J. Appl. Prob. 17, 33–46 (1980)
Printed in Israel
0021-9002/80/010033-14 $01.65

TIME REVERSAL AND AGE DISTRIBUTIONS, I. DISCRETE-TIME MARKOV CHAINS

S. TAVARÉ,* *University of Sheffield*

Abstract

The connection between the age distribution of a discrete-time Markov chain and a certain time-reversed Markov chain is exhibited. A method for finding properties of age distributions follows simply from this approach. The results, which have application in several areas in applied probability, are illustrated by examples from population genetics.

MARKOV CHAINS; AGE DISTRIBUTION; DIFFUSION APPROXIMATION; GENETICS; TIME REVERSAL; DUALITY

1. Introduction

Several authors have considered statistical approaches to estimating the 'age' of a Markov process (Stigler (1970), Thompson (1976)). In this context, the 'age' is the absolute time origin of the process, and estimation is based on some 'current' observations. More recently, Levikson (1977) considered the problem of finding the age distribution of Markov processes with state space S; a set of absorbing states, C, and transient states $B = S - C$. His method is to restart the process whenever C is visited by forcing instantaneous return to the nearest point of B. Between such returns, the probabilistic description of the two processes is identical. Given 'current' value $j \in B$, the age G_j has distribution determined by the limiting distribution of the time that has elapsed since C was last visited.

We now consider discrete-time Markov chains $\{X_n\}$, with corresponding return processes $\{\bar{X}_n\}$. We make a minor modification to Levikson's restarting procedure. Instead of requiring instantaneous returns, we assume that it takes one step for a return to take place. This ensures that both processes have the same state space, which is a useful simplification (cf. Tavaré (1978a), Pakes (1978)). It is assumed that the return process is irreducible, aperiodic, and

Received 8 August 1978; revision received 13 December 1978.

* Present address: Department of Mathematics, University of Utah, Salt Lake City, UT 84112, U.S.A.

positive recurrent (as will usually be the case in genetic applications). The distribution of G_j is then given by

$$(1.1)\quad P(G_j = n) = g_j^{(n)} = \lim_{m\to\infty} P(\bar{X}_{m-n} \in C, X_{m-k} \notin C, 0 < k < n \mid \bar{X}_m = j).$$

Levikson used renewal theory methods to derive the distribution (1.1) in terms of the original chain, and some interesting properties of such an age. The purpose of this note is to give a simple method of deriving such results for an arbitrary restarting distribution. The idea is to relate the age distribution defined by (1.1) to a property of a certain time-reversed Markov chain. Such a representation considerably simplifies the analytical methods required to derive results about ages, and suggests some interesting problems in the theory of diffusion processes and weak convergence. Time-reversal has been used before in a genetic context to explain properties of asymptotic conditional distributions of the type

$$\lim_{m\to\infty} P(X_m = j \mid X_m \in B),$$

and

$$\lim_{m\to\infty}\lim_{n\to\infty} P(X_m = j, m < n \mid X_n \in B).$$

See, for example, Darroch and Seneta (1965), and Seneta (1966).

2. Two-barrier models

We suppose that $S = \{0, 1, \cdots, M\}$, and $C = \{0, M\}$. This is the natural description of a wide variety of models that arise in population genetics. We begin with some notation. Suppose $\{X_n\}$ has transition matrix $P = (p_{ij})$, which we write in the form

$$(2.1)\qquad P = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \boldsymbol{p}_0' & Q & \boldsymbol{p}_M' \\ 0 & \mathbf{0} & 1 \end{bmatrix}.$$

Here, $'$ denotes transpose. The fundamental matrix is denoted $N = (n_{ij})$, where $n_{ij} = E$ (number of visits to $j \mid X_0 = i$); $i, j \in B$. Let $\pi_{ik} = P(X\text{-chain absorbed at } k \mid X_0 = i)$, $i \in B$, $k \in C$, and set $\pi_{00} = \pi_{MM} = 1$, $\pi_{0M} = \pi_{M0} = 0$. Let $\boldsymbol{\pi}_k = (\pi_{1k}, \pi_{2k}, \cdots, \pi_{M-1,k})$.

Following Kemeny and Snell (1960), we have the results

$$(2.2)\qquad \text{and} \qquad \begin{aligned} N &= (I - Q)^{-1}, \\ \boldsymbol{\pi}_k' &= N\boldsymbol{p}_k', \quad k \in C. \end{aligned}$$

To construct the transition matrix $\bar{P} = (\bar{p}_{ij})$ of the return process $\{\bar{X}_n\}$ we have to specify the elements $\{\bar{p}_{0l}, \bar{p}_{Ml} : l \in S\}$. The remaining elements of $\bar{P}$ are identical to the corresponding elements of P, since probabilistic descriptions of the two processes are identical as long as the processes are in the set B. For notational convenience, we will set $\bar{p}_{0l} = r_{0l}, \bar{p}_{Ml} = r_{Ml}, l \in S$, and define

$$r_0 = (r_{01}, r_{02}, \cdots, r_{0,M-1}),$$

(2.3) and

$$r_M = (r_{M1}, r_{M2}, \cdots, r_{M,M-1}).$$

We can write $\bar{P}$ in the form

$$\bar{P} = \begin{bmatrix} r_{00} & r_0 & r_{0M} \\ p_0' & Q & p_M' \\ r_{M0} & r_M & r_{MM} \end{bmatrix}. \tag{2.4}$$

The n-step transition probabilities of X and $\bar{X}$ are denoted $p_{ij}^{(n)}$, and $\bar{p}_{ij}^{(n)}$, respectively, and the stationary distribution of the return process is denoted by $(\alpha_0, \alpha_1, \cdots, \alpha_M)$.

It is then easy to check that for $0 < n \leqq m$,

$$P(\bar{X}_{m-n} \in C, \bar{X}_{m-k} \notin C, 0 < k < n \mid \bar{X}_m = j)$$

$$= ({}_c\bar{p}_{0j}^{(n)} p_0^{(m-n)} / p_j^{(m)}) + ({}_c\bar{p}_{Mj}^{(n)} p_M^{(m-n)} / p_j^{(m)}), \tag{2.5}$$

where $\{{}_c\bar{p}_{ij}^{(n)}\}$ are the n-step, C-avoiding transition probabilities of the return process, and $p_l^{(m)} = P(\bar{X}_m = l)$ is taken to be generated by any initial distribution for $\bar{X}_0$. It follows from (1.1) and (2.5) that

$$g_j^{(n)} = {}_c\bar{p}_{0j}^{(n)} \frac{\alpha_0}{\alpha_j} + {}_c\bar{p}_{Mj}^{(n)} \frac{\alpha_M}{\alpha_j}. \tag{2.6}$$

One method of deriving the form of (2.6) is to relate the quantities on the right-hand side to properties of the original absorbing chain; for example, moments of the age distribution can be computed by matrix-generating function methods. However, it is clear that the limit derived in (2.6) is independent of the distribution of $\bar{X}_0$. Hence we may assume that $\bar{X}_0$ has the stationary distribution, $(\alpha_0, \cdots, \alpha_M)$; the chain $\{\bar{X}_n\}$ is then stationary. We may now reverse time, to obtain the dual chain $\{\bar{Y}_n\}$, which has the transition matrix $\bar{P}^* = (\bar{p}^*_{ij})$, where

$$\bar{p}^*_{ij} = \bar{p}_{ji} (\alpha_j / \alpha_i). \tag{2.7}$$

If we let $D_0 = \operatorname{diag}\{\alpha_1, \cdots, \alpha_{M-1}\}$, then it is easy to show that we can write

$$\bar{P}^* = \begin{bmatrix} r_{00} & \alpha_0^{-1} p_0 D_0 & \alpha_0^{-1}\alpha_M r_{M0} \\ \alpha_0 D_0^{-1} r_0' & D_0^{-1} Q' D_0 & \alpha_M D_0^{-1} r_M' \\ \alpha_0 \alpha_M^{-1} r_{0M} & \alpha_M^{-1} p_M D_0 & r_{MM} \end{bmatrix}. \tag{2.8}$$

We also introduce the reversed absorbing process $\{Y_n\}$ which is derived from $\{\bar{Y}_n\}$ by making states in C absorbing. The corresponding transition matrix $P^* = (p^*_{ij})$ is

$$P^* = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \alpha_0 D_0^{-1} r_0' & D_0^{-1} Q' D_0 & \alpha_M D_0^{-1} r_M' \\ 0 & \mathbf{0} & 1 \end{bmatrix}. \tag{2.9}$$

It should now be clear that the age, G_j, of the $\{X_n\}$ process is the absorption time of the reversed absorbing process, with $Y_0 = j$. By means of the representation of the age in terms of a well-specified Markov chain, we can now use standard Markov chain theory to derive a number of interesting properties of our age distribution. We begin with two lemmas.

Lemma 1. Let $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_{M-1})$. Then

$$\alpha_0 \propto r_M \cdot \boldsymbol{\pi}_0' + r_{M0},$$

$$\alpha_M \propto r_0 \cdot \boldsymbol{\pi}_M' + r_{0M},$$

and

$$\alpha_j \propto (\alpha_0 r_0 N + \alpha_M r_M N)_j, \qquad j \in B.$$

Proof. To find the stationary distribution of $\bar{P}$, we solve $(\alpha_0, \boldsymbol{\alpha}, \alpha_M) = (\alpha_0, \boldsymbol{\alpha}, \alpha_M)\bar{P}$. Using (2.4), write this as

$$\begin{aligned} &\alpha_0 r_{00} + \boldsymbol{\alpha} p_0' + \alpha_M r_{M0} = \alpha_0, \\ &\alpha_0 r_{0M} + \boldsymbol{\alpha} p_M' + \alpha_M r_{MM} = \alpha_M, \end{aligned} \tag{2.10}$$

$$\alpha_0 r_0 + \boldsymbol{\alpha} Q + \alpha_M r_M = \boldsymbol{\alpha}. \tag{2.11}$$

From (2.11) and (2.2),

$$\boldsymbol{\alpha} = \alpha_0 r_0 N + \alpha_M r_M N. \tag{2.12}$$

Substituting (2.12) into (2.10), and using (2.2) again yields

$$\alpha_0(r_{0M} + r_0 \boldsymbol{\pi}_M') = \alpha_M(r_M \boldsymbol{\pi}_0' + r_{M0}). \tag{2.13}$$

The conclusion of the lemma follows by substituting (2.13) into (2.12).

In what follows, denote by * the corresponding properties of the reversed absorbing chain $\{Y_n\}$. For instance, N^* is the fundamental matrix of the reversed absorbing process and T_j^* is the absorption time distribution of Y, given $Y_0 = j$.

Lemma 2. (i) $N^* = D_0^{-1}N'D_0$.

(ii) $\pi_k^{*\prime} = \alpha_k D_0^{-1}N'r_k', \qquad k \in C.$

Proof. Immediate from (2.2), (2.9).

To find the age distribution, we require the absorption-time distribution of the Y process. One can show that the n-step transition matrix of this process is given by

$$P^{*n} = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ B_n^* p_0^{*\prime} & Q^{*n} & B_n^* p_M^{*\prime} \\ 0 & \mathbf{0} & 1 \end{bmatrix}, \tag{2.14}$$

where $B_n^* = I + Q^* + \cdots + Q^{*(n-1)}$. Then $P(T_j^* = n) = P$ (age $= n$ | 'now' at $j) = g_j^{(n)}$. Let $g_n = (g_1^{(n)}, \cdots, g_{M-1}^{(n)})$. We then have the following result.

Lemma 3. For $n \geqq 1$,

$$\begin{aligned} g_n' &= Q^{*(n-1)}(p_0^{*\prime} + p_M^{*\prime}) \\ &= D_0^{-1}(Q')^{n-1}(\alpha_0 r_0' + \alpha_M r_M'). \end{aligned}$$

Proof. The first statement is just by definition of g_n, and the final statement follows from (2.8).

The absorption probabilities of the reverse chain, Y, give the (limiting) probabilities of the X process having been restarted from a particular end of the state space, and will therefore be termed the restarting probabilities of X. A genetic interpretation, initially due to Levikson, will be given in the examples. We can now combine Lemmas 2 and 3 to give the following theorem.

Theorem 4. For $j \in B, n \geqq 1$,

(i) $$P(G_j = n) = g_j^{(n)} = \frac{\left(\sum_{k=0}^{M-1} r_{Mk}\pi_{k0}\right)\left(\sum_{k=1}^{M-1} r_{0k}p_{kj}^{(n-1)}\right) + \left(\sum_{k=1}^{M} r_{0k}\pi_{kM}\right)\left(\sum_{k=1}^{M-1} r_{Mk}p_{kj}^{(n-1)}\right)}{\left(\sum_{k=0}^{M-1} r_{Mk}\pi_{k0}\right)\left(\sum_{k=1}^{M-1} r_{0k}n_{kj}\right) + \left(\sum_{k=1}^{M} r_{0k}\pi_{kM}\right)\left(\sum_{k=1}^{M-1} r_{Mk}n_{kj}\right)}$$

(ii) P(last restart of X process was at M | 'now' at $j) = \pi_{jM}^*$

$$= \frac{\left(\sum_{k=1}^{M} r_{0k}\pi_{kM}\right)\left(\sum_{k=1}^{M-1} r_{Mk}n_{kj}\right)}{\left(\sum_{k=0}^{M-1} r_{Mk}\pi_{k0}\right)\left(\sum_{k=1}^{M-1} r_{0k}n_{kj}\right)+\left(\sum_{k=1}^{M} r_{0k}\pi_{kM}\right)\left(\sum_{k=1}^{M-1} r_{Mk}n_{kj}\right)}.$$

Proof. (i) follows from Lemma 3, and Lemma 1. (ii) follows from Lemma 2, and Lemma 1.

From the representation of the age of X as the absorption time of Y, we can readily deduce the moments of the age from standard Markov chain theory. Let $\boldsymbol{T} = (T_1, \cdots, T_{M-1})$ be the vector of absorption times for the X process. Then, following Kemeny and Snell (1960), p. 49, we have

$$\begin{aligned} \boldsymbol{ET}' &= \boldsymbol{Nl}', \\ \operatorname{Var} \boldsymbol{T}' &= (2N - I)\boldsymbol{Nl}' - (\boldsymbol{Nl}')^2, \end{aligned} \tag{2.15}$$

where $\boldsymbol{l} = (1, 1, \cdots, 1)$ and, for a vector $s = (s_1, s_2, \cdots, s_{M-1})$, $s^2 = (s_1^2, s_2^2, \cdots, s_{M-1}^2)$. The moments of the age distributions $\boldsymbol{G} = (G_1, \cdots, G_{M-1})$ are just the moments of the absorption-time distributions $\boldsymbol{T}^* = (T_1^*, \cdots, T_{M-1}^*)$, and these are computed using the following lemma.

Lemma 5.

$$\boldsymbol{ET}^{*\prime} = \boldsymbol{D}_0^{-1}\boldsymbol{N}'\boldsymbol{D}_0\boldsymbol{l}',$$

and

$$\operatorname{Var} \boldsymbol{T}^{*\prime} = \boldsymbol{D}_0^{-1}(2N' - I)N'\boldsymbol{D}_0\boldsymbol{l}' - (\boldsymbol{D}_0^{-1}\boldsymbol{N}'\boldsymbol{D}_0\boldsymbol{l}')^2.$$

Another interesting problem related to the age process is the following. What is the (limiting) distribution of the number of visits to state $i \in B$, between successive visits to C, for a given 'current' position j? This will give us added information on how the 'age process' is moving through its paths. We can easily reinterpret this distribution in terms of the reverse absorbing process. We are looking for the distribution of the number of visits to i before absorption, given $Y_0 = j$. Denote this random variable by N_{ji}^*. Then the distribution of N_{ji}^* is given by

$$\begin{aligned} P(N_{ji}^* = m) &= 1 - \frac{n_{ji}^*}{n_{ii}^*}, \qquad m = 0 \\ &= \frac{n_{ji}^*}{n_{ii}^*} \cdot \frac{1}{n_{ii}^*} \cdot \left(1 - \frac{1}{n_{ii}^*}\right)^{m-1}, \qquad m \geqq 1. \end{aligned} \tag{2.16}$$

3. A class of restarting distributions of use in genetics

We consider the special case that arises from (2.3) by setting, for some $i \in B$,

$$r_{0j} = 0, j \neq i; = 1, j = i,$$

(3.1) and

$$r_{Mj} = 0, j \neq M - i; = 1, j = M - i.$$

The preceeding results then specialise to:

$$\alpha_0 \propto \pi_{M-i,0},$$

(3.2)

$$\alpha_M \propto \pi_{i,M},$$

$$\alpha_j \propto \pi_{M-i,0} n_{ij} + \pi_{iM} n_{M-i,j}, \qquad j \in T,$$

(3.3) $$g_j^{(n)} = \frac{\pi_{M-i,0} p_{ij}^{(n-1)} + \pi_{i,M} p_{M-i,j}^{(n-1)}}{\pi_{M-i,0} n_{ij} + \pi_{i,M} n_{M-i,j}}, \qquad n \geqq 1, \qquad j \in B.$$

(3.4) $$P(\text{last restart of } X \text{ process at } M \mid \text{'now' at } j) = \frac{\pi_{i,M} n_{M-i,j}}{\pi_{M-i,0} n_{ij} + \pi_{i,M} n_{M-i,j}},$$

and

(3.5) $$n_{kj}^{*} = \frac{n_{jk}(\pi_{M-i,0} n_{ij} + \pi_{i,M} n_{M-i,j})}{(\pi_{M-i,0} n_{ik} + \pi_{i,M} n_{M-i,k})}.$$

The case $i = 1$ corresponds to the case considered by Levikson. The result of (3.2) suggests a new interpretation of a problem which has appeared in the genetics literature before. Diffusion processes are often used as approximations to the type of model described in the introduction. By suitably rescaling the time and state spaces, we arrive at a diffusion process $X(t)$ on $[0, 1]$, with drift and diffusion coefficients $m(x)$ and $v(x)$ respectively. In what follows, we will assume that

$$v(x) = x(1-x),$$

(3.6)

$$m(x) = x(1-x)\phi(x),$$

where $\phi(x)$ is an arbitrary polynomial. The specification in (3.6) ensures that 0 and 1 are exit boundaries, corresponding to the absorbing barriers of the original model. Now write

(3.7) $$G(x) = \exp\left\{-2\int^{x} \frac{m(y)}{v(y)}\, dy\right\}.$$

Then formal application of Wright's formula for stationary distributions (Wright (1937)) shows that X should have a stationary distribution $\alpha(x)$ of the form

(3.8) $$\alpha(x) \propto \{G(x)v(x)\}^{-1}, \qquad 0 < x < 1.$$

Since the specification (3.6) ensures that no stationary distribution can exist, what interpretation can be placed on (3.8)? This problem has received attention

before; for a fuller discussion, see Ewens (1963). We can now give another interpretation. The elements n_{ij} correspond to the pseudo-transient function of the diffusion, given by

$$\begin{aligned} n(p,x) &= \frac{2P_0(p)}{G(x)v(x)} \int_0^x G(y)dy, \qquad 0 < x \leqq p \\ &= \frac{2P_1(p)}{G(x)v(x)} \int_x^1 G(y)dy, \qquad p \leqq x < 1 \end{aligned} \tag{3.9}$$

where $P_1(p) = 1 - P_0(p) = P(X \text{ absorbed at } 1 \,|\, X(0) = p)$, and

$$P_1(p) = \int_0^p G(y)dy \Big/ \int_0^1 G(y)dy. \tag{3.10}$$

We introduce the instantaneous return process (IRP) $\bar{Y}(t)$, which is derived from X by jumping instantaneously to ε if the barrier at 0 is hit, and to $1-\varepsilon$ if the barrier at 1 is hit. We may take $0 < \varepsilon < \frac{1}{2}$. This process is again a diffusion with state space $(0,1)$. It clearly has a stationary distribution, $\alpha^*(x)$ and it can be shown from Feller (1954), p. 23 that it satisfies

$$\frac{1}{2}\frac{d}{dx}\{v(x)\alpha^*(x)\} - m(x)\alpha^*(x) = \begin{cases} D_1, & 0 < x < \varepsilon \\ D_2, & \varepsilon < x < 1-\varepsilon \\ D_3, & 1-\varepsilon < x < 1, \end{cases} \tag{3.11}$$

where D_i are constants, determined by requiring α^* to be continuous at ε and $1-\varepsilon$, and integrable on $(0,1)$. By analogy with (3.2), a substitution of the form

$$\alpha^*(x) \propto P_0(1-\varepsilon)n(\varepsilon, x) + P_1(\varepsilon)n(1-\varepsilon, x)$$

has the required property, and it follows that

$$\alpha^*(x) = \begin{cases} \dfrac{\left\{\int_0^\varepsilon G(y)dy\right\}^{-1}}{G(x)v(x)} \int_0^x G(y)dy, & 0 < x \leqq \varepsilon \\ \dfrac{1}{G(x)v(x)}, & \varepsilon < x < 1-\varepsilon \\ \dfrac{\left\{\int_{1-\varepsilon}^1 G(y)dy\right\}^{-1}}{G(x)v(x)} \int_x^1 G(y)dy, & 1-\varepsilon \leqq x < 1 \end{cases} \tag{3.12}$$

with

$$D_1 = \left\{2\int_0^\varepsilon G(y)dy\right\}^{-1}, \quad D_2 = 0, \quad D_3 = \left\{2\int_{1-\varepsilon}^1 G(y)dy\right\}^{-1}. \tag{3.13}$$

Wright's formula is derived by solving

$$\frac{1}{2}\frac{d}{dx}\{v(x)\alpha(x)\} - m(x)\alpha(x) = 0, \qquad 0 < x < 1. \tag{3.14}$$

It follows by comparing (3.12), (3.13) with (3.14), that as $\varepsilon \to 0$, the stationary measure of the IRP is given by the solution of (3.14). A simple genetic explanation of such a process is as follows: we can suppose that rare mutation events reintroduce a particular allele into the population whenever it is lost. As $\varepsilon \to 0$, this corresponds to reintroducing a single copy of an allele at frequency $1/M$, or $1 - 1/M$, where M is the (large) population size.

4. Age distributions for a class of genetic models

We specialise the results of (3.1)–(3.5) to the case $i = 1$, in the context of a class of genetic models introduced by Cannings (1974). Consider a single locus in a population of fixed size M in each generation. There are two possible alleles, denoted A and B. Let X_n be the number of A-alleles at time n. In the absence of mutation pressures, $\{X_n\}$ is a Markov chain of the required type. Two models in this class are:

(i) Wright–Fisher model.

$$p_{ij} = \binom{M}{j}\left(\frac{i}{M}\right)^j\left(1 - \frac{i}{M}\right)^{M-j}, \qquad i, j \in S. \tag{4.1}$$

(ii) Moran's model (Moran (1958)).

$$\begin{aligned} p_{i,i+1} &= p_{i,i-1} = \frac{i(M-i)}{M^2}, \qquad i \in B, \\ p_{ii} &= 1 - p_{i,i+1} - p_{i,i-1}, \qquad i \in B, \\ p_{ij} &= 0, \qquad |i-j| > 1. \end{aligned} \tag{4.2}$$

These models both satisfy $\pi_{i,0} = 1 - i/M = 1 - \pi_{i,M}$. However, the elements of N are only of a manageable form in the case of Moran's model. (The recent paper of Piva and Holgate (1977) may produce an explicit, but complicated, expression for N in the case of the Wright–Fisher model.) For Moran's model one obtains

$$n_{ij} = \frac{M(M-i)}{(M-j)}, \quad 1 \leqq j \leqq i, \quad \text{and} \quad n_{ij} = M\frac{i}{j}, \quad i \leqq j \leqq M-1. \tag{4.3}$$

The elements of N for other members of the class are usually given by (rescaled) diffusion approximations (cf. Ewens (1969)). For the Wright model, we have

$$n_{ij} \approx \frac{2(M-i)}{(M-j)}, \quad 1 \leqq j \leqq i, \quad \text{and} \quad n_{ij} \approx 2\frac{i}{j}, \quad i \leqq j \leqq M-1. \tag{4.4}$$

These are of course, just multiples of the exact result in (4.3). It follows from the previous formulae that, for the Moran model,

$$\pi^*_{jM} = \frac{n_{M-1,j}}{n_{M-1,j} + n_{1j}} = \frac{j}{M} = \pi_{jM},$$

(4.5) and

$$ET^*_j = EG_j = \sum_{l=1}^{M-1} \frac{(n_{1l} + n_{M-1,l})n_{lj}}{n_{1j} + n_{M-1,j}} = \sum_{l=1}^{M-1} n_{jl} = ET_j.$$

These results could be anticipated by noticing that for the Moran model with restarting points 1 and $M-1$, we have $\bar{P}^* \equiv \bar{P}$, i.e. the process is completely reversible, and so the age distribution of X is just the absorption distribution of X (cf. Watterson (1977)). The quantities π^*_{jM} have been interpreted by Levikson as the probability of the A allele being the oldest, given current frequency j. For the Moran model, $\pi^*_{j,M} = \pi_{j,M}$, as predicted by reversibility. However, for Wright's model, the process is not reversible, and so (4.5) is only the 'diffusion approximation' to what we want. It is therefore of interest to determine how accurate such an approximation is. In Table 1, exact results for the last restart probabilities, the mean and variance of the age, and mean and variance of the absorption times are compared with the corresponding diffusion approximations. It can be seen that the results are surprisingly good. For further comments on the adequacy of diffusion approximations to N in this context, see, for example, Tavaré (1979), Pollak and Arnold (1975).

5. One-barrier models

The state space of the X-chain is now $S = \{0, 1, \cdots, M\}$, but $M = \infty$ is allowed. The absorbing state is $C = \{0\}$, and again $B = S - C$. The notation of the previous sections will be used again. In order for the return process to be recurrent, it is obvious that we must have the absorption probabilities of the absorbing process identically 1. This will be assumed throughout. We are again interested in the limiting distribution specified by (1.1). The transition matrix P of the X-chain is written

$$P = \begin{pmatrix} 1 & \mathbf{0} \\ \boldsymbol{p}_0' & Q \end{pmatrix} \tag{5.1}$$

while the return process transition matrix is given by

$$\bar{P} = \begin{pmatrix} r_{00} & \boldsymbol{r}_0 \\ \boldsymbol{p}_0' & Q \end{pmatrix}, \tag{5.2}$$

where $\boldsymbol{r}_0 = (r_{01}, r_{02}, \cdots)$. $\bar{P}$ is assumed to be the transition matrix of a positive-recurrent, irreducible, aperiodic Markov chain, whose stationary distribution we denote $(\alpha_0, \alpha_1, \cdots)$. The limit in (1.1) is now given by

TABLE 1

Exact and diffusion approximations (DA) for Wright's model

	$M = 10$			$M = 15$		
	π^*_{jM} (DA)	EG_j (DA)	Var G_j (DA)	π^*_{jM} (DA)	EG_j (DA)	Var G_j (DA)
j		ET_j	Var T_j		ET_j	Var T_j
1	0.0771	4.950	52.180	0.0505	5.548	91.511
	(0.1000)	(5.658)	(62.880)	(0.0667)	(6.503)	(114.313)
		5.753	60.866		6.624	111.113
3	0.2918	11.517	90.655	0.1921	14.164	193.443
	(0.3000)	(11.279)	(90.119)	(0.2000)	(14.070)	196.481
		11.031	91.009		13.825	195.873
5	0.5000	13.299	92.543	0.3297	18.477	214.381
	(0.5000)	(12.913)	(90.487)	(0.3333)	(18.134)	(212.541)
		12.509	92.655		17.751	214.669
7	0.7082	11.517	90.655	0.4659	20.158	216.394
	(0.7000)	(11.279)	(90.119)	(0.4667)	(19.761)	(212.966)
		11.031	91.009		19.327	216.490
9	0.9229	4.950	52.180	0.6022	19.604	216.131
	(0.9000)	(5.658)	(62.880)	(0.6000)	(19.225)	(213.215)
		5.753	60.866		18.808	216.261
11				0.7388	16.697	208.519
				(0.7333)	(16.443)	(208.541)
					16.116	209.272
13				0.8701	11.240	164.772
				(0.8667)	(10.863)	(169.467)
					10.743	167.420

$$g_j^{(n)} = {}_0\bar{p}_{0j}^{(n)} \frac{\alpha_0}{\alpha_j}; \qquad n \geqq 1, \quad j \in B,$$

where ${}_0\bar{p}_{0j}^{(n)}$ is the n-step, 0-avoiding transition probability of the return process. Under the stated conditions, the limit is again independent of the distribution of $\bar{X}_0$, which will now be taken to be $(\alpha_0, \alpha_1, \cdots)$. The corresponding return processes $\{\bar{Y}_n\}$ and $\{Y_n\}$ have transition matrices given by

$$\bar{P}^* = \begin{pmatrix} r_{00} & \alpha_0^{-1}\boldsymbol{p}_0 D_0 \\ \alpha_0 D_0^{-1}\boldsymbol{r}_0' & D_0^{-1}Q'D_0 \end{pmatrix}, \tag{5.3}$$

and

$$P^* = \begin{pmatrix} 1 & \mathbf{0} \\ \alpha_0 D_0^{-1} \boldsymbol{r}_0' & D_0^{-1} Q' D_0 \end{pmatrix} \tag{5.4}$$

respectively. Here $D_0 = \text{diag}\{\alpha_1, \alpha_2, \cdots\}$, and $D_0^{-1} = \text{diag}\{1/\alpha_1, 1/\alpha_2, \cdots\}$. The result concerning the fundamental matrix N is given by

$$N(I - Q) = (I - Q)N = I. \tag{5.5}$$

If $M < \infty$, then (5.5) reduces to (2.2). If $M = \infty$, then N is the minimal positive right (and left) inverse of $(I - Q)$. See, for example, Kemeny, Snell and Knapp (1966), p. 108. Again the age distribution of the X-process is related to the absorption time of the reverse process. The following lemma is useful in determining when the (recurrent) return process is positive, and in finding the stationary distribution in terms of N.

Lemma 6. (i) The return chain is positive if, and only if, $\Sigma_{k=1}^{\infty} r_{0k} E T_k < \infty$,

(ii) $$\alpha_0 \propto 1, \quad \alpha_j \propto \sum_{k=1}^{\infty} r_{0k} n_{kj}, \qquad j \in B.$$

Proof. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots)$. The chain is positive if, and only if, $(\alpha_0, \boldsymbol{\alpha}) = (\alpha_0, \boldsymbol{\alpha})\bar{P}$ has a non-trivial non-negative solution, with $\boldsymbol{\alpha} \boldsymbol{l}' < \infty$, where $\boldsymbol{l} = (1, 1, \cdots)$. From (5.2), (5.5) we see that $\boldsymbol{\alpha} = \alpha_0 \boldsymbol{r}_0 N$ so that $\boldsymbol{\alpha} \boldsymbol{l}' < \infty$ if, and only if, $\boldsymbol{r}_0 N \boldsymbol{l}' < \infty$, i.e. $\Sigma_{k=1}^{\infty} r_{0k} E T_k < \infty$. (ii) follows immediately.

Theorem 7. (i) $N^* = D_0^{-1} N' D_0$.

(ii) $$g_j^{(n)} = \frac{\sum_{k=1}^{\infty} r_{0k} p_{kj}^{(n-1)}}{\sum_{k=1}^{\infty} r_{0k} n_{kj}}, \qquad n \geqq 1, \quad j \in B.$$

(iii) $$E\boldsymbol{T}^{*\prime} = D_0^{-1} N' D_0 \boldsymbol{l}'.$$

(iv) $$\text{Var}\, \boldsymbol{T}^{*\prime} = D_0^{-1}(2N' - I) N' D_0 \boldsymbol{l}' - (D_0^{-1} N' D_0 \boldsymbol{l}')^2.$$

Proof. These follow by arguments analogous to the equivalent results in the two-barrier case, with the help of Lemma 6.

Theorem 7 gives us the age distribution and some of its properties. If we specialise to the case $\boldsymbol{r}_0 = (1, 0, \cdots)$, so that the return is made to state 1, then we arrive at the equivalent of Levikson's models. In fact, we have the following corollary.

Corollary 8 (Levikson-type models). If $E T_1 < \infty$, then for $j, k \in B$,

(i) $$g_j^{(n)} = P(G_j = n) = \frac{p_{1j}^{(n-1)}}{n_{1j}}, \qquad n \geqq 1.$$

(ii) $$n^*_{jk} = \frac{n_{1k} n_{kj}}{n_{1j}}.$$

(iii) $$EG_j = \sum_{k=1}^{\infty} \frac{n_{1k} n_{kj}}{n_{1j}}.$$

We can again use (2.15) to give the distribution of the number of visits to a state i for the reversed process for a given starting state j. The special case $i = j$ does not give us a good measure of the differences in behaviour of the two absorbing processes, since $n^*_{jj} = n_{jj}$. If $\bar{P} = \bar{P}^*$, then the two recurrent processes behave in an identical way. Then, of course, we will also have $N = N^*$. The role of such reversibility is discussed in the next section.

6. Reversibility and comments

The role of complete reversibility in the determination of properties of age distributions is apparent, since if $\bar{P} = \bar{P}^*$ the process is reversible, and then the two absorbing processes are identical. The idea of complete reversibility has been discussed in the case of certain diffusion results by Watterson (1977). In any case, we can derive age results in a simple way via time-reversal, and the limiting operation of (1.1) can then be given a simple intuitive meaning. In the one-barrier model, it is possible for the (recurrent) return process to be null, in which case the limit in (1.1) need not exist, and need not be independent of the initial distribution. Pakes (1978) discusses such chains in the case of Levikson-type return boundaries. However, if we are prepared to believe that the return process is stationary, which seems reasonable from the point of view of applications, then the results of Section 5 still follow. They will be identical to the results of Pakes in the cases where the strong ratio limit property holds. It is possible that time reversal may also give some results in the case of transient return chains.

For an application of reversibility to the infinite-allele Moran model, see Kelly (1977), and for another approach to age distributions see Sawyer (1977). The methods of this paper can be extended, under some restrictions, to continuous-time Markov chains. The results will be presented in Tavaré (1978b).

Acknowledgements

I should like to thank Geoff Watterson and Tony Pakes for helpful remarks on an earlier draft of this paper, and the referee, whose comments have considerably improved the presentation.

References

CANNINGS, C. (1974) The latent roots of certain Markov chains arising in genetics. I. Haploid models. *Adv. Appl. Prob.* **6,** 260–290.

DARROCH, J. N. AND SENETA, E. (1965) On quasi-stationary distributions in absorbing, discrete time, finite Markov chains. *J. Appl. Prob.* **2,** 88–100.

EWENS, W. J. (1963) The diffusion equation and a pseudo-distribution in genetics. *J. R. Statist. Soc.* **B25,** 405–412.

EWENS, W. J. (1969) *Population Genetics.* Methuen, London.

FELLER, W. (1954) Diffusion processes in one dimension. *Trans. Amer. Math. Soc.* **77,** 1–31.

KELLY, F. P. (1977) Some exact results for the Moran neutral-allele model. *Adv. Appl. Prob.* **9,** 197–201.

KEMENY, J. G. AND SNELL, J. L. (1960) *Finite Markov Chains.* Van Nostrand, Princeton, NJ.

KEMENY, J. G., SNELL, J. L. AND KNAPP, A. W. (1966) *Denumerable Markov Chains.* Van Nostrand, Princeton, NJ.

LEVIKSON, B. (1977) The age distribution of Markov processes. *J. Appl. Prob.* **14,** 492–506.

MORAN, P. A. P. (1958) Random processes in genetics. *Proc. Camb. Phil. Soc.* **54,** 60–71.

PAKES, A. (1978) The age distribution of a Markov chain. *J. Appl. Prob.* **15,** 65–77.

POLLAK, E. AND ARNOLD, B. C. (1975) On sojourn times at particular gene frequencies. *Genet. Res., Camb.* **25,** 89–94.

PIVA, M. AND HOLGATE, P. (1977) The eigenvectors of a finite population model. *Ann. Hum. Genet.* **41,** 103–106.

SAWYER, S. (1977) On the past history of an allele now known to have frequency *p. J. Appl. Prob.* **14,** 439–450.

SENETA, E. (1966) Quasi-stationary distributions and time reversion in genetics. *J. R. Statist. Soc.* **B28,** 253–277.

STIGLER, S. M. (1970) Estimating the age of a Galton–Watson branching process. *Biometrika* **57,** 505–512.

TAVARÉ, S. (1978a) Age distributions for Markov processes in genetics (abstract). *Adv. Appl. Prob.* **10,** 17–19.

TAVARÉ, S. (1978c) Time reversal and age distributions. II. Continuous-time Markov chains.

TAVARÉ, S. (1979) Sojourn times for conditioned Markov chains. *Theoret. Popn Biol.* **15,** 108–113.

THOMPSON, E. A. (1976) Estimation of age and rate increase of rare mutants. *Amer. J. Hum. Genet.* **28,** 442–452.

WATTERSON, G. A. (1976) Reversibility and the age of an allele I. Moran's infinitely many neutral alleles model. *Theoret. Popn Biol.* **10,** 239–253.

WATTERSON, G. A. (1977) Reversibility and the age of an allele II. Two allele models with selection and mutation. *Theoret. Popn Biol.* **12,** 179–196.

WRIGHT, S. (1937) The distribution of gene frequencies in populations. *Proc. Nat. Acad. Sci. USA* **23,** 307–320.