Adv. Appl. Prob. **13**, 681–703 (1981)
Printed in N. Ireland
0001–8678/81/040681–23$02.55

COMMENTS ON THE AGE DISTRIBUTION OF MARKOV PROCESSES

ANTHONY G. PAKES,* *University of Western Australia*
SIMON TAVARÉ,** *University of Utah*

Abstract

Previous work on the concept of a limiting conditional age distribution of a discrete-state continuous-time Markov process with one absorbing state is generalised. The generalisation allows this process to have a finite number of absorbing states and the associated return process to have an arbitrary initial distribution on the transient states of the absorbing process. If the return process is ρ-recurrent, possesses the strong ratio limit property and satisfies some further requirements then the limiting age distribution exists. The proof of this result requires a new representation of the ρ-invariant measure of the return process.

The following examples are treated, (a) finite state space birth–death processes, (b) Markov branching processes and the linear death process, and (c) the linear birth and death process with killing.

MARKOV PROCESS; LIMITING AGE; ρ-CLASSIFICATION; STRONG RATIO LIMIT PROPERTY; BIRTH AND DEATH PROCESSES; MARKOV BRANCHING PROCESSES; LIMIT THEOREMS

1. Introduction

Consider a Markov process $\mathcal{Y}=\{Y(t), t\geqq 0\}$ with irreducible state space $\mathcal{S}=\{0,1,2,\cdots\}$ or $\mathcal{S}=\{0,1,\cdots,N\}$, where $N\in\mathbb{N}$. Fix $a\in\mathcal{S}$ and define the last-exit time from $\{a\}$ by time t to be $\gamma_a(t,\omega)=\sup \overline{S_a(\omega)}\cap[0,t]$ where $S_a(\omega)=\{t: Y(t,\omega)=a\}$ and $\sup\phi=0$. We define the 'age' of the process by $T(t,\omega)=t-\gamma_a(t,\omega)$. Pakes (1979) gave conditions ensuring the existence of the (limiting) conditional age distribution

$$a_j(t)=\lim_{\tau\to\infty} P_i(T(\tau)\leqq t \mid Y(\tau)=j),$$

Received 6 January 1981.
* Postal address: Department of Mathematics, University of Western Australia, Nedlands, WA6009, Australia.
** Present address: Department of Statistics, Colorado State University, Fort Collins, Colorado, 80523, U.S.A.

where $P_i(\cdot)=P(\cdot \mid Y(0)=i)$. The required conditions are that $\mathcal{Y}$ is ρ-recurrent and that it possesses the strong ratio limit property.

This result was applied to a return process $\mathcal{Y}$ constructed from a regular minimal absorbing process $\mathcal{X}=\{X(t), t\geqq 0\}$ whose state space is $\mathcal{S}$ and $\{0\}$ is absorbing and accessible from the irreducible transient set $\mathcal{T}=\mathcal{S}\backslash\{0\}$. The return process is conservative, has the same generator on $\mathcal{T}$ as has $\mathcal{X}$, and if $\{0\}$ is hit then the next jump is into $\{1\}$ after an almost surely positive sojurn in $\{0\}$. The state $\{1\}$ has no special significance, and any other single state in $\mathcal{T}$ can take its place.

Recently, Tavaré (1980) has considered variants of this scheme. He allows an arbitrary initial distribution on $\mathcal{T}$ for the return process, and jumps from $\{0\}$ to $\mathcal{T}$ following a given distribution. He also considered the case of an absorbing process for which $|\mathcal{S}|<\infty$, $\{0\}$ and $\{N\}$ are absorbing, and the corresponding return process allows jumps from $\{0\}$ and $\{N\}$ into $\mathcal{T}=\mathcal{S}\backslash\{0, N\}$. He imposes the restriction that the return processes are positive recurrent.

It is our intention here to give a unified treatment which combines the essence of the work of Pakes and Tavaré. We accomplish this by letting $\mathcal{Y}$ be as in the first paragraph. Let H be a finite subset of $\mathcal{S}$, $\mathcal{T}=\mathcal{S}\backslash H$, and $\{c_i, i\in\mathcal{T}\}$ be the initial distribution of $\mathcal{Y}$. If $S_H(\omega)=\{t: Y(t,\omega)\in H\}$, then define $\gamma_H(t,\omega)=\sup \overline{S_H(\omega)}\cap[0,t]$, and $T(t,\omega)=t-\gamma_H(t,\omega)$ to be the time from t to the previous exit from H. We seek conditions ensuring the existence of

$$a_j(t)=\lim_{\tau\to\infty} P(T(\tau)\leqq t \mid Y(\tau)=j), \qquad j\in\mathcal{T},$$

where $P(\cdot)$ corresponds to the fixed initial distribution $\{c_i, i\in\mathcal{T}\}$. It transpires that two ingredients are required for this. We shall require an extension of the strong ratio limit property which takes account of general starting conditions. This will follow from the discrete-time results of Orey (1971), p. 79. Secondly, we must extend Chung's work on last-exit times (Chung (1967), §II 12), and use this to generalise the representation for the ρ-invariant measure obtained by Pakes (1979). This is carried out in Sections 2 and 3 and the results are applied to absorbing processes in Section 4. In Section 5 we discuss the last-hitting times of states in the transient set of an absorbing process. Finally in Section 6 we present some examples, viz. (a) finite-state-space birth–death processes; (b) Markov branching processes and the linear death process, and (c) the linear birth and death process with killing.

2. The strong ratio limit property

Let $p_j(t)=P(Y(t)=j)$, and ${}_Hp_{ij}(t)=P_i(Y(t)=j,\ Y(s)\notin H;\ 0<s<t)$, so that if i or $j\in H$, then ${}_Hp_{ij}(t)=0$. If $j\in\mathcal{T}$, then

$$(2.1)\qquad a_j(t,\tau)=P(T(\tau)\leqq t \mid Y(\tau)=j)=1-\sum_i p_i(\tau-t)\,{}_Hp_{ij}(t)/p_j(\tau).$$

Suppose that $\mathscr{Y}$ is ρ-recurrent, with ρ-invariant measure $\{m_j\}$, and function $\{x_j\}$; cf. Kingman (1963). These are positive on $\mathscr{S}$, unique up to multiplicative constants, and hence can be normalised by setting $m_a = x_a = 1$ for a fixed $a \in H$. The strong ratio limit property asserts that for $i, j, k, l \in \mathscr{S}$, $t \in \mathbb{R}$,

$$\lim_{\tau \to \infty} p_{ij}(t+\tau)/p_{kl}(\tau) = e^{-\rho t} x_i m_j / x_k m_l, \tag{2.2}$$

and it is valid iff there exists $\varepsilon > 0$ such that

$$\limsup_{\tau \to \infty} p_{00}(t+\tau)/p_{00}(\tau) \leqq e^{-\rho t}, \qquad 0 \leqq t < \varepsilon.$$

The strong ratio limit property suggests that for $j, k, l \in \mathscr{S}$, $t \in \mathbb{R}$

$$p_j(t+\tau)/p_{kl}(\tau) \to e^{-\rho t} c m_j / x_k m_l, \tag{2.3}$$

where $c = \sum_{i \in \mathscr{T}} c_i x_i$. This is clearly true if $\{c_i\}$ has finite support, and if $c = \infty$, then (2.3) follows from (2.2) and Fatou's lemma. When $c < \infty$, the following is true.

Lemma 1. Suppose that $\mathscr{Y}$ is ρ-recurrent, $c = \sum_i c_i x_i < \infty$, and that $\mathscr{Y}$ has the strong ratio limit property (2.2). Then (2.3) holds for a given $j, k, l \in \mathscr{S}$ iff there is a finite positive constant M such that for all $i \in \mathscr{S}$ and sufficiently large τ

$$p_{ij}(t+\tau)/p_{kl}(\tau) \leqq M(x_i x_l / x_k x_j) e^{-\rho t}. \tag{2.4}$$

Remarks. The proof does not require that $\{c_i\}$ be concentrated on $\mathscr{T}$. If $\rho = 0$ then $x_i \equiv 1$, and hence $c < \infty$.

Proof. Let $q_{ij}(t) = e^{\rho t} p_{ij}(t) x_j / x_i$. Then (2.2) is equivalent to

$$q_{ij}(t+\tau)/q_{kl}(\tau) \to x_j m_j / x_l m_l; \qquad i, j, k, l \in \mathscr{S}, \qquad t \in \mathbb{R}. \tag{2.5}$$

Notice that $\{x_i m_i\}$ is an invariant measure for $[q_{ij}(t)]$. The left-hand side of (2.3) becomes

$$c e^{-\rho t} (x_l / x_k x_j)(q_j(t+\tau)/q_{kl}(\tau)),$$

where $q_j(\tau) = \sum_i a_i q_{ij}(\tau)$, and $a_i = c_i x_i / c$, which defines a distribution on $\mathscr{S}$. Finally, Condition (2.4) is equivalent to

$$q_{ij}(t+\tau)/q_{kl}(\tau) \leqq M, \qquad i \in \mathscr{S}, \qquad \tau \geqq \tau_0 > 0. \tag{2.6}$$

Thus if (2.6) holds, then dominated convergence and (2.5) show that

$$q_j(t+\tau)/q_{kl}(\tau) \to x_j m_j / x_l m_l, \tag{2.7}$$

and (2.3) follows. Conversely, if (2.6), and hence (2.4), is not valid, it is easy to modify Orey's ((1971), p. 79) argument to show that (2.3) fails.

Nummelin (1979) has recently given conditions ensuring that Markov chains with general state spaces possess the strong ratio limit property. In the next lemma, we adapt his idea of small measures to give a condition which ensures that the assertion of Lemma 1 is true. This condition is stronger than the assumptions of Lemma 1, but does not involve ratios as in (2.4) above.

Lemma 2. Suppose that $\mathscr{Y}$ is ρ-recurrent and has the strong ratio limit property. Suppose also that there exists $T>0$, a finite set $A\subset\mathscr{S}$ and $\delta>0$ such that for all sufficiently large i,

$$\int_0^T \sum_{r\in A} e^{\rho s}p_{ri}(s)\,ds \geqq \delta c_i.$$

Then $c<\infty$, and (2.3) holds.

Proof. Since $e^{\rho s}\sum_i p_{ri}(s)x_i = x_r$, it is clear that $c<\infty$. The proof can be completed simply by showing that

$$\overline{\lim_{J\to\infty}}\ \overline{\lim_{\tau\to\infty}} \sum_{i>J} c_i p_{ij}(t+\tau)/p_{kl}(\tau)=0. \tag{2.8}$$

But the sum is dominated by

$$\delta^{-1}\int_0^T \sum_{r\in A} m_r e^{\rho s}\Big(p_{rj}(s+t+\tau)-\sum_{i\leqq J} p_{ri}(s)p_{ij}(t+\tau)/p_{kl}(\tau)\Big)\,ds$$
$$\to (m_j/x_k x_l\delta)e^{-\rho t}\int_0^T \sum_{r\in A} m_r\Big(x_r - e^{\rho s}\sum_{i\leqq J} p_{ri}(s)x_i\Big)\,ds \qquad (t\to\infty)$$

where we have used the Chapman–Kolmogorov equation to obtain the first line and then the strong ratio limit property and dominated convergence. Now let $J\to\infty$, and use monotone convergence to obtain (2.8).

Theorem 1. Suppose that $\mathscr{Y}$ is ρ-recurrent and has the strong ratio limit property. If the additional assumptions in either Lemma 1 or Lemma 2 are satisfied then $\lim_{\tau\to\infty} a_j(t,\tau)$ exists, and is given by

$$a_j(t)=1-\sum_{i\in\mathscr{T}}(m_i/m_j)e^{\rho t}{}_H p_{ij}(t).$$

Proof. This is similar to that in Pakes (1979), p. 283, but extends a simplification due to Tavaré (1980). The hypotheses show that

$$\lim_{\tau\to\infty} p_i(\tau-t)/p_j(\tau)=e^{\rho t}m_i/m_j, \qquad i,j\in\mathscr{S}, \qquad t\geqq 0.$$

Equation (2.1) and Fatou's lemma yield

$$\liminf_{\tau\to\infty}\,(1-a_j(t,\tau))\geqq \sum_{i\in\mathscr{T}} e^{\rho t}(m_i/m_j)_H p_{ij}(t).$$

The Chapman–Kolmogorov equations yield

$$a_j(t,\tau)=\Big[p_j(\tau)-\sum_{i\in\mathcal{T}}p_i(\tau-t)_Hp_{ij}(t)\Big]\Big/p_j(\tau)$$
$$=\Big\{\sum_{i\in H}p_i(\tau-t)p_{ij}(t)+\sum_{i\in\mathcal{T}}p_i(\tau-t)[p_{ij}(t)-{}_Hp_{ij}(t)]\Big\}\Big/p_j(\tau).$$

Since the summands are non-negative, Fatou's lemma is applicable, yielding

$$\liminf_{\tau\to\infty}a_j(t,\tau)\geqq\sum_{i\in\mathcal{S}}e^{\rho t}(m_i/m_j)p_{ij}(t)-\sum_{i\in\mathcal{T}}e^{\rho t}(m_i/m_j)_Hp_{ij}(t)$$
$$=1-\sum_{i\in\mathcal{T}}e^{\rho t}(m_i/m_j)_Hp_{ij}(t),$$

by ρ-invariance of $\{m_i\}$. The proof is now complete.

If $\mathcal{Y}$ is ρ-positive, then $[q_{ij}(t)]$ is a positive recurrent matrix and (2.6), whence (2.4), is automatically satisfied. If also $c_i=O(m_i)$ then $c<\infty$.

3. The calculation of $a_j(t)$

We begin by reviewing some preliminary material. Remember that $\{a\}$ is a fixed state in H. For $j\neq a$, let

$$\phi_{aj}(\delta;s)=\delta^{-1}\sum_k p_{ak}(\delta)_ap_{kj}(s).$$

Chung (1967), p. 201, shows that as $\delta\downarrow 0$ $\phi_{aj}(\delta;s)$ converges to a measurable bounded function $g_{aj}(s)$ and

$$p_{aj}(t)=\int_0^t p_{aa}(t-s)g_{aj}(s)\,ds\qquad(j\neq a).\tag{3.1}$$

It was shown by Pakes (1979) that the sequence $\{m_j\}$ defined by $m_a=1$, $m_j=\int_0^\infty e^{\rho s}g_{aj}(s)\,ds$ $(j\neq a)$ is a ρ-invariant measure for $\mathcal{Y}$. It was then shown that if $H=\{a\}$, than $a_j(t)=m_j^{-1}\int_0^t e^{\rho s}g_{aj}(s)\,ds$ $(j\neq a)$.

Now define

$$\phi_{ij}^{(H)}(\delta;s)=\delta^{-1}\sum_{k\in\mathcal{T}}p_{ik}(\delta)_Hp_{kj}(s),\qquad(i\in H,\quad j\in\mathcal{T}).$$

The details of the proofs of Theorems 2 and 3 of Chung (1967), §II.12, can be modified to prove the following result.

Proposition. If $i\in H$, $j\in\mathcal{T}$ then

$$p_{ij}(t)=\lim_{m\to\infty}\sum_{v=1}^{[2^mt]}\sum_{k\in H}p_{ik}\Big(\frac{v-1}{2^m}\Big)\sum_{l\in\mathcal{T}}p_{kl}(2^{-m})_Hp_{lj}(t-v2^{-m}),$$

and for each m, the equality may be replaced by $\geqq$ without the limit. As $\delta \downarrow 0$, $\phi_{ij}^{(H)}(\delta; s)$ converges to a bounded measurable function $g_{ij}^{(H)}(s)$, and

$$p_{ij}(t) = \int_0^t \sum_{k \in H} p_{ik}(t-s) g_{kj}^{(H)}(s)\, ds \qquad (i \in H, \quad j \in \mathcal{T}). \tag{3.2}$$

Theorem 2. The conditional age distribution function $a_j(t)$ has the form

$$a_j(t) = m_j^{-1} \int_0^t \sum_{k \in H} m_k g_{kj}^{(H)}(s) e^{\rho s}\, ds, \qquad j \in \mathcal{T}, \tag{3.3}$$

and it is non-defective.

Proof. Let T_H be the hitting time of H, and for $i \in \mathcal{T}$, $j \in H$ let $F_{ij}(t) = P_i(T_H \leqq t, Y(T_H) = j)$. Let $\hat{F}_{ij}(\theta) = \int_0^\infty e^{-\theta t} F_{ij}(dt)$, and $\hat{\psi}(\theta) = \int_0^\infty \psi(t) e^{-\theta t}\, dt$ for any function $\psi : [0, \infty) \to [0, \infty)$, provided this transform exists. Thus (3.2) yields

$$\hat{p}_{ij}(\theta) = \sum_{k \in H} \hat{p}_{ik}(\theta) \hat{g}_{kj}^{(H)}(\theta), \qquad i \in H, \qquad j \in \mathcal{T}. \tag{3.4}$$

If $i \in \mathcal{T}$, then

$$p_{ij}(t) = {}_H p_{ij}(t) + \sum_{k \in H} \int_0^t p_{kj}(t-s) F_{ik}(ds), \qquad j \in \mathcal{S},$$

whence from (3.4) and the last relation

$$\begin{aligned}
\sum_{i \in \mathcal{T}} m_i \, {}_H\hat{p}_{ij}(\theta) &= \sum_{i \in \mathcal{T}} m_i \hat{p}_{ij}(\theta) - \sum_{k \in H} \sum_{i \in \mathcal{T}} m_i \hat{F}_{ik}(\theta) \hat{p}_{kj}(\theta) \\
&= \sum_{i \in \mathcal{T}} m_i \hat{p}_{ij}(\theta) - \sum_{k \in H} \sum_{i \in \mathcal{T}} m_i \hat{F}_{ik}(\theta) \sum_{l \in H} \hat{p}_{kl}(\theta) \hat{g}_{lj}^{(H)}(\theta) \qquad (3.5) \\
&= \sum_{i \in \mathcal{T}} m_i \hat{p}_{ij}(\theta) - \sum_{i \in \mathcal{T}} \sum_{l \in H} m_i \hat{p}_{il}(\theta) \hat{g}_{lj}^{(H)}(\theta), \qquad j \in \mathcal{T}.
\end{aligned}$$

The Laplace transform of the invariant equations for $\{m_j\}$ yield $\sum_{i \in \mathcal{T}} m_i \hat{p}_{ij}(\theta) = (\rho + \theta)^{-1} m_j$, whence from (3.5) and then (3.4)

$$\begin{aligned}
\sum_{i \in \mathcal{T}} m_i \, {}_H\hat{p}_{ij}(\theta) &= (\rho + \theta)^{-1} m_j - \sum_{i \in H} m_i \hat{p}_{ij}(\theta) \\
&\quad - \sum_{l \in H} \left[(\rho + \theta)^{-1} m_l - \sum_{i \in H} m_i \hat{p}_{il}(\theta) \right] \hat{g}_{lj}^{(H)}(\theta) \\
&= (\rho + \theta)^{-1} \left\{ m_j - \sum_{l \in H} m_l \hat{g}_{lj}^{(H)}(\theta) \right\}.
\end{aligned}$$

Inverting this transform equation shows that

$$e^{\rho t}\sum_{i\in\mathcal{T}} m_{i\,H}p_{ij}(t)=m_j-\int_0^t e^{-\rho t}\sum_{l\in H} m_l g_{lj}^{(H)}(s)\,ds,$$

and (3.3) follows.

By letting $i=a$ in (3.2), and equating this with (3.1) we obtain for $j\in\mathcal{T}$

$$\hat{p}_{aa}(\theta)\hat{g}_{aj}(\theta)=\sum_{k\in H}\hat{p}_{ak}(\theta)\hat{g}_{kj}^{(H)}(\theta).$$

Now it is shown in Pakes (1979) that $\hat{p}_{ak}(\theta)$, $\hat{g}_{aj}(\theta)$ exist for $\theta>-\rho$ $(k\in\mathcal{S}, j\neq a)$, that $m_j=\hat{g}_{aj}(-\rho)$ $(j\neq a)$ and $m_k=\lim_{\theta\downarrow-\rho}\hat{p}_{ak}(\theta)/\hat{p}_{aa}(\theta)$. It follows that

$$m_j=\sum_{k\in H} m_k\int_0^\infty e^{\rho t}g_{kj}^{(H)}(t)\,dt,\qquad j\in\mathcal{T},$$

and hence that $a_j(\cdot)$ is non-defective.

4. Return processes

Suppose now that $\mathcal{X}=\{X(t), t\geqq 0\}$ is a Markov process with state space $\mathcal{S}$, H is a finite set of absorbing states, $\mathcal{T}=\mathcal{S}\backslash H$ is irreducible and transient, and H is accessible from $\mathcal{T}$. We shall assume that $\mathcal{X}$ is a regular minimal process corresponding to the generator $[u_{ij}]$ and let $[r_{ij}(t), t\geqq 0]$ denote its transition semigroup. We construct a return process $\mathcal{Y}$ as follows. For each $i\in H$, let $q_i\in(0,\infty)$ be given, and a distribution $\{p_{ij}, j\in\mathcal{T}\}$ be given. Define the generator $Q=[q_{ij}]$ by $q_{ij}=u_{ij}$ $(i\in\mathcal{T})$ and if $i\in H$ then $q_{ii}=-q_i$, $q_{ij}=q_ip_{ij}$ $(j\in\mathcal{T})$. Let $\mathcal{Y}$ be the minimal process constructed from Q; it is unique and regular. Thus if $Y(t)\in\mathcal{T}$, $\mathcal{Y}$ evolves according to the construction of $\mathcal{X}$ until it next hits a state $i\in H$. It sojourns there for a time which has an exponential distribution with parameter q_i, and then jumps to $\{j\}\in\mathcal{T}$ with probability p_{ij}, and then evolves as before.

If $k, j\in\mathcal{T}$, then ${}_Hp_{kj}(t)=r_{kj}(t)$, and hence

$$\phi_{ij}^{(H)}(\delta;s)=\delta^{-1}\sum_{k\in\mathcal{T}} p_{ik}(\delta)r_{kj}(s),\qquad i\in H,\qquad j\in\mathcal{T}.$$

If $i\in H$, $k\in\mathcal{T}$ then

$$p_{ik}(\delta)=\sum_{l\in\mathcal{T}}\int_0^\delta q_ie^{-q_i(\delta-u)}p_{il}r_{lk}(u)\,du,$$

whence

$$\begin{aligned}\phi_{ij}^{(H)}(\delta;s)&=\delta^{-1}q_i\sum_{l\in\mathcal{T}}\int_0^\delta e^{-q_i(\delta-u)}p_{il}r_{lj}(u+s)\,du\\&=\delta^{-1}(1-e^{-q_i\delta})\sum_{l\in\mathcal{T}}p_{il}r_{lj}(u_l'+s)\end{aligned}$$

where $0 < u_l' < \delta$, and we have used an integral mean-value theorem. Dominated convergence shows that

$$g_{ij}^{(H)}(s) = \sum_{l \in \mathcal{T}} q_{il} r_{lj}(s), \qquad i \in H, \qquad j \in \mathcal{T},$$

and hence

$$a_j(t) = m_j^{-1} \sum_{i \in H} \sum_{k \in \mathcal{T}} m_i q_{ik} \int_0^t e^{\rho s} r_{kj}(s)\, ds. \tag{4.1}$$

Suppose that $\mathcal{Y}$ is ρ-recurrent and satisfies (2.3). Then (3.3) holds. Let $\tilde{\mathcal{Y}} = \{\tilde{Y}(t), t \geqq 0\}$ be the reversed process whose transition semi-group is defined by

$$\tilde{p}_{ij}(t) = e^{\rho t} m_j p_{ji}(t)/m_i.$$

Let $\tilde{T}_H$ be the hitting time of H by $\tilde{\mathcal{Y}}$, and $\tilde{F}_i(t) = \tilde{P}_i(\tilde{T}_H \leqq t)$.

Theorem 3. Suppose that $\mathcal{Y}$ is a return process as defined earlier. Then

$$a_j(t) = \tilde{F}_j(t), \qquad t > 0, \qquad j \in \mathcal{T}.$$

Proof. Let $\tilde{\mathcal{X}} = \{\tilde{X}(t), t \geqq 0\}$ be the process obtained from $\tilde{\mathcal{Y}}$ by stopping at $\tilde{T}_H$, and let $[\tilde{r}_{ij}(t), t \geqq 0]$ be the transition semi-group of $\tilde{\mathcal{X}}$. Then ${}_H\tilde{p}_{ij}(t) = \tilde{r}_{ij}(t)$, $i, j \in \mathcal{T}$.

By its construction, $\tilde{\mathcal{X}}$ is the regular minimal process corresponding to its generator, which is $\tilde{u}_{ij} = m_j q_{ji}/m_i$ $(i \in \mathcal{T}, i \neq j)$, $\tilde{u}_{ii} = q_{ii} + \rho$ $(i \in \mathcal{T})$, $\tilde{u}_{ij} = 0$ $(i \in H)$. If $i, j \in \mathcal{T}$, the backward and forward systems for $\tilde{\mathcal{X}}$ are

$$\tilde{r}_{ij}'(t) = \sum_{k \in \mathcal{T}} \tilde{u}_{ik} \tilde{r}_{kj}(t) \quad \text{and} \quad \tilde{r}_{ij}'(t) = \sum_{k \in \mathcal{T}} \tilde{r}_{ik}(t) \tilde{u}_{kj}, \tag{4.2}$$

respectively. Let $\alpha_{ij}(t) = e^{\rho t} m_j r_{ji}(t)/m_i$, $i, j \in \mathcal{T}$. Then $\alpha_{ij}(0^+) = \delta_{ij}$, $\alpha_{ij}(t)$ satisfies the systems (4.2), since they transform into the forward and backward systems for $\mathcal{X}$, respectively.

If $i \in \mathcal{T}$, $j \in H$, the backward and forward systems for $\tilde{\mathcal{X}}$ are

$$\tilde{r}_{ij}'(t) = \sum_{k \in \mathcal{T}} \tilde{u}_{ik} \tilde{r}_{kj}(t) + \tilde{u}_{ij}, \tag{4.3}$$

and

$$\tilde{r}_{ij}'(t) = \sum_{k \in \mathcal{T}} \tilde{r}_{ik}(t) \tilde{u}_{kj}. \tag{4.4}$$

Equation (4.4) suggests the definitions $\alpha_{ij}(t) = \sum_{k \in \mathcal{T}} \int_0^t \alpha_{ik}(s)\, ds \tilde{u}_{kj}$, $i \in \mathcal{T}$, $j \in H$, and $\alpha_{ij}(t) = \delta_{ij}$ $(i \in H)$. Then $\alpha_{ij}(0^+) = \delta_{ij}$, and the $\alpha_{ij}(\cdot)$ satisfy the forward system for $\tilde{\mathscr{X}}$. Using (4.2), a little algebra shows that $\alpha_{ij}(t)$ also satisfies (4.3), with the same initial condition. It follows that $\alpha_{ij}(t) = \tilde{r}_{ij}(t)$, $i, j \in \mathscr{S}$.

Hence

$$\tilde{F}_j(t) = 1 - \sum_{i \in \mathscr{T}} {}_H\tilde{p}_{ji}(t) = 1 - \sum_{i \in \mathscr{T}} \tilde{r}_{ji}(t)$$

(4.5)

$$= 1 - m_j^{-1} \sum_{i \in \mathscr{T}} e^{\rho t} m_i r_{ij}(t) = a_j(t)$$

by Theorem 1, and the proof is completed.

The theorem has the intuitively appealing consequence that the limiting conditional age distribution is just the hitting time distribution of the reversed process $\tilde{\mathscr{Y}}$. If $\mathscr{Y}$ is reversible, that is $p_{ij}(t) = \tilde{p}_{ij}(t)$ then the limiting conditional age distribution of $\mathscr{Y}$ is the same as its hitting distribution. A necessary condition for this is $\rho = 0$, for if $\rho > 0$ then the $p_{ij}(t)$ decay exponentially fast, whereas the $\tilde{p}_{ij}(t)$ do not.

The representation of $a_j(t)$ in terms of a hitting time can be used to compute the moments of the age distribution. Define

$$G_{ij}(\rho) = \int_0^\infty e^{\rho t} r_{ij}(t)\, dt, \qquad i, j \in \mathscr{T},$$

and $G_{ij}^{(n+1)}(\rho) = \sum_{k \in \mathscr{T}} G_{ik}^{(n)}(\rho) G_{kj}(\rho)$. Then (4.5) leads eventually to

$$\text{(4.6)} \quad \mu_j^{(n)} = n \int_0^\infty t^{n-1}(1 - a_j(t))\, dt = n!\, m_j^{-1} \sum_{i \in H} m_i \sum_{k \in \mathscr{T}} q_{ik} G_{kj}^{(n+1)}(\rho), \qquad n = 0, 1, \cdots.$$

The same result can be derived by the method of Pakes (1979). Let $U = [u_{ij}, i, j \in \mathscr{T}]$ be given by $u_{ij} = q_{ij}$ $(i \neq j)$, $= q_{ii} + \rho$ $(i = j)$. Then it follows from (4.2) that $G = [G_{ij}(\rho), i, j \in \mathscr{T}]$ is the minimal non-negative right (and left) inverse of $-U$.

Finally, define $\tilde{b}_{jl} = \tilde{P}_j(\tilde{Y}(\tilde{T}_H) = l)$, $j \in \mathscr{T}$, $l \in H$. Then it is straightforward to show that

$$\text{(4.7)} \qquad \tilde{b}_{jl} = m_l m_j^{-1} \sum_{k \in \mathscr{T}} q_l p_{lk} G_{kj}(\rho) \qquad j \in \mathscr{T}, \qquad l \in H.$$

$\tilde{b}_{jl}$ is interpreted as the (limiting) conditional probability that H was last visited at l; cf. (6.1).

5. Last-hit distributions

Consider the absorbing process $\mathscr{X}$ defined in the last section. For $j \in \mathscr{T}$, let $L_j = \sup\{t : X(t) = j\}$, a possibly defective random variable. Let $\beta = \{\beta_j, j \in \mathscr{T}\}$ be the initial distribution of $\mathscr{X}$.

Theorem 4. $\Lambda_j(t)=P(L_j\leqq t\mid L_j<\infty)=A^{-1}\sum_{i\in\mathcal{T}}\beta_i\int_0^t r_{ij}(s)\,ds$ where $A=\sum_{i\in\mathcal{T}}\beta_iG_{ij}$, and $G_{ij}=\int_0^\infty r_{ij}(t)\,dt$.

Proof. Let ρ_{ij} be the probability of hitting $\{j\}$ from $\{i\}$. Then

$$P(t<L_j<t+dt)=\sum_{\substack{i\in\mathcal{S}\\ i\neq j}}\sum_{k\in\mathcal{T}}\beta_kr_{kj}(t)u_{ji}(1-\rho_{ij})\,dt+o(dt),$$

since the ith summand is the probability that $\mathcal{X}$ has reached $\{j\}$ by time t, then jumps to $\{i\}$ during $(t, t+dt)$ and never again returns to $\{j\}$. The sum is given by

$$\sum_{k\in\mathcal{T}}\beta_kr_{kj}(t)\left\{\sum_{i\in H}u_{ji}+\sum_{\substack{i\in\mathcal{T}\\ i\neq j}}u_{ji}(1-\rho_{ij})\right\}dt+o(dt)$$

(5.1)

$$=\sum_{k\in\mathcal{T}}\beta_kr_{kj}(t)\left\{-u_{jj}-\sum_{\substack{i\in\mathcal{T}\\ i\neq j}}u_{ji}\rho_{ij}\right\}dt+o(dt),$$

since $\sum_{i\in H}u_{ji}=-\sum_{i\in\mathcal{T}}u_{ji}$.

The backward equation $r'_{jj}(t)=\sum_{i\in\mathcal{T}}u_{ji}r_{ij}(t)$ can be immediately integrated to give

$$-1=\sum_{i\in\mathcal{T}}u_{ji}G_{ij}$$

whence, since $\rho_{ij}=G_{ij}G_{jj}^{-1}$ $(i\neq j)$,

(5.2)
$$-u_{jj}-\sum_{\substack{i\in\mathcal{T}\\ i\neq j}}u_{ji}\rho_{ij}=G_{jj}^{-1}.$$

Combining (5.1) and (5.2) gives

$$P(t<L_j<t+dt)=G_{jj}^{-1}\sum_{k\in\mathcal{T}}\beta_kr_{kj}(t)\,dt+o(dt).$$

Finally,

$$P(L_j<\infty)=P(\{j\}\text{ ever visited})=\beta_j+\sum_{\substack{k\in\mathcal{T}\\ k\neq j}}\beta_k\rho_{kj}=G_{jj}^{-1}\sum_{k\in\mathcal{T}}\beta_kG_{kj},$$

and the assertion follows.

We shall now establish a connection between the age distribution and the last-exit distribution as follows. Let $\mathcal{X}$ be as defined at the start of this section, and let $\mathcal{Y}$ be the corresponding return process, assumed to possess the properties above. Let $\bar{\mathcal{Y}}$ be the dual process whose transition semi-group is given by

(5.3)
$$\bar{p}_{ij}(t)=e^{\rho t}p_{ij}(t)x_j/x_i,$$

and $\bar{\mathscr{X}}$ be obtained by stopping $\bar{\mathscr{Y}}$ when it hits H. The generator of $\bar{\mathscr{Y}}$ is $q_{ij}x_j/x_i$ $(i \neq j)$, and $q_{ii}+\rho$ if $i=j$, and hence that of $\bar{\mathscr{X}}$ is $u_{ij}x_j/u_i$ $(i \neq j)$, and $u_{ii}+\rho$ if $i=j$. If the transition semi-group of $\bar{\mathscr{X}}$ is $[\bar{r}_{ij}(t),\ t \geqq 0,\ i, j \in \mathscr{T}]$ then it follows as in the proof of Theorem 3 that

$$\bar{r}_{ij}(t) = e^{\rho t} r_{ij}(t) x_j/x_i \qquad i, j \in \mathscr{T}$$

and

$$\bar{r}_{ij}(t) = (x_j/x_i) \int_0^t \sum_{k \in \mathscr{T}} e^{\rho s} r_{ik}(s) u_{kj}\, ds, \qquad i \in \mathscr{T}, \qquad j \in H.$$

Let $\bar{L}_j$ be the last-exit time from $\{j\}$ for $\bar{\mathscr{X}}$. Then its distribution function is given by

$$\bar{\Lambda}_j(t) = A^{-1} \sum_{i \in \mathscr{T}} (\beta_i/x_i) \int_0^t e^{\rho s} r_{ij}(s)\, ds$$

where $A = \sum_{i \in \mathscr{T}} (\beta_i/x_i) G_{ij}(\rho) < \infty$, and $\{\beta_i,\ i \in \mathscr{T}\}$ is the initial distribution for $\bar{\mathscr{X}}$.

Now choose $\{\beta_i\}$ so that $\beta_i = Bx_i \sum_{k \in H} m_k q_{ki}$ $(i \in \mathscr{T})$, where B is chosen so that $\sum_{i \in \mathscr{T}} \beta_i = 1 = B \sum_{i \in \mathscr{T}} x_i \sum_{k \in H} q_{ki} m_k$. The invariance equation for $\{x_i\}$ shows that $\sum_{i \in \mathscr{T}} q_{ki} x_i = -\rho x_k - \sum_{i \in H} q_{ki} x_k$ (cf. Tweedie (1974), Proposition 2), and by construction of $\mathscr{Y}$ we have $q_{ki} = -\delta_{ki} q_i$ $(i, k \in H)$, whence

$$B^{-1} = \sum_{k \in H} m_k x_k (q_k - \rho) > 0, \tag{5.4}$$

and the required $\{\beta_i\}$ can indeed be chosen. With this choice of $\{\beta_i\}$ it follows that $a_j(t) = \bar{\Lambda}_j(t)$, and hence we have the following result.

Theorem 5. If the dual absorbing process $\bar{\mathscr{X}}$ has the initial distribution defined by $\beta_i = B \sum_{k \in H} m_k q_{ki} x_i$, where B is given in (5.4), then the last exit distribution of $\bar{\mathscr{X}}$ is just the conditional age distribution given in the statement of Theorem 1.

Remarks. If $\rho = 0$, then $x_i \equiv 1$, $\bar{\mathscr{X}} \overset{\mathscr{D}}{=} \mathscr{X}$, and the initial distribution of $\bar{\mathscr{X}}$ prescribed by Theorem 5 is

$$\beta_i = \sum_{k \in H} m_k q_k p_{ki} \Big/ \sum_{k \in H} m_k q_k, \qquad i \in \mathscr{T}.$$

An important special case is that in which the return distribution sends $\mathscr{Y}$ to a fixed state $\{a\} \in \mathscr{T}$, no matter where $\mathscr{Y}$ hits H. If $\rho = 0$ this covers the situation in Pakes (1979) where $H = \{0\}$, $a = 1$. In this case the last exit distribution of $\mathscr{X}$, when started at $\{a\}$ coincides with the conditional age distribution. For a similar property in the context of diffusion processes, see Nagasawa and Maruyama (1979).

6. Examples

(a) *Finite birth-death processes.* For a variety of genetic models, we can take $\mathcal{S}=\{0, 1, \cdots, N\}$, $H=\{0, N\}$ and $\rho=0$. It is perhaps worth highlighting some of the previous results for this case. Let $\mathcal{X}=\{X(t), t\geqq 0\}$ describe the evolution of the number of A-alleles in a population of size N at time t. The probabilities $\tilde{b}_{jl}$ of (4.7) can be used to determine the probability that the A-allele is the oldest. The restarting distribution is given by $p_{01}=1$, $p_{0k}=0$ $(k\neq 1)$ and $p_{N,N-1}=1$, $p_{Nk}=0$ $(k\neq N-1)$. Then the (limiting) probability that allele A is the oldest is given by

$$\begin{aligned}\tilde{b}_{jN} &= m_N m_j^{-1} q_N G_{N-1,j}\\ &= m_N q_N G_{N-1,j}/\{m_N q_N G_{N-1,j}+m_0 q_0 G_{1j}\}\\ &= b_{1N} G_{N-1,j}/\{b_{1N} G_{N-1,j}+b_{N-1,0} G_{1j}\}\end{aligned} \tag{6.1}$$

where $b_{jl}=P_j(X(T_H)=l)$, $l\in H$, $j\in\mathcal{T}$ and $G_{ij}=G_{ij}(0)$. If the return process $\mathcal{Y}$ is reversible then $\tilde{b}_{jN}=b_{jN}$. An alternative approach to the discrete-time version of (6.1) can be found in Levikson (1977).

Finally, to complete the connection between the age distribution and the last-exit time described by Theorem 5, we note that the initial distribution $\{\beta_i, i\in\mathcal{T}\}$ specified there is given by $\beta_1=m_0q_0/\{m_0q_0+m_Nq_N\}$ and $\beta_{N-1}=m_Nq_N/\{m_0q_0+m_Nq_N\}$; these can be simplified to $\beta_1=b_{N-1,0}/\{b_{N-1,0}+b_{1N}\}$, $\beta_{N-1}=b_{1N}/\{b_{N-1,0}+b_{1N}\}$. If $b_{j0}=1-jN^{-1}=1-b_{jN}$, then $\beta_1=\beta_{N-1}=\frac{1}{2}$.

In these applications it is often the case that $\mathcal{Y}$ is a birth–death process and then the return process is a birth–death process iff $p_{01}=p_{N,N-1}=1$. If this is the case,

$$m_j/q_0m_0=\frac{\lambda_1\cdots\lambda_{j-1}}{\mu_1\cdots\mu_j}\qquad (j=1,\cdots,N)$$

where $\lambda_1,\cdots,\lambda_N>0$ are the birth parameters of $\mathcal{X}$ and $\mu_0,\cdots,\mu_{N-1}>0$ are its death parameters.

More generally if $p_{0j}=\delta_{aj}$, $p_{Nj}=\delta_{bj}$ where, $0<a$, $b<N$ and $a\neq b$, we have

$$a_j'(t)=(m_0q_0r_{aj}(t)+m_Nq_Nr_{bj}(t))/m_j$$

and if $\sigma_{jj}=1/\mu_j$ and $\sigma_{ij}=(\lambda_i\cdots\lambda_{j-1})/(\mu_i\cdots\mu_j)$, $i<j$, then, when for example $a<b$,

$$\begin{aligned} m_j/m_0q_0 &= \sum_{i=1}^{j}\sigma_{ij} \qquad (1\leqq j\leqq a),\\ m_j/m_0q_0 &= \sum_{i=1}^{a}\sigma_{ij} \qquad (a<j\leqq b),\end{aligned} \tag{6.2}$$

$$m_j/m_0q_0=\sum_{i=1}^{a}\sigma_{ij}-(m_Nq_N/m_0q_0)\sum_{i=b+1}^{N-1}\sigma_{ij}\qquad (b<j\leqq N-1) \tag{6.3}$$

and

$$m_N q_N / m_0 q_0 = \left[\lambda_{N-1} \sum_{i=1}^{a} \sigma_{i,N-1}\right]\left[1+\lambda_{N-1} \sum_{i=b+1}^{N-1} \sigma_{i,N-1}\right]^{-1}. \tag{6.4}$$

The expressions are derived in the usual way, viz., by successively adding the equations $\sum m_i q_{ij} = 0$ and iteratively solving the resulting difference equations. A similar procedure may be adopted if $b < a$.

If $a = b$ then

$$a_j'(t) = ((m_0 q_0 + m_N q_N)/m_j) r_{aj}(t) = G_{aj}^{-1} r_{aj}(t)$$

and the m_j are given by Equations (6.2)–(6.4) above with $a = b$. This density function is similar to that of the limiting age distribution for a process with a single absorbing state. Indeed if the two absorbing states of $\mathcal{Y}$ are amalgamated and returns to $\{a\}$ occur with rate $q_0 + q_N$, then the two return processes have the same limiting conditional age distribution. This follows since the limiting age distribution is specified by $[u_{ij}]$.

As a specific example consider the continuous-time version of Moran's model for the number of A-alleles in a population of alleles at a single diallelic locus in a haploid population of fixed size N; see Karlin and McGregor (1962). This is a finite birth and death process for which

$$\lambda_i = \frac{i(N-i)}{N}\alpha_2, \qquad \mu_i = \frac{i(N-i)}{N}\alpha_1 \qquad (0 \leqq i \leqq N)$$

where α_1, α_2 are positive constants which reflect selective differences between the alleles. For simplicity we shall let $\alpha_1 = \alpha_2 = \alpha$—there are no selective differences. Our formulation differs from that of Karlin and McGregor because we assume that the rate of birth–death events is proportional to the population size, whereas they assume this rate is independent of N.

It is not difficult to show that

$$G_{ij} = (N-i)/\alpha(N-j) \quad (1 \leqq j \leqq i), \qquad = i/\alpha j \quad (i \leqq j \leqq N-1),$$

(see Tavaré (1980)) and if $a = 1$, $b = N-1$ then

$$m_j/m_0 = q_0 N/\alpha j(N-j) \quad (1 \leqq j \leqq N-1), \qquad m_N/m_0 = q_0/q_N.$$

Referring to (6.1), we find that $\tilde{b}_{jN} = j/N$. The spectral expansions developed by Karlin and McGregor (1962) may be used to write down expressions for $a_j(t)$. These are uninformative in themselves and so we shall omit them. However, see Watterson (1976) for the discrete-time analogue. Substituting the quantities above into (4.6) and using a little algebra shows that

$$\mu_j^{(1)} = \alpha^{-1}\left[j \sum_{k=j+1}^{N-1} k^{-1} + (N-j) \sum_{k=N-j}^{N-1} k^{-1}\right].$$

Asymptotic expressions may be derived for large populations by allowing $N \to \infty$ and j to behave in various ways. Thus if $j = o(N)$ as $N \to \infty$ then

$$\mu_{N-j}^{(1)} = \mu_j^{(1)} \sim (j/\alpha) \log N,$$

but if $j \sim \theta N$, $0 < \theta < 1$, then

$$\mu_j^{(1)} \sim (N/\alpha)(\theta \log \theta^{-1} + (1-\theta) \log (1-\theta)^{-1}).$$

The expression on the right is just that obtained for the diffusion approximation of the hitting distribution of H for the absorbing process.

If $a = b$ we obtain

$$\alpha\mu_j^{(1)} = \begin{cases} \dfrac{N-a}{a} \displaystyle\sum_{k=1}^{a-1} \frac{k}{N-k} + j - a + \frac{j}{N-j} \sum_{k=j}^{N-1} \frac{N-k}{k} & (a < j) \\ \dfrac{N-j}{j} \displaystyle\sum_{k=1}^{j-1} \frac{k}{N-k} + a - j + \frac{a}{N-a} \sum_{k=a}^{N-1} \frac{N-k}{k} & (a \geqq j) \end{cases}$$

and if $a \sim Np$, $j \sim N\theta$ as $N \to \infty$ where $0 < p$, $\theta < 1$ then

$$\alpha\mu_j^{(1)} \sim \begin{cases} -N\left[\dfrac{\theta \log \theta}{1-\theta} + 1 + \dfrac{(1-p) \log (1-p)}{p}\right] & (p < \theta) \\ -N\left[\dfrac{p \log p}{1-p} + 1 + \dfrac{(1-\alpha) \log (1-\alpha)}{\alpha}\right] & (\theta < p). \end{cases}$$

This expression for the case $p < \theta$ was derived by Kimura and Ohta (1973), Equation (11), as the mean age of the approximating diffusion process.

(b) *Markov branching process.* Let $\{X(t)\}$ be the Markov branching process whose generator is $u_{ij} = \nu i p_{j-i+1}$ $(i \neq j)$ and $u_{ii} = -\sum_{k \neq i} u_{ik}$ where $\nu > 0$, $p_j \geqq 0$, $\sum_{j \geqq 0} p_j = 1$, $0 < p_0 < 1$ and $p_1 = 0$; see Athreya and Ney ((1972), Chapter 3). As is well known $X(t)$ represents the size of a population of individuals whose lifetimes are independent and exponentially distributed with mean ν^{-1}, and at the end of its lifetime an individual produces j progeny with probability p_j. All individuals reproduce independently. Clearly $\mathcal{T} = \mathbb{N}$ is irreducible and $H = \{0\}$ is accessible from $\mathcal{T}$ since $p_0 > 0$. Let $f(s) = \sum_{j \in \mathcal{S}} p_j s^j$. Regularity of $\{X(t)\}$ is equivalent to the condition $\int_{1-\varepsilon}^{1} ds/(f(s) - s) = \infty$ for each ε in $(0, 1-q)$ where q is the probability of eventual extinction when $X(0) = 1$ and is the least positive root of $f(s) = s$. We always assume this condition. Let $m = f'(1-)$.

We define a return process $\{Y(t)\}$ by setting $\lambda = q_0$, $h_j = p_{0j}$ $(j \in \mathcal{T})$ and $\sum h_j = 1$. The return process is just the Markov branching process with a state-dependent immigration component as defined by Stewart (1976) and Yamazato (1975). Pakes (1979) considered the age distribution for the special case $h_1 = 1$.

When $m \leqq 1$ Stewart (1976) and Yamazato (1975) have shown that the return process is recurrent and hence if it possesses the strong ratio limit property and property (2.4) then the limiting age distribution exists for any initial distribution and

$$a_j(t)=\Big(\sum_{i\geq 1} h_i \int_0^t r_{ij}(u)\,du\Big)\Big/\Big(\sum_{i\geq 1} h_i G_{ij}\Big).$$

When $m<1$ the condition $\sum h_j \log j<\infty$ is necessary and sufficient for the positive recurrence of $\mathcal{Y}$, whence the strong ratio limit property. The following lemma gives a sufficient condition for the null recurrent case.

Lemma 3. Let $m<1$ and $\sum h_j \log j=\infty$. The return process has the strong ratio limit property if

$$G(x)=1-h(1-e^{-x})=1/x^{\delta}L(x)$$

where $\frac{1}{2}<\delta\leq 1$ and L is slowly varying at ∞.

Proof. Let $F_{00}(t)$ be the distribution function of $\inf\{t: Y(t)=0,\ t>\Delta_0\}$ where Δ_0 is the hitting time of $\mathbb{N}$ from $\{0\}$. Clearly

$$\begin{aligned}F_{00}(t)&=\lambda \sum_{j\geq 1}\int_0^t h_j r_{j0}(t-u)e^{-\lambda u}\,du\\&=\lambda\int_0^t h(r_{10}(t-u))e^{-\lambda u}\,du\end{aligned}$$

where $h(s)=\sum h_j s^j$. Moreover,

$$p_{00}(t)=\int_0^t e^{-\lambda(t-u)}\mathcal{F}_{00}(du)$$

where $\mathcal{F}_{00}$ is the renewal function generated by F_{00}.

Now

$$1-F_{00}(t)=\lambda\int_0^t e^{-\lambda u}G[-\log(1-r_{10}(t-u))]\,du+e^{-\lambda t}$$

and $1-r_{10}(t)\sim\text{const.}\,e^{-\nu(1-m)t}$. It is easy to show from these that $1-F_{00}(t)\sim(\text{const.})/(t^{\delta}L(t))$ (see Pakes (1979), p. 287) and Erickson's (1970) key renewal theorem shows that $\lim_{t\to\infty} m(t)p_{00}(t)$ exists and is positive, where $m(t)=\int_0^t(1-F_{00}(u))\,du$. In particular $p_{00}(t)$ is regularly varying at infinity, whence $p_{00}(t+\tau)/p_{00}(\tau)\to 1$, which is equivalent to the strong ratio limit property.

When $m=1$ the condition

$$\int_0^1 \frac{1-h(s)}{f(s)-s}\,ds<\infty$$

is necessary and sufficient for positive recurrence of $\mathcal{Y}$. The following lemma is analogous to that above and extends the treatment given by Pakes (1979).

Lemma 4. Suppose $m = 1$,

$$f(s) - s = (1-s)^{1+\delta} L((1-s)^{-1}), \qquad 1 - h(s) = (1-s)^{\nu} M((1-s)^{-1})$$

where $0 \leqq \nu \leqq \delta \leqq 1$, $\delta > 0$ and L, M are slowly varying at ∞. The return process has the strong ratio limit property if $\nu > \delta/2$.

Remark. $\mathcal{Y}$ is positive recurrent if $\nu > \delta$.

Proof. The integrated backward equation for $\mathcal{X}$, $\int_0^{r_{10}(t)} du/a(u) = t$, where $a(s) = \nu(f(s) - s)$, can be written as $V((1 - r_{10}(t))^{-1}) = \nu t$ where

$$V(t) = \int_1^t dx/x^{1-\delta} L(x).$$

It follows that $1 - r_{10}(t) = 1/t^{1/\delta}\lambda(t)$ where λ is slowly varying at ∞. Thus $t^{\nu/\delta}(1 - F_{00}(t))$ is slowly varying at infinity and now the proof proceeds as before.

As an aside we mention that the recurrence-time distribution of $\{0\}$ for the return process constructed from the Bienaymé–Galton–Watson process is given by $f_{00}^{(n)} = h(f_{n-1}) - h(f_{n-2})$, where f_n is the nth iterate of f evaluated at 0; see Pakes (1971). It is quite easy to place conditions on h which ensure that $p_{00}^{(n+1)}/p_{00}^{(n)} \to 1$ (Garsia (1963)) and hence obtain the strong ratio limit property under fewer restrictions than we have been able to achieve for the continuous-time case. These discrete-time conditions cannot be taken over to the present case since the discrete skeletons of $\mathcal{Y}$ are not chains of the form considered by Pakes (1971). This may be contrasted to the situation for the Markov branching process with unrestricted immigration: its discrete skeletons are the simple branching processes allowing immigration (Pakes (1975)).

It is obvious that (2.4) is satisfied when $\mathcal{Y}$ is positive recurrent. We shall not address this question in other cases and assume, say, that $\{c_i\}$ has a finite support.

When $m > 1$, Stewart (1976) has shown that the function

$$g(\theta) = \theta\left[1 - \lambda \int_0^{\infty} e^{-\theta t}(1 - h(r_{10}(t))\, dt\right]$$

exists and is strictly increasing for $\theta > d = \max\,(-\lambda(1 - h(q)), a'(q))$ and there exists $\rho \in (0, -d)$ such that $g(-\rho) = 0$. He also shows that $\mathcal{Y}$ is ρ-positive and hence the limiting conditional age distribution always exists if $m > 1$.

When $h_1 = 1$, Pakes (1979) made use of local limit theorems for the $r_{1j}(t)$ to obtain limit theorems for the limiting age distribution as $j \to \infty$. His result for $m = 1$ extends in a fairly straightforward manner, as we shall now demonstrate.

Theorem 6. Let $m = 1$, $\sum p_j j^2 \log^+ j < \infty$, $\sigma = \nu f''(1-)/2$ and $\beta = \sum h_i i < \infty$.

Then

$$\lim_{j\to\infty} P(A(j)\leqq jt)=\exp(-1/\sigma t) \qquad (t>0).$$

Remark. The same distribution has been obtained by Watterson (1977) from a hitting distribution of a diffusion approximation for critical branching processes. Presumably the critical branching process with state dependent immigration has a recurrent diffusion approximation. Since this diffusion is reversible, it follows that the limiting conditional age distribution, when j is large, will behave like that of the extinction time of the corresponding critical branching process having a large number of initial ancestors. It is obvious that extreme value distributions are the appropriate weak limits of the extinction time.

Proof. First observe that the conditions of Lemma 4 are satisfied under the conditions of the theorem. Pakes (1979) proved that

$$\nu j G_{ij} = v_{j+1}+\cdots+v_{(j-i)^+} \qquad (i,j\geqq 1)$$

where $\{v_j\}$ is the renewal sequence corresponding to the 'lifetime' distribution whose generating function is $(1-f(s))/(1-s)$. In particular $v_j\to\alpha^{-1}$ $(j\to\infty)$, where $\alpha=f''(1-)/2$. The representation given above for the Green's functions yields

$$\nu j\sum h_i G_{ij}=\sum_{i=1}^{j} h_i(v_j+\cdots+v_{j-i})+v_j\sum_{i>j} h_i.$$

The second term on the right $\to 0$ as $j\to\infty$ and the key renewal theorem shows that the first $\to\beta/\alpha$. Thus $j\sum h_iG_{ij}\to\beta/\sigma$.

Now consider

$$j\sum_{i\geqq 1} h_i\int_{jt}^{\infty} r_{ij}(u)\,du=\sum_{i\geqq 1} h_i\int_t^{\infty} j^2 r_{ij}(jy)\,dy.$$

A result of Kesten, Ney and Spitzer ((1966), (4.22)) enables us to infer that as $j\to\infty$,

$$j^2 r_{ij}(jy)=i[j(1-r_{10}(jy))/\sigma y][r_{10}(jy)]^{i-1}e^{-1/\sigma y}$$
$$+y^{-1}\varepsilon(jy)\min(i/y,j)+O\{(j/y)[1-(r_{10}(jy))^i-i(r_{10}(jy))^{i-1}(1-r_{10}(jy))]\}$$

for all $i\geqq 1$ and $y\in[t,\infty)$, $t>0$, and where $\varepsilon(t)\to 0$ as $t\to\infty$.

Given $\delta>0$, then for all j sufficiently large,

$$\sum h_i\int_t^{\infty} y^{-1}\varepsilon(jy)\min(i/y,j)\,dy\leqq\delta\sum h_i\int_t^{\infty} y^{-1}\min(i/y,j)\,dy$$

$$\to\delta\beta\int_t^{\infty} y^{-2}\,dy \qquad (j\to\infty),$$

by monotone convergence. It follows that, as $j \to \infty$,

$$j \sum h_i \int_{jt}^{\infty} r_{ij}(u)\,du = \int_t^{\infty} h'(r_{10}(jy)) j(1-r_{10}(jy))(\sigma y)^{-1} e^{-1/\sigma y}\,dy$$
$$+ O\left\{\int_t^{\infty} j[1-h(r_{10}(jy))-(1-r_{10}(jy))h'(r_{10}(jy))]\,dy/y\right\} + o(1).$$

Since $j(1-r_{10}(jy)) \to 1/\sigma y$, it is now clear that

$$j \sum h_i \int_{jt}^{\infty} r_{ij}(u)\,du \to \beta \int_t^{\infty} e^{-1/\sigma y}(\sigma y)^{-2}\,dy$$
$$= (\beta/\sigma)(1-e^{-1/\sigma t}),$$

and the theorem now follows.

When $m \neq 1$ and $h_1 = 1$, Pakes (1979) used the local limit theorem for the supercritical Markov branching process to obtain limit theorems for the limiting age, $A(j)$, as $j \to \infty$. The limit distributions which arose involved the density of the strong limit of a supercritical Markov branching process, that is, it is a function of the detailed structure of the offspring distribution $\{p_j\}$. The following simple examples indicate that when we use a general return distribution, the limit distributions of $A(j)$ also depend on the structure of $\{h_j\}$.

We shall consider the degenerate case where $\mathscr{X}$ is the linear death process; $p_0 = 1$. Then, if $i \geqq 1$,

$$r_{ij}(t) = \binom{i}{j} e^{-\nu jt}(1-e^{-\nu t})^{i-j} \quad (j \leqq i), \qquad = 0 \quad (j > i)$$

and

$$G_{ij} = 1/\nu j \quad (j \leqq i), \qquad = 0 \quad (j > i).$$

The limiting age distribution function is

$$a_j(t) = \left\{\nu j \sum_{i \geqq j} h_i \binom{i}{j} \int_0^t e^{-\nu jy}(1-e^{-\nu y})^{i-j}\,dy\right\} \Big/ \sum_{i \geqq j} h_i$$
$$= \left\{\sum_{i \geqq j} h_i \binom{i}{j} \int_0^{j(1-\exp(-\nu t))} (1-x/j)^{j-1}(x/j)^{i-j}\,dx\right\} \Big/ \sum_{i \geqq j} h_i. \tag{6.5}$$

Suppose first that the return distribution has a Poisson tail:

$$h_j \sim \text{const.}\ \delta^j/j! \qquad (j \to \infty),$$

and without loss of generality we can take the constant as unity. In this case $\sum_{i \geqq j} h_i \sim \delta^j/j!$ and some computation shows that

$$a_j(t) \sim \int_0^{j(1-\exp(-\nu t))} (1-y/j)^{j-1} e^{\delta y/j}\,dy \qquad (j \to \infty),$$

whence

$$a_j(t/j) \to \int_0^{\nu t} e^{-y}\, dy = 1 - e^{-\nu t}.$$

Equation (6.5) can be written as

(6.6) $$a_j(t) = \left\{ \int_0^{j(1-\exp(-\nu t))} (1 - x/j)^{j-1} h^{(j)}(x/j)\, dx \right\} \Big/ j! \sum_{i \geqq j} h_i.$$

If we suppose that the h_j are eventually proportional to the negative binomial probabilities

$$\binom{N+j-1}{N-1} \zeta^j (1-\zeta)^N, \qquad 0 < \zeta < 1,$$

then it is not hard to show that

$$a_j(t) \sim \frac{\Gamma(N+j)}{\Gamma(j+1) j^{N-1}} \int_0^{j(1-\exp(-\nu t))} (1 - x/j)^{j-1} (1 - \zeta x/j)^{-N-j}\, dx$$

whence

(6.7) $$a_j(t/j) \to 1 - e^{-\nu(1-\zeta)t} \qquad (j \to \infty).$$

Some insight into these results can be obtained from the following considerations. Since both return distributions have rapidly decreasing tails, an observation at a large state $\{j\}$ means that the last return from $\{0\}$ is likely to be to a state close to $\{j\}$. For the Poisson-type distribution, $h_{j+k} = o(h_j)$, as $j \to \infty$, and $k \geqq 1$ hence we expect that

$$P(A(j) \leqq t) = P(T_j \leqq t) = 1 - \exp(-\nu j t)$$

where T_j is the sojourn time of the death process in $\{j\}$, and this is indeed consistent with the limit theorem we obtained. The negative-binomial probabilities satisfy $h_{j+k} \sim h_j \zeta^k$ as $(j \to \infty)$ and $k \geqq 1$. It follows that the conditional age distribution is close to a geometric mixture of the distribution functions of the first-passage time from $\{i\}$ to $\{j-1\}$:

$$E(e^{-\theta A(j)}) \simeq (1-\zeta) \sum_{i=0}^{\infty} \zeta^i \prod_{k=j}^{i+j} \left(\frac{\nu k}{\nu k + \theta} \right),$$

and this leads to (6.7).

These observations lead us to expect rather different limiting behaviour of $A(j)$ for long-tailed return distributions. Suppose for example that $\{h_j\}$ is defined by $h^{(a)}(s) = c(1-s)^{-1}$ for some $a \in \{2, 3, \cdots\}$, $c > 0$. Thus $h_j \sim c j^{-a}$ $(j \to \infty)$. It is easily shown that

$$a_j(t) = [(a-1) j^{a-1} (j-a)!/(j-1)!] \int_0^{j(1-\exp(-\nu t))} (1 - x/j)^{a-2}\, dx$$
$$\to 1 - e^{-\nu(a-1)t} \qquad (j \to \infty).$$

We now consider the special case of a returned pure death process defined by setting $h_N = 1$ where $N \geqq 1$. Equation (4) yields

$$a_j'(t) = \nu j \binom{N}{j} e^{-\nu j t}(1-e^{-\nu t})^{N-j}.$$

This is just the density of $V_{(N-j+1)}$ where $V_{(1)} \leqq \cdots \leqq V_{(N)}$ denotes the order statistics of N independent random variables having the exponential distribution with parameter ν. This observation allows us to list a number of limit theorems for $A(j)$ as $N \to \infty$ and with j satisfying a number of growth conditions.

Theorem 7. In the following assertions, $N \to \infty$, Λ_j has the standard gamma density $t^{j-1}e^{-t}/(j-1)!$ $(t \geqq 0)$ and $N(0, 1)$ has the standard normal distribution.

(a) $\nu A(j) - \log N \xrightarrow{\mathscr{D}} \log \Lambda_j^{-1}$.

(b) $\nu N A(N-j) \xrightarrow{\mathscr{D}} \Lambda_j$.

(c) If $0 < a < 1$, then

$$\sqrt{\frac{aN}{1-a}}\,[\nu A([aN]) + \log a] \xrightarrow{\mathscr{D}} N(0, 1).$$

(d) If $j \to \infty$ but $j/N \to 0$ then

$$\nu j^{\frac{1}{2}} A(j) - \log (N/j) \xrightarrow{\mathscr{D}} N(0, 1).$$

(e) Suppose there exists $M \in \mathbb{N}_+$ satisfying

$$j^{\frac{1}{2}}(j/N)^M \to \infty \quad \text{but} \quad j^{\frac{1}{2}}(j/N)^{M+1} \to \gamma$$

for some $0 \leqq \gamma < \infty$. Then

$$j^{\frac{1}{2}}\left\{(\nu N/j)A(N-j) - \sum_{l=0}^{M}\left(\frac{j}{N}\right)^l (1+l)^{-1}\right\} \xrightarrow{\mathscr{D}} \gamma + N(0, 1).$$

Assertions (a) and (b) are direct consequences of long established limit theorems on extreme order statistics (Galambos (1978), p. 102) and (c) is a special case of Mosteller's limit theorem; see David (1970), p. 201. By using the well-known representation of order statistics of exponentially distributed samples (Feller (1971), p. 19) we can write

$$A(j) \stackrel{\mathscr{D}}{=} \sum_{i=j}^{N} V_i/i \tag{6.8}$$

where the V_i are independent and have density $\nu e^{-\nu t} I(t \geqq 0)$. Assertion (b) follows by inspection. Assertion (d) was suggested by Hall's (1979) Theorem 3, but our special assumptions permit the use of a simpler centering and norming.

Both (d) and (e) are easy to prove directly from (7) by using moment generating function techniques.

(c) *A linear birth–death process with killing.* Let $\mathcal{Y}$ be a linear birth–death process on $\mathcal{S}=\mathbb{N}_+\cup\{\Delta\}$ where Δ is a fictitious state into which the process can be absorbed from any state of $\mathcal{T}=\mathbb{N}$. The generator of $\mathcal{Y}$ is $u_{ii}=-(\lambda+\mu+\gamma)i$, $u_{i,i+1}=\lambda i$, $u_{i,i-1}=\mu i$, $u_{i\Delta}=\gamma i$ $(i\in\mathbb{N})$ and $u_{ij}=0$ for all other $i,j\in\mathcal{S}$. The process can simply vanish from $\mathbb{N}_+$ and is then regarded as having being absorbed into $\{\Delta\}$. Processes of this type occur as approximations to certain Markov chain models in population genetics, see Karlin and Tavaré (1980).

Let

$$a(s)=\lambda s^2-(\lambda+\mu+\gamma)s+\mu \qquad (s>0).$$

The generating function $F_i(s,t)=\sum_{j=0}^{\infty} r_{ij}(t)s^j$ satisfies the forward equation

$$\frac{\partial F_i}{\partial t}=a(s)\frac{\partial F_i}{\partial s} \quad \text{and} \quad F_i(s,0)=s^i \qquad (i\in\mathbb{N}),$$

whose solution is

$$F_i(s,t)=\left\{\frac{(\mu/\lambda)(1-e^{-\delta t})-s(s_0-s_1e^{-\delta t})}{s_1-s_0e^{-\delta t}-s(1-e^{-\delta t})}\right\}^i$$

where $0<s_0<1<s_1$ are the roots of $a(s)=0$ and $\delta=\lambda(s_1-s_0)>0$. The Green functions are given by

$$G_{ij}=\begin{cases}(\delta j)^{-1}s_0^i(s_0^{-j}-s_1^{-j}) & (j=1,2,\cdots,i)\\ (\delta j)^{-1}s_1^{-j}(s_1^i-s_0^i) & (j\geqq i).\end{cases}$$

Interest in the genetic problem focuses attention on the return process for which $p_{01}=p_{\Delta 1}=1$. In this case

$$a_j(t)=G_{1j}^{-1}\int_0^t r_{1j}(y)\,dy$$

which can be explicitly computed by using the expressions above.

It is of some interest to observe here that the return process is not reversible, thus the expected hitting time T_H of $\{0,\Delta\}$, given by

$$E_j(T_H)=\sum_{k=1}^{\infty}G_{jk}=\delta^{-1}\left[(s_0^j-s_1^j)\log(1-s_1^{-1})+\sum_{k=1}^{j}k^{-1}[s_0^{j-k}-s_1^{j-k}]\right],$$

differs from

$$E(A(j))=-\delta^{-1}\left[((s_1/s_0)^j-1)\log(1-s_0/s_1)+\sum_{k=1}^{j}k^{-1}[1-(s_0/s_1)^{k-j}]\right].$$

The case of interest to Karlin and Tavaré (1980) was a model describing the time to formation of a particular genotype in a population undergoing selfing. In their case, $\lambda = \frac{1}{2}$, $\mu = 1$ and $\gamma = \frac{1}{4}$ giving $s_0 = 0\cdot7142$, $s_1 = 2\cdot7808$ and $\delta = 1\cdot0308$. Thus $E_1(T_H) = 0\cdot891$, whereas the mean age is $E(A(1)) = 0\cdot833$.

It is quite straightforward to show that

$$\lim_{j\to\infty} P(\delta A(j) - \log j \leqq t) = \exp\left(-(1 - s_0/s_1)e^{-t}\right).$$

Comparing this result with that for the subcritical linear birth–death process (Pakes (1979), §6.4) shows that with respect to the age distribution, the killed process behaves like the ordinary subcritical linear birth–death process.

References

Athreya, K. B. and Ney, P. E. (1972) *Branching Processes.* Springer-Verlag, Berlin.

Chung, K. L. (1967) *Markov Chains with Stationary Transition Probabilities,* 2nd edn. Springer-Verlag, Berlin.

David, H. A. (1970) *Order Statistics,* Wiley, New York.

Erickson, K. B. (1970) Strong renewal theorems with infinite mean. *Trans. Amer. Math. Soc.* **151,** 263–291.

Feller, W. (1971) *An Introduction to Probability Theory and its Applications,* Vol. 2, 2nd edn. Wiley, New York.

Galambos, J. (1978) *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York.

Garsia, A. M. (1963) Some Tauberian theorems and the asymptotic behaviour of probabilities of recurrent events. *J. Math. Anal. Appl.* **7,** 146–162.

Hall, P. (1979) On the relative stability of large order statistics. *Math. Proc. Camb. Phil. Soc.* **86,** 467–475.

Karlin, S. and McGregor, J. (1962) On a genetics model of Moran. *Prob. Camb. Phil. Soc.* **58,** 299–311.

Karlin, S. and Tavaré, S. (1980) Non-random mating and its effect on the time to production at particular genotypes in finite populations. In preparation.

Kesten, H., Ney, P. and Spitzer, F. (1966) The Galton–Watson process with mean one and finite variance. *Theory Prob. Appl.* **11,** 513–540.

Kimura, M. and Ohta, T. (1973) The age of a neutral mutant persisting in a finite population. *Genetics* **75,** 199–212.

Kingman, J. F. C. (1963) The exponential decay of Markov transition probabilities. *Proc. Lond. Math. Soc.* (3) **13,** 337–358.

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

Nagasawa, M. and Maruyama, T. (1979) An application of time reversal of Markov processes to a problem of population genetics. *Adv. Appl. Prob.* **11,** 457–478.

Nummelin, E. (1979) Strong ratio limit theorems for ϕ-recurrent Markov chains. *Ann. Prob.* **7,** 639–650.

Orey, S. (1971) *Lecture Notes on Limit Theorems for Markov Chain Probabilities.* Van Nostrand-Reinhold Mathematical Studies No. 34, London.

Pakes, A. G. (1971) A branching process with a state dependent immigration component. *Adv. Appl. Prob.* **3,** 301–314.

Pakes, A. G. (1975) On Markov branching processes with immigration. *Sankhyā* A **37,** 129–138.

Pakes, A. G. (1979) The age of a Markov process. *Stoch. Proc. Appl.* **8,** 277–303.

STEWART, D. B. (1976) Branching Processes with State-dependent Immigration. Unpublished Ph.D. thesis, Monash University.

TAVARÉ, S. (1979) Some Results for Markov Processes with Application to Genetic Models. Unpublished Ph.D. thesis, University of Sheffield.

TAVARÉ, S. (1980) Time reversal and age distributions I. Discrete Markov chains. *J. Appl. Prob.* **17,** 33–46.

TWEEDIE, R. L. (1974) Some ergodic properties of the Feller minimal process. *Quart. J. Math. Oxford.* **25,** 485–495.

WATTERSON, G. A. (1976) Reversibility and the age of an allele. I. Moran's infinitely many neutral alleles model. *Theoret. Popn Biol.* **10,** 139–253.

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.

YAMAZATO, M. (1975) Some results on continuous time branching process with state-dependent immigration. *J. Math. Soc. Japan.* **17,** 479–496.