(t)\mathbf{1} = 1$ for all $t \in R$, the proof of the theorem is complete.

We note here that by taking $\Phi(t) = \sum_{k=c+1}^{s} \sum_{j=0}^{c} (\lambda_{jk}\pi_j/\lambda)\varphi_{jk}(t)x_k(t)$, we can write the ch.f. G as $G(t) = \exp\{-\lambda(1 - \Phi(t))\}$. ($\lambda > 0$ is the value which makes $\Phi(0) = 1$.) So G is the ch.f. of the compound Poisson distribution with parameter λ and compounding distribution defined by the ch.f. Φ. The limit distribution of Y_{nn} is the convolution of two independent random variables, one of them being compound Poisson.

If the long-run initial probability $\boldsymbol{\rho}$ is degenerate to the ergodic subspace C, that is $\mathbf{p}_n \to \boldsymbol{\rho} = (\rho_0, \ldots, \rho_c, 0, \ldots, 0)$, then, $\boldsymbol{\rho}\mathbf{x}(t) = 1$ for all $t \in R$, and the limiting ch.f. of Y_{nn} in (25) is simply $\lim_{n\to\infty} G_{nn}(t) = G(t)$, $t \in R$.

A special case of interest is the partial sum $S_{nn} = X_{n0} + \cdots + X_{nn}$, with X_{n0} not degenerate at 0. By using the ch.f. $G_{nn}(t)$ as defined by (9) and $\mathbf{P}_n(t)$ as by (7), we can prove the next corollary in an identical manner.

Corollary 8. *Under the limit conditions (with $C = \{0\}$) and in the case that the Markov sequence $\{X_n\}$ is not completely stationary, the limiting ch.f. of S_{nn} is*

$$\lim_{n\to\infty} E\left(e^{it(X_{n0}+\cdots+X_{nn})}\right) = \boldsymbol{\rho}(t)\mathbf{x}(t)G(t), \quad t \in R, \tag{26}$$

where $\boldsymbol{\rho}(t) = (\rho_0, \rho_1 e^{it}, \ldots, \rho_s e^{ist}), \mathbf{x}(t)$ is defined by (13) and the $\varphi_{jk}(t)$ in the entries in $G(t)$ are $\varphi_{jk}(t) = e^{ikt}$.

In case the Markov chain $\{X_n\}$ is completely stationary with transition matrix defined by (32) below, the limiting ch.f. of S_{nn} is of form (33).

4. Applications

We illustrate the applications of Theorem 7 and Corollary 8 with some examples.

In the two-state case, if we take $c = 0$ and $C = \{0\}$, with $p_{n1} \to \rho$, $np_{n01} \to \lambda$, $p_{n10} \to \tau$ and $\mathbf{x}(t) = (1, \frac{\tau\varphi_{10}(t)}{1-(1-\tau)\varphi_{11}(t)})^T$, (25) reduces to

$$\boldsymbol{\rho}(t)\mathbf{x}(t)G(t) = \Big(1 - \rho + \frac{\rho\tau\varphi_{10}(t)}{1-(1-\tau)\varphi_{11}(t)}\Big) \exp\Big\{-\lambda\Big(1 - \frac{\varphi_{01}(t)\tau\varphi_{10}(t)}{1-(1-\tau)\varphi_{11}(t)}\Big)\Big\}. \tag{27}$$

Such ch.f.'s as $G(t) = \exp\{-\lambda(1-e^{it})\}$ and $G(t) = \exp\{-\lambda(1-\varphi(t))\}$, where $\varphi(t)$ is a ch.f., can be deduced from (27).

By Corollary 8, the limiting ch.f. of S_{nn} is

$$\Big(1 - \rho + \frac{\varphi_{01}(t)\rho\tau\varphi_{10}(t)}{1-(1-\tau)\varphi_{11}(t)}\Big) \exp\Big\{-\lambda\Big(1 - \frac{\varphi_{01}(t)\tau\varphi_{10}(t)}{1-(1-\tau)\varphi_{11}(t)}\Big)\Big\}, \tag{28}$$

which is slightly different from (27). (Here $\varphi_{01}(t) = e^{it} = \varphi_{11}(t)$, $\varphi_{10}(t) = 1$.)

For the completely stationary Markov Bernoulli sequence in Edwards (1960), with initial probability $(1 - p_n, p_n)$, $p_{n01} = (1 - \pi)p_n$ and $p_{n10} = (1 - \pi)(1 - p_n)$, where $0 \le \pi < 1$ is a fixed constant. Under the simple limit condition "$np_n \to \lambda > 0$", the ch.f. of S_{nn} is

$$\exp\Big\{ -(1-\pi)\lambda\Big(1 - \frac{(1-\pi)e^{it}}{1-\pi e^{it}}\Big)\Big\}. \tag{29}$$

This result is the main theorem in Wang (1981) and Gani (1982).

Consider the general Markov Bernoulli with initial probability $(1 - p_n, p_n)$. If $p_n \to \rho$, $np_{n01} \to \lambda > 0$, $p_{n10} \to \tau$ $(0 < \tau, \rho < 1)$, then by (28) and $\varphi_{j0}(t) = 1$, $\varphi_{j1}(t) = e^{it}$, for $j = 0$, 1, the limiting ch.f. of S_{nn} is

$$\Big(1 - \rho + \frac{\rho\tau e^{it}}{1-(1-\tau)e^{it}}\Big) \exp\Big\{ -\lambda\Big(1 - \frac{\tau e^{it}}{1-(1-\tau)e^{it}}\Big)\Big\}.$$

This result is Theorem 2.1 in Lin and Wang (1994).

Denote by Y_{nn} the number of runs of type "$1 \to 1$" occurring in n transitions. The ch.f.'s φ are the same as in the previous case except $\varphi_{01}(t) = 1$, instead of e^{it}. The limiting distribution of Y_{nn}, by (28), is

$$\Big(1 - \rho + \frac{\rho\tau}{1-(1-\tau)e^{it}}\Big) \exp\Big\{ -\lambda\Big(1 - \frac{\tau}{1-(1-\tau)e^{it}}\Big)\Big\}.$$

The following are some known results in the literature for the general finite-state case.

First let us consider the simplest case $-$ the i.i.d. multi-state sequence. In a Markovian context, the (nth) i.i.d. sequence $\{X_{ni}\}$ has transition probabilities $p_{njk} = \Pr(X_{nj} = k) = \Pr(X_{n(j+1)} = k \mid X_{nj} = i) = p^*_{nk} \ge 0$, with $\sum_{k=0}^{s} p^*_{nk} = 1$, and initial probability $\mathbf{p}_n$=$(p_{n0}, \ldots, p_{ns})$. The simple limit conditions are "$np^*_{nk} \to \lambda_{0k} = \lambda_k > 0$, for $k = 1, \ldots, s$" which means $\boldsymbol{\rho} = (1, 0, \ldots, 0)$, $\mathbf{R}$ =$\{1\}$, $\mathbf{T} = \mathbf{1}$ and $\mathbf{Q} = \mathbf{0}$. Consider S_{nn}, for which $\varphi_{j0}(t) = 1$, $\varphi_{jk}(t) = e^{ikt}$ $(k = 1, \ldots, s$ and $j = 0, \ldots, s)$ and $\mathbf{x}(t) = \mathbf{1}$. By (26), the limiting ch.f. of S_{nn} is

$$G(t) = \exp\Big\{ -\sum_{k=1}^{s} \lambda_k(1 - e^{ikt})\Big\} = \prod_{k=1}^{s} \exp\Big\{ -\lambda_k(1 - e^{ikt})\Big\}, \tag{30}$$

which is the ch.f. of the sum $W_1 + 2W_2 + \cdots + sW_s$, where the W_k are independent Poisson random variables with parameters λ_k.

We can rewrite (30) as $G(t) = \exp\{-\lambda(1 - \Phi(t))\}$, where $\lambda = \sum_{k=1}^{s} \lambda_k$ and $\Phi(t) = \sum_{k=1}^{s} (\lambda_k/\lambda)e^{ikt}$, which is the ch.f. of a multi-state random variable Z with $\Pr(Z = k) = \lambda_k/\lambda$, $k = 1, \ldots, s$. Thus the limiting distribution of S_{nn} is a compound Poisson distribution, not obvious in the form (30). (See Wang and Ji (1993) for some other related results.)

Actually, this is a rather straight-forward case. The distribution of S_{nn} has the ch.f.

$$g(t) = \Big(p_0 + \sum_{k=1}^{s} p_k e^{ikt}\Big)^{n+1}. \tag{31}$$

It is easy to see that, under the limit conditions, (31) converges to (30). (We remark that the distribution corresponding to (31) is sometimes known as "univariate multinomial distribution". (See Steyn (1951).)

Edwards' completely stationary Markov Bernoulli model, stated earlier, was extended to a general finite-state case by Wang and Yang (1995). The transition matrix $\mathbf{P}_n$=(p_{njk}) for the general case is

$$p_{njk} = \pi\delta_{jk} + (1-\pi)p_{nk}, \quad j,k \in S, \tag{32}$$

where δ_{jk} is the Kronecker delta and $\mathbf{p}_n$= $(p_{n0},\dots,p_{ns})$ is a probability distribution. If we take $\mathbf{p}_n$ as the initial probability, then we have $\mathbf{p}_n\mathbf{P}_n$= $\mathbf{p}_n$, and the Markov chain becomes completely stationary. (See their paper for detailed discussions about many properties of this model. In that paper, Edwards' model was extended in two directions, one in the multivariate direction and the other as above.)

The limit conditions are, as in the previous case, "$np_{nk} \to \lambda_k > 0$, for $k = 1,\dots,s$" which leads to $np_{n0k} \to \lambda_{0k} = (1-\pi)\lambda_k$, $\mathbf{T}$ =$(1-\pi,\dots,1-\pi)^T$ and $\mathbf{Q}$ an $s \times s$ diagonal matrix with identical diagonal entry π. Consider S_{nn}, so $\varphi_{jk}(t) = e^{ikt}$, for $j,\ k = 0,\dots,s$ and $x_k(t) = (1-\pi)e^{ikt}/(1-\pi e^{ikt})$, for $k = 1,\dots,s$. Putting these together, we get

$$G(t) = \exp\Big\{-\sum_{k=1}^{s}(1-\pi)\lambda_k\Big(1 - e^{ikt}\frac{(1-\pi)e^{ikt}}{1-\pi e^{ikt}}\Big)\Big\}. \tag{33}$$

In p.g.f. form, (33) is found on p.48 of Wang and Yang (1995).

Finally, we consider a reliability system known as "consecutive k-out-of-n: F system", which consists of n linearly ordered components. The component failure times are assumed to be i.i.d. and the system fails if and only if at least k out of the n components fail. It is known (see Chao and Fu (1989)) that such a system forms a Markov chain with $k+1$ states: 0 represents a system in perfect condition, 1 to $k-1$ are levels of deterioration and the state k indicates that the system has failed. The event $\{X_n = j\}$ signifies that at time n the system is in the state j. Two cases arise when the system is in state k. (1) There is a stand-by system which automatically replaces the failed system at the next time period after the kth failure occurred; then the state k is a reflexive state and

$\Pr(X_{n+1} = 0 \mid X_n = k) = 1$ for all n. (2) There is no stand-by system and the failed system can not be repaired; the state k is then an absorbing state with $\Pr(X_{n+1} = k \mid X_n = k) = 1$ for all n. In the first case the transition matrix is

$$\mathbf{P} = \begin{pmatrix} 1-\alpha & \alpha & 0 \cdots 0 \\ 1-\alpha & 0 & \alpha \cdots 0 \\ \vdots & \vdots & \vdots \; \vdots \; \vdots \\ 1-\alpha & 0 & 0 \cdots \alpha \\ 1 & 0 & 0 \cdots 0 \end{pmatrix}, \tag{34}$$

take initial probabilities $\Pr(X_1 = j) = p_j$, $j = 0, \ldots, k$, $\sum_{j=0}^{k} p_j = 1$. It turns out that the second case, whose transition matrix is as (34) except the 1 in the last row appears at the very end of the row, has the same limiting distribution for S_n as in the first case. (See Chao and Fu (1989).) (For ease of notation, we have suppressed the dependence on the subscript n, the nth row of a triangular array.)

We use Corollary 8 to find the limit distribution of S_n in the first case. Here $c = 0$, $s = k$, $\mathbf{Q} = \mathbf{0}_{s\times s}$, $\mathbf{T} = \mathbf{x}(t) = (1, 1, \ldots, 1)^T$, $\lim_{n\to\infty} n\alpha = \lambda_{01} = \lambda > 0$, $\lambda_{0j} = 0$ for $j = 1, \ldots, s$, $\boldsymbol{\rho}(t) = (\rho_0, \rho_1 e^{it}, \ldots, \rho_s e^{ist})$, $\varphi_{j0}(t) = 1$ and $\varphi_{jk}(t) = e^{ikt}$ for all $j \geq 0$, $k \geq 1$. The limiting ch.f. is

$$(\rho_0 + \rho_1 e^{it} + \cdots + \rho_s e^{ist}) \exp\left\{-\lambda \left(1 - e^{it}\right)\right\}. \tag{35}$$

The distribution can be identified as that of the sum of two independent random variables Z and Y, where $\Pr(Z = j) = \rho_j$, for $j = 0, \ldots, s$, and Y is Poisson with parameter λ. The ch.f. (35), in p.g.f. form, is found in Theorem 3.1 in Lin and Wang (1994).

In all the above examples $C = \{0\}$. As a demonstration to show that the ergodic set does not have to be the singleton set $\{0\}$, let us consider a case with $c = 1$, $s = 3$ and $C = \{0, 1\}$. All $\mathbf{R}$, $\mathbf{T}$, and $\mathbf{Q}$ are 2×2 matrices with $\mathbf{R} = (r_{jk})$ satisfying $0 < r_{jk} < 1$, and $r_{j0} + r_{j1} = 1$ for j, $k = 0$, 1. The stationary distribution $\mathbf{\Pi} = (\pi_0, \pi_1)$ of $\mathbf{R}$ is easily seen to be $\pi_0 = r_{10}/(r_{10} + r_{01})$, and $\pi_1 = r_{01}/(r_{10} + r_{01})$, so that $\mathbf{y}(t) = (\pi_0, \pi_1, 0, 0)$ and $\mathbf{x}(t) = (1, 1, x_2(t), x_3(t))^T$, where

$$\begin{pmatrix} x_2(t) \\ x_3(t) \end{pmatrix} = (\mathbf{I} - \mathbf{Q}(t))^{-1} \, \mathbf{T}(t)\mathbf{1}.$$

The four parameters $\lambda_{jk} = \lim_{n\to\infty} np_{jk} \geq 0$, for $j = 0$, 1 and $k = 2$, 3.

The limiting ch.f. $G(t)$ is

$$G(t) = (\rho_0 + \rho_1 + \rho_2 x_2(t) + \rho_3 x_3(t)) \exp\left\{-\lambda \left(1 - \Phi(t)\right)\right\}, \tag{36}$$

where $\Phi(t) = \sum_{k=2}^{3} \sum_{j=0}^{1} (\lambda_{jk}\pi_j/\lambda)\varphi_{jk}(t)x_k(t)$, with $\lambda = \sum_{k=2}^{3} \sum_{j=0}^{1} \lambda_{jk}\pi_j$.

We consider further two instances of the sequence $\{Y_n\}$.

(a) Let Y_n be the number of times "runs" of type "$2 \to 2$" or "$3 \to 3$" occur in n transitions. Then $\varphi_{jk}(t) = 1$, for all j, $k = 0, 1, 2, 3$, except $\varphi_{22}(t) = \varphi_{33}(t) = e^{it}$. To obtain a manageable expression for G, we take

$$\mathbf{T} = \begin{pmatrix} 1-\alpha & 0 \\ 0 & 1-\beta \end{pmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \quad 0 < \alpha,\ \beta < 1.$$

Then $x_2(t) = (1-\alpha)/(1-\alpha e^{it})$ and $x_3(t) = (1-\beta)/(1-\beta e^{it})$. And

$$G(t) = \left(\rho_0 + \rho_1 + \rho_2 \frac{(1-\alpha)}{1-\alpha e^{it}} + \rho_3 \frac{(1-\beta)}{1-\beta e^{it}}\right) \exp\left\{-\upsilon_2\left(1 - \frac{(1-\alpha)}{1-\alpha e^{it}}\right) - \upsilon_3\left(1 - \frac{(1-\beta)}{1-\beta e^{it}}\right)\right\}, \tag{37}$$

where $\upsilon_2 = \lambda_{02}\pi_0 + \lambda_{12}\pi_1$ and $\upsilon_3 = \lambda_{03}\pi_0 + \lambda_{13}\pi_1$. The distribution corresponding to (37) is that of $Z + U_1 + U_2$, where Z, U_1 and U_2 are independent with

$$\begin{cases} P(Z=0) = \rho_0 + \rho_1 \\ P(Z=k) = \rho_2 \alpha (1-\alpha)^k + \rho_3 \beta (1-\beta)^k, \quad k = 1, 2, \dots , \end{cases}$$

while U_1 and U_2 have compound Poisson distributions with parameters υ_2 and υ_3 and geometric compounding distributions $\{\alpha(1-\alpha)^k;\ k = 0, 1, \dots,\}$ and $\{\beta(1-\beta)^k;\ k = 0,\ 1, \dots,\}$, respectively.

(b) Let Y_n be the number of times the chain visits the states $\{2,\ 3\}$ in n transitions, so that $\varphi_{j2}(t) = \varphi_{j3}(t) = e^{it}$ and $\varphi_{j0}(t) = \varphi_{j1}(t) = 1$, for all $j = 0$, 1, 2, 3. For simplicity, we let the limit $\boldsymbol{\rho} = (\rho_0, \rho_1, 0, 0)$ and

$$\mathbf{T} = \begin{pmatrix} 1-\alpha & 0 \\ 0 & 1-\alpha \end{pmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{pmatrix} \alpha/2 & \alpha/2 \\ \alpha/2 & \alpha/2 \end{pmatrix}, \quad 0 < \alpha < 1.$$

It can be shown that $(\mathbf{I} - \mathbf{Q}(t))^{-1}\mathbf{T}(t)\mathbf{1} = ((1-\alpha)^2, (1-\alpha)^2)^T$, so that

$$\sum_{k=2}^{3}\sum_{j=0}^{1} \lambda_{jk}\pi_j\varphi_{jk}(t)x_k(t) = \{(1-\alpha)^2 \sum_{k=2}^{3}\sum_{j=0}^{1} \lambda_{jk}\pi_j\}e^{it}.$$

Letting $\lambda = (1-\alpha)^2 \sum_{k=2}^{3}\sum_{j=0}^{1} \lambda_{jk}\pi_j$, we obtain $G(t) = \exp(-\lambda(1-e^{it}))$, which is the ch.f. of the Poisson distribution with parameter λ.

References

Chao, M. T. and Fu, J. C. (1989). A limit theorem of certain repairable systems. *Ann. Inst. Statist. Math.* **41**, 809-818.

Cox, D. R. and Miller, H. D. (1965). *The Theory of Stochastic Processes.* Chapman and Hall, London.

Deheuvels, P. and Pfeifer, D. (1986). A semi-group approach to Poisson approximation. *Ann. Probab.* **14**, 663-676.

Edwards, A. W. F. (1960). The meaning of binomial distribution. *Nature* **186**, 1074.

Gabriel, K. R. (1959). The distribution of the number of successes in a sequence of dependent trials. *Biometrika* **46**, 454-460.

Gani, J. (1982). On the probability generating function of the sum of Markov Bernoulli random variables. *J. Appl. Probab.* **19A**, 321-326.

Graham, A. (1987). *Non-Negative Matrices and Applicable Topics in Linear Algebra.* Ellis Horwood Limited, Chichester.

Hsiau, S. R. (1997). Compound Poisson limit theorems for Markov chains. *J. Appl. Probab.* **34**, 24-34.

Iosifescu, M. (1980). *Finite Markov Processes and Their Applications.* Wiley, Chichester.

Kelson, J. and Wishart, D. M. G. (1964). A central limit theorem for processes defined on a finite Markov chain. *Proc. Camb. Phil. Soc.* **60**, 547-567.

Kolmogorov, A. N. (1956). Deux theoremes asymptotiques pour les sommes de variables aleatoires. (Russian). *Teoriia Veroiatnosteei* **1**, 426-436.

Koopman, B. O. (1950). A generalization of Poisson's distribution for Markov chains. *Proc. Nat. Acad. Sci.* **36**, 202-207.

Lin, Z. Y. and Wang, Y. H. (1994). Limit theorems on multi-state Markov chains. *Chinese Ann. Math.* **15A**, 511-517, (in Chinese). Also appeared in English in *Chinese J. Contemp. Math.* **15**, 307-316.

Mathews, J. P. (1970). A central limit theorem for absorbing Markov chains. *Biometrika* **57**, 129-139.

Miller, H. D. (1961). A convexity property in the theory of random variables defined on a finite Markov chain. *Ann. Math. Statist.* **32**, 1260-1270.

Miller, H. D. (1962a). A matrix factorization problem in the theory of random variables defined on a finite Markov chain. *Proc. Camb. Phil. Soc.* **58**, 268-285.

Miller, H. D. (1962b). Absorption probabilities for sums of random variables defined on a finite Markov chain. *Proc. Camb. Phil. Soc.* **58**, 286-298.

Pedler, J. (1971). Occupation times for two state Markov chain. *J. Appl. Probab.* **8**, 381-390.

Pedler, J. (1978). The occupation times, number of transitions, and waiting time for two-state Markov chains. Master's thesis, The Flinders University of South Australia.

Prohorov, Yu. V. (1953). Asymptotic behavior of the binomial distribution. (Russian). *Uspekhii Mate. Hauk.* **8**, 135-142.

Pyke, R. (1961). Markov renewal processes: definitions and preliminary properties. *Ann. Math. Statist.* **32**, 1231-1242.

Steyn, H. S. (1951). On discrete multivariate probability functions. *Proc. Kon. Nederl. Akad. Wetensch. Ser. A* **54**, 23-30.

Volkov, I. S. (1958). On the distribution of sums of random variables defined on a homogeneous Markov chain with a finite number of states. *Teor. Veroyat. i ee Prim.* **3**, 413-429. (in Russian).

Wang, Y. H. (1981). On the limit of the Markov binomial distribution. *J. Appl. Probab.* **18**, 937-942.

Wang, Y. H. and Liang, Z. Y. (1993). The probability of occurrences of runs of length k in n Markov Bernoulli trials. *Math. Sci.* **18**, 105-112.

Wang, Y. H. and Ji, S. X. (1993). Derivations of the compound Poisson distributions and processes. *Statist. Probab. Lett.* **18,** 1-7.

Wang, Y. H. and Yang, Z. J. (1995). On a Markov multinomial distribution. *Math. Sci.* **20**, 40-49.

Wang, Y. H. and Ji, S. X. (1995). Limit theorems for the number of occurrences of consecutive k successes in n Markov trials. *J. Appl. Probab.* **32**, 727-735.

Wang, Y. H. and Tang, L. Q. (1997). A completely stationary Markov chain with infinite state space. *Taiwanese J. Math.* **1**, 517-525.

Department of Statistics, Tunghai University, Taichung, Taiwan.

E-mail: yhwang@mail.thu.edu.tw

Health Services Research Center, UCLA Neuropsychiatric Institute, Los Angeles, U.S.A.

E-mail: lqtang@ucla.edu

(Received May 2000; accepted July 2002)