

RANDOM PERMUTATIONS AND NEUTRAL EVOLUTION MODELS

Paul JOYCE and Simon TAVARÉ

Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113, USA

Received 29 June 1989

Revised 28 December 1989

Permutation-valued Markov processes provide a convenient way to describe the genealogical structure of certain population models that allow immigration or mutation. Distinct cycles of the permutation correspond to binary branching trees that describe relationships among members of a particular family (or copies of an allele in the genetics setting), and the ordering of the cycles corresponds to families (or alleles) in the order of their appearance in the population. Building on the simple combinatorial structure of the Yule process with immigration, we describe the tree-valued processes that arise from linear birth and death processes and a population genetics model of Moran. This approach simplifies and explains much of the combinatorial structure of such processes, and relates genealogical (or time-reversed) processes with those running forward in time.

infinite alleles models * random permutations * genealogy

1. Introduction

In this paper we study several permutation-valued processes that describe the familial and mutational structure of some classical models from population genetics and population growth. The recent work of Kingman (1982a, b) has made it abundantly clear that explicit study of the genealogy of individuals provides a natural and powerful way to study many types of population genetic process. In simple terms, the method involves sampling from the population at a given time, and then tracing back its ancestry to time zero. The effects of mutation are then superimposed on this structure. This technique is essentially retrospective. In contrast, in our approach the state-space of the process is extended to include ancestral and mutational relationships; the process drags its history along with it into the future. This idea leads to simple, direct methods for calculating many quantities of interest for such processes.

The results in the following sections were motivated by earlier work of Joyce and Tavaré (1987), who studied a permutation-valued formulation of the classical linear birth process with immigration. To fix the terminology and the ideas, we will describe this representation below. We begin, however, by giving our notation for permutations.

A permutation of the first n integers is a one-to-one onto function from the set $\{1, 2, \dots, n\}$ to itself. We denote by S_n the set of such permutations. We will typically write a permutation $p \in S_n$ as an *ordered* product of cycles in the following way: If $p = c_1 c_2 \cdots c_k$, then c_1 is the cycle containing the integer 1, c_2 contains the smallest integer not in c_1 and so on. For consistency with Joyce and Tavaré (1987), we will find it convenient to read permutations from right to left. For example, $p = (5\ 3\ 2\ 4\ 1)$ is interpreted as the permutation $1 \rightarrow 4 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 1$, so that $p(1) = 4$, $p(2) = 3$, $p(3) = 5$, $p(4) = 2$, $p(5) = 1$. If p and $q \in S_n$ then the composition pq of p and q is defined by $pq(l) = p(q(l))$, so that we are also composing permutations from right to left. Let $p, q \in S_n$. We say that p is conjugate to q in S_n if there exists a permutation $s \in S_n$ such that $p = sqs^{-1}$. The partition of a permutation $p \in S_n$ is a vector (a_1, a_2, \dots, a_n) where $\sum_{i=1}^n ia_i = n$ and a_i is the number of cycles in p containing i elements. Conjugacy is an equivalence relation; two permutations are in the same conjugacy equivalence class if they have the same partition. It will sometimes be convenient not to display the singleton cycles (for example the permutation $(5\ 3\ 1)(2)(4)$ might be written $(5\ 3\ 1)$), and we will sometimes write $p(k) = p_k$.

The Yule (or linear birth) process with immigration may be represented as a permutation-valued Markov process $\{\Pi(t), t \geq 0\}$. Informally, suppose that the population at time t has $m-1$ members in k families, and that $\Pi(t) = c_1 c_2 \cdots c_k \in S_{m-1}$. The next individual, m , that appears in the process is either an immigrant, in which case the set state is $c_1 \cdots c_k(m)$, or it is an offspring of the existing individual labelled j . In this case, the new state is formed by inserting the integer m in the cycle in which j belongs, immediately to the left of j . This state-space of permutations provides complete details of the family relationships through which the process has passed.

The jump-chain of $\Pi(\cdot)$ is a Markov chain $\{\tilde{\Pi}_m, m = 0, 1, 2, \dots\}$ on $S = \bigcup_{n=0}^{\infty} S_n$, with $\tilde{\Pi}_0 = (0)$, the permutation of no elements. When the birth-rate is 1 and the immigration rate is θ , the jump-chain evolves as follows. Let $\{Y_1, Y_2, \dots\}$ be a sequence of independent random variables, such that Y_m is S_m -valued and is concentrated on transpositions (cycles of length 2). Let I_m be the identity in S_m , and let $i_m: S_m \rightarrow S_{m+1}$ be in the inclusion map $i_m(p) = c_1 c_2 \cdots c_k(m+1)$ if $p = c_1 c_2 \cdots c_k$. Let $Y_1 = I_1$ and for $m > 1$, let Y_m be distributed as:

$$P[Y_m = (mk)] = 1/(m-1+\theta) \quad \text{if } 1 \leq k < m,$$

$$P[Y_m = I_m] = \theta/(m-1+\theta).$$

Then $\{\tilde{\Pi}_m, m = 1, 2, \dots\}$ is generated by:

$$\begin{aligned} \tilde{\Pi}_1 &= I_1, \\ \tilde{\Pi}_{m+1} &= i_m(\tilde{\Pi}_m) Y_{m+1}, \quad m = 2, 3, \dots \end{aligned} \tag{1}$$

Multiplying the permutation $\tilde{\Pi}_m$ on the right by the transposition Y_{m+1} puts the next individual into the process in just the right way.

It will be convenient to let $|p|$ denote the number of cycles in the permutation p , and we define $x_{(m)} = x(x+1) \cdots (x+m-1)$. We will need the distribution of $\tilde{\Pi}_n$, which is given in Joyce and Tavaré (1987):

$$P[\tilde{\Pi}_n = p] = \theta^k / \theta_{(n)} \quad \text{if } p \in S_n \text{ and } |p| = k. \tag{2}$$

This result provides an alternative method for studying the Ewens Sampling Formula (Ewens, 1972), and properties of certain urn models that arise in the study of neutral mutation in population genetics theory (Donnelly, 1986; Hoppe, 1987). A beautiful review of the interplay between these ideas appears in Ewens (1990).

In this paper, we extend the range of applicability of these ideas to several other models. In Sections 2 and 3, we describe a permutation-valued version of the infinitely-many-neutral-alleles Moran model, and derive in Section 4 a basic property of genealogical trees sampled from such a process. Section 5 studies the case of no mutation and relates the results to the coalescent (Kingman, 1982a, b). In Sections 6 and 7 we study corresponding properties of the linear birth-and-death process, both with and without immigration.

2. An S_n -valued Moran model

In Section 1 we saw that a permutation of length n describes the entire family history of the first n individuals in a birth with immigration process. A graphical representation of this family history may be described by a collection of family trees. Each time a birth occurs a branch is added to the branch of the parent. Each time an immigration occurs a new tree is started. Each permutation p gives rise to a unique set of trees. Figure 1 shows the tree associated with the permutation $p = (5\ 3\ 1)(4\ 2)$.

We use the convention that when a birth occurs and a branch of the tree splits into two, the branch to the left is that of the child and the one to the right is the branch of the parent; compare Fig. 1. Although the time scale is removed we still have the relative order in which the individuals arose in the population. One can

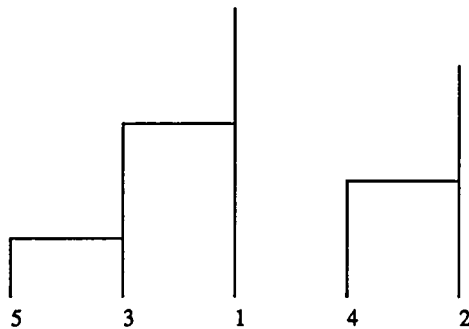


Fig. 1. Trees for birth with immigration.

therefore view the collection of trees associated with an n -permutation as a description of the genealogy of n individuals.

We may use an analogous idea to describe the genealogical and mutational structure of the infinitely-many neutral alleles Moran model in continuous time. The original version of this process may be described as follows (see Moran, 1958; Kelly, 1979). Individuals die at rate μ , and when a death occurs an individual, chosen at random from amongst the remaining $n - 1$ individuals, gives birth. The offspring is of the same type as his parent with probability $1 - u$ and is a mutant with probability u . Since the transitions in this process occur at the points of Poisson process with rate $n\mu$, the stationary distribution of the process is the same as that of its jump chain. For this reason we will concentrate on the embedded discrete-time version of the model, in which generation $k + 1$ is determined from generation k in the following way: the population has n individuals, from which one is chosen at random to die. A new individual is then born to one of the remaining $n - 1$ individuals. The individual is of the same type as his parent with probability $1 - u$ and is a mutant with probability u . The infinitely-many alleles assumption means that when a mutation occurs the mutant individual is a novel type that has not previously existed in the population.

The state space of the jump-chain will be S_n , and a sample path of the process will be a collection of permutations p_1, p_2, \dots , a permutation $p_k \in S_n$ representing the history of the n individuals alive at time k . For instance, if $n = 5$ the permutation $(531)(42)$ is interpreted the following way: The oldest allele is represented by three individuals, the second oldest by two, and the second oldest allele arose before the line of descent of the first type split; see Fig. 1 again.

We use the same convention for labelling the individuals on the tree as we did in the birth-with-immigration process but we interpret trees differently. In the birth process case the individual labelled 1 was the oldest individual in the population. In the present setting the individual labelled 1 need not be the oldest individual in the population, but he is a direct descendent of the oldest mutant.

The permutation-valued Moran model may be viewed as follows: When an individual dies remove his branch from the tree, relabel all the individuals so that the new tree (with one less branch) is consistent with our permutation description, and is described by a permutation of the first $n - 1$ integers. Then split one of the remaining branches in two or start a new tree. This is done by placing integer n to the left of his parent or in a cycle by itself to the right. Each branch is equally likely to be removed and with probability u the new branch is a mutant (which starts its own tree) or with probability $(1 - u)$ the new branch is added to one of the remaining trees. The following examples may be instructive.

Example 1. The tree corresponding to $p = (5\ 3\ 1)(4\ 2)$ appears in Fig. 1. If individual 4 is chosen to die, the resulting tree with the branch associated with 4 missing is given in Fig. 2(a). We now relabel the branches in a way that is consistent with the permutation process, giving us the permutation $(4\ 3\ 1)(2)$; see Fig. 2(b). Notice that

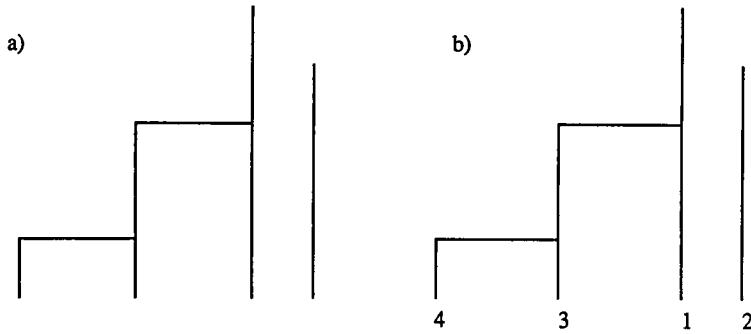


Fig. 2. Trees after a death I.

the label of individual 5 changes to 4. Suppose that the newborn individual is the nonmutant offspring of the individual currently labelled 1. We then place 5 to the left of 1. The new permutation is $(4\ 3\ 5\ 1)(2)$, and its tree is given by Fig. 3. In summary, if individual 4 dies and individual 1 has a non-mutant offspring, then the permutation $(5\ 3\ 1)(4\ 2)$ becomes $(4\ 3\ 5\ 1)(2)$.

Example 2. Again suppose that $p = (5\ 3\ 1)(4\ 2)$, but now assume that individual 1 dies. The tree with branch 1 removed and after relabelling appears in Fig. 4. Notice

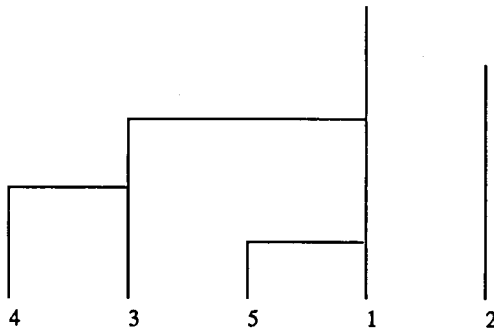


Fig. 3. Trees after a birth I.

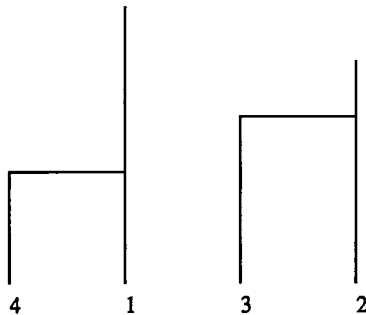


Fig. 4. Trees after a death II.

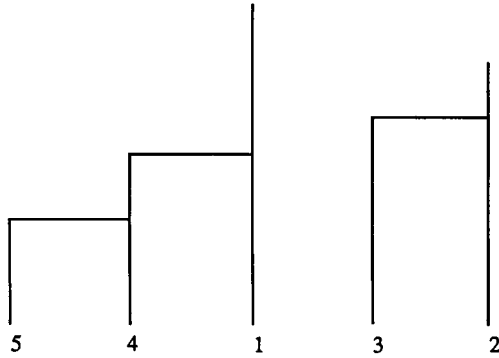


Fig. 5. Trees after a birth II.

that all individuals labelled 3 and above in Fig. 1 have their labels changed. Suppose that the individual now labelled 4 has an offspring. The new tree is given by Fig. 5. The permutation $(5\ 3\ 1)(4\ 2)$ has become $(5\ 4\ 1)(3\ 2)$.

3. Results for the Moran model

We can now describe the structure of the model in detail. We use conjugation to relabel a population described by a permutation. Conjugation will leave the family sizes the same but will change the labels of the individuals. Recall that each individual is represented by a branch on a tree, such a branch being either a 'left branch' or a 'right branch'. When a birth occurs and the tree splits by convention we make the branch to the left that of the child and the branch to the right that of the parent. Thus right branches have smaller labels than their corresponding left branches; compare Examples 1 and 2. If a left branch with label k is removed, then we rotate the labels from k to n , so that label n becomes $n-1$, $n-1$ becomes $n-2$, \dots , $k+1$ becomes k , and k becomes n . Define $s_k \in S_n$ by

$$s_k = (n\ n-1\ \dots\ k+1\ k).$$

Then $s_k^{-1} = (k+1\ k+2\ \dots\ n-1\ n\ k)$ and $s_k^{-1}ps_k$ will relabel p in the way that was described above. For example if $p = (3\ 5\ 1)(4\ 2)$, $s_3 = (5\ 4\ 3)$ then $s_3^{-1} = (4\ 5\ 3)$ and $s_3^{-1}ps_3 = (5\ 4\ 1)(3\ 2)$.

Now suppose a right branch with label k is removed from $p \in S_n$. In order to be consistent with the previous conventions, we rotate the labels from p_k up to n , so that n becomes $n-1$, \dots , p_k becomes n . The resulting permutation is $s_{p_k}^{-1}ps_{p_k}$.

The individual that dies is replaced by a new individual who always takes the last label, n . To achieve this, define $\hat{n}: S_n \rightarrow S_n$ by

$$\hat{n}(p) = (n\ p_n)p. \quad (3)$$

\hat{n} removes the n th integer from the permutation p and places it in a singleton cycle by itself. For example if $n = 5$ and $p = (5\ 2\ 1)(4\ 3)$, then $\hat{n}(p) = (5\ 1)(5\ 2\ 1)(4\ 3) = (2\ 1)(4\ 3)(5) = (2\ 1)(4\ 3)$. It is clear that

$$\hat{n}(p) = (n\ p_n)p = p(n\ p^{-1}(n)). \quad (4)$$

Define $h_k : S_n \rightarrow S_n$ by

$$h_k(p) = \hat{n}(s_r^{-1}ps_r), \quad \text{where } r = \max\{k, p_k\}.$$

The permutation $h_k(p)$ arises from p after relabelling following the death of individual k .

Suppose first that $k > p_k$. After relabelling, $n \rightarrow p_k$ in $s_k^{-1}ps_k$, implying that $\hat{n}(s_k^{-1}ps_k) = (n\ p_k)s_k^{-1}ps_k$ by (3). If $k < p_k$ then $k \rightarrow n$ in $s_{p_k}^{-1}ps_{p_k}$, so by (4), $\hat{n}(s_{p_k}^{-1}ps_{p_k}) = s_{p_k}^{-1}ps_{p_k}(k\ n)$. Hence an alternative way of writing h_k is the following:

$$h_k(p) = \begin{cases} (p_k\ n)(s_k^{-1}ps_k), & p_k < k, \\ s_{p_k}^{-1}ps_{p_k}(k\ n), & p_k > k, \\ s_k^{-1}ps_k, & p_k = k. \end{cases} \quad (5)$$

Example 3. Let $p = (5\ 3\ 1)(4\ 2)$. If $k = 4$ then $h_4(p) = (4\ 3\ 1)$ and we see from Example 1 that $h_4(p)$ is exactly the tree one gets by removing the branch labelled 4.

Example 4. Let $p = (5\ 3\ 1)(4\ 2)$ and $k = 1$. Then $h_1(p) = (4\ 1)(3\ 2)$, and we see from Example 2 that $h_1(p)$ is exactly the tree one gets from removing the branch labelled 1, after relabelling.

Now we are ready to describe the Moran model using permutations. Let Z_1, Z_2, \dots be independent and identically distributed random variables having a uniform distribution on $\{1, 2, \dots, n\}$, with

$$P[Z_m = k] = n^{-1}, \quad k = 1, \dots, n.$$

Define

$$\theta = \frac{(n-1)u}{1-u},$$

and let Y_1, Y_2, \dots be independent and identically distributed random permutations on S_n having distribution

$$\begin{aligned} P[Y_m = (n\ k)] &= 1/(n-1+\theta), \quad k = 1, \dots, n-1, \\ P[Y_m = (n\ n)] &= u = \theta/(n-1+\theta). \end{aligned} \quad (6)$$

Notice that Y_m concentrates on transpositions. Suppose further that the Y_m 's are independent of the Z_m 's. The process of interest is defined as follows:

$$X_m = h_{Z_m}(X_{m-1})Y_m, \quad m = 1, 2, \dots \quad (7)$$

Z_m is the individual chosen at random to die, and $h_{Z_m}(X_{m-1})$ is the relabelling after the death. Multiplying by $Y_m = (nk)$ on the right inserts the new individual to the left of his parent if $k < n$ (so he is not a mutant), and inserts the new individual in a cycle on his own if $k = n$, corresponding to the new-born individual being a mutant.

Clearly $\{X_m, m \geq 0\}$ is a homogeneous Markov chain on S_n , and we note that if $\theta > 0$ then the process is irreducible. The transition probabilities are given by

$$R_{pq} = P[X_m = q | X_{m-1} = p] = P[h_{Z_m}(p) = \hat{n}(q)]P[Y_m = (q_n n)]. \quad (8)$$

Let us define

$$D_{pq} = \#\{k: h_k(p) = \hat{n}(q)\}.$$

Since the Z_m 's are uniform on $\{1, \dots, n\}$ we have

$$P[h_{Z_m}(p) = \hat{n}(q)] = D_{pq}/n. \quad (9)$$

The evaluation of certain sums of the D_{pq} is crucial in what follows. In particular, it can be shown that

$$\sum_p D_{pq} = n^2 \quad (10)$$

and

$$\sum_{\{p: |p| = |\hat{n}(q)|\}} D_{pq} = n. \quad (11)$$

The proofs of these two results, which are of a technical nature, are given in the appendix. With these results in hand, we can calculate the stationary measure of the chain.

Theorem 1. *If X_1, X_2, \dots is a Markov chain on S_n defined by*

$$X_m = h_{Z_m}(X_{m-1})Y_m, \quad m = 1, 2, \dots,$$

where h_k is defined in (5), the Z_i are independent and identically distributed random variables, uniform on $\{1, 2, \dots, n\}$, independent of the Y_i , which are themselves independent with distribution given by (6), and $\theta > 0$, then the stationary distribution $r_q = P[X_m = q]$ is given by

$$r_q = \theta^k / \theta_{(n)} \quad \text{if } |q| = k.$$

Proof. Assume that $r_q = \theta^k / \theta_{(n)}$ whenever $|q| = k$. There are two cases to consider. The first case arises when q ends with a singleton cycle, that is $n \rightarrow n$ in q or $\hat{n}(q) = q$.

In this case the only way to make a transition from p to q is if the offspring is a mutant. Thus for $|\hat{n}(q)| = |q| = k$ we have

$$\begin{aligned} \sum_p r_p R_{pq} &= \sum_{|p|=k-1} r_p R_{pq} + \sum_{|p|=k} r_p R_{pq} \\ &= \sum_{|p|=k-1} \frac{\theta^{k-1}}{\theta_{(n)}} \frac{D_{pq}}{n} \frac{\theta}{n-1+\theta} \\ &\quad + \sum_{|p|=k} \frac{\theta^k}{\theta_{(n)}} \frac{D_{pq}}{n} \frac{\theta}{n-1+\theta} \quad (\text{from (6) and (8)}) \\ &= \frac{\theta^k}{\theta_{(n)}} \frac{1}{n(n-1+\theta)} \left[\sum_{|p|=k-1} D_{pq} + \theta \sum_{|p|=k} D_{pq} \right] \\ &= \frac{\theta^k}{\theta_{(n)}} \frac{1}{n(n-1+\theta)} [n^2 - n + n\theta] \quad (\text{from (10) and (11)}) \\ &= \theta^k / \theta_{(n)} = r_q. \end{aligned}$$

The second case arises when $|\hat{n}(q)| = |q| + 1 = k + 1$. In this case we have

$$\begin{aligned} \sum_p r_p R_{pq} &= \sum_{|p|=k} r_p R_{pq} + \sum_{|p|=k+1} r_p R_{pq} \\ &= \sum_{|p|=k} \frac{\theta^k}{\theta_{(n)}} \frac{D_{pq}}{n} \frac{1}{n-1+\theta} \\ &\quad + \sum_{|p|=k+1} \frac{\theta^{k+1}}{\theta_{(n)}} \frac{D_{pq}}{n} \frac{1}{n-1+\theta} \quad (\text{from (6) and (8)}) \\ &= \sum_{|p|=k} \frac{\theta^k}{\theta_{(n)}} \frac{D_{pq}}{n(n-1+\theta)} + \sum_{|p|=k+1} \frac{\theta^k}{\theta_{(n)}} \frac{D_{pq}\theta}{n-1+\theta} \\ &= \frac{\theta^k}{\theta_{(n)}} \frac{1}{n(n-1+\theta)} \left[\sum_{|p|=k} D_{pq} + \theta \sum_{|p|=k+1} D_{pq} \right] \\ &= \frac{\theta^k}{\theta_{(n)}} \frac{1}{n(n-1+\theta)} [n^2 - n + n\theta] \quad (\text{from (10) and (11)}) \\ &= \theta^k / \theta_{(n)} = r_q. \end{aligned}$$

Thus $r_q = \sum_p r_p R_{pq}$ implying that $\{r_p\}$ is the stationary distribution. \square

4. Sampling

Now suppose a sample of size r is taken from a population of size n evolving according to a stationary Moran model. It is important in practice to be able to describe the mutational history of such a sample. The natural way to do this is to randomly remove $n - r$ branches from the set of trees that describe the population, then relabel the ‘reduced set of trees’ so as to be consistent with the permutation process. Equivalently, we can remove the branches one at a time and relabel after each branch is removed.

Example 5. Suppose $n = 5$, $r = 3$, and $p = (5\ 3\ 1)(4\ 2)$ (see Fig. 1 again). Let us first remove the branch labelled 4. That is, $h_4(p) = (4\ 3\ 1)(2)(5)$. Now remove the branch

labelled 3, to obtain $h_3 \circ h_4(p) = (3\ 1)(2)(4)(5)$. Each application of an h removes another branch and relabels the tree. We want to view this as a permutation in S_3 , which we can do by projection.

Let $Z_n, Z_{n-1}, \dots, Z_{r+1}$ be independent random variables, Z_i uniformly distributed over the set $\{1, 2, \dots, i\}$. Let $\pi_j: S_j \rightarrow S_{j-1}$ be the projection map. Let $h_k^{(j)} = \pi_j \circ h_k$ where $h_k: S_j \rightarrow S_j$ defined by (5). (For ease of notation, we are implicitly using the definition in (5) with n replaced by j ; there should be no cause for confusion.) Define $H^{(r)}: S_n \rightarrow S_r$ by

$$H^{(r)} = h_{Z_{r+1}}^{(r+1)} \circ h_{Z_{r+2}}^{(r+2)} \circ \dots \circ h_{Z_n}^{(n)}. \tag{12}$$

We interpret $H^{(r)}(p)$ as the collection of trees one gets after randomly removing $n - r$ branches from p .

Theorem 2. *If X is an S_n -valued random variable with distribution*

$$P[X = p] = \theta^k / \theta_{(n)}, \quad \text{if } |p| = k,$$

if the Z_i are independent of X , and if $q \in S_r$ and $|q| = l$, then

$$P[H^{(r)}(X) = q] = \theta^l / \theta_{(r)}. \tag{13}$$

Proof. The proof is by backward induction on r . Let $r = n - 1$, and assume that $q \in S_{n-1}$ and $|q| = l$. Then

$$\begin{aligned} P[H^{(n-1)}(X) = q] &= P[h_{Z_n}^{(n)}(X) = q] \\ &= \sum_p P[h_{Z_n}^{(n)}(p) = q | X = p] P[X = p] \\ &= \sum_{|p|=l+1} P[h_{Z_n}^{(n)}(p) = q] \frac{\theta^{l+1}}{\theta_{(n)}} + \sum_{|p|=l} P[h_{Z_n}^{(n)}(p) = q] \frac{\theta^l}{\theta_{(n)}} \\ &= \sum_{|p|=l+1} \frac{D_{pq}}{n} \frac{\theta^{l+1}}{\theta_{(n)}} + \sum_{|p|=l} \frac{D_{pq}}{n} \frac{\theta^l}{\theta_{(n)}} \\ &= \frac{\theta^l}{\theta_{(n)}} \left[\frac{\theta n + n^2 - n}{n} \right] \quad (\text{from (10) and (11)}) \\ &= \theta^l / \theta_{(n-1)}. \end{aligned}$$

By induction suppose (13) is true for some $r < n - 1$. If $q \in S_{r-1}$ and $|q| = l$ then

$$\begin{aligned} P[H^{(r-1)}(X) = q] &= P[h_{Z_r}^{(r)} \circ H^{(r)}(X) = q] \\ &= \sum_{p \in S_r} P[h_{Z_r}^{(r)}(p) = q | H^{(r)}(X) = p] P[H^{(r)}(X) = p] \\ &= \sum_{p \in S_r} P[h_{Z_r}^{(r)}(p) = q] P[H^{(r)}(X) = p] \\ &= \theta^l / \theta_{(r-1)}. \end{aligned}$$

The last line follows by the same argument used in the $r = n - 1$ case and by the induction hypothesis. \square

This theorem shows that the sampled trees have the same probabilistic structure as the trees associated with the entire (stationary) population. This may be viewed

as the analog of a theorem of Kelly (1979, Theorem 7.1) for the standard representation of the Moran model. As a consequence, other sampling distributions associated with the Moran model (which arise when less detailed information about the process is recorded, such as its partition structure) follow readily by counting. For example, the cycle representation preserves the age structure of the types and so will yield an alternative derivation of Donnelly and Tavaré's (1986) age-ordered Ewens sampling formula. Further examples appear in Joyce and Tavaré (1987). The present approach also gives alternative derivations for many of the sampling results described in Hoppe (1987) and Ewens (1990).

5. The Moran model with no mutation

When $u = 0$ (so there is no mutation in the model) the permutation-valued process $\{X_m, m \geq 0\}$ is defined as in (7), but since $\theta = 0$, Y_1, Y_2, \dots are now independent and identically distributed random permutations on S_n having distribution

$$P[Y_m = (nk)] = 1/(n-1), \quad k = 1, \dots, n-1. \quad (14)$$

The interpretation of permutations as genealogical trees is just as before, but the process is no longer irreducible. From (14) it is clear that $|X_m|$ is non-increasing in m , and that any limit distribution for the process must concentrate on C_n , the set of $p \in S_n$ comprising a single cycle.

If we allow the process to start from any $p \in S_n$, then the fact that eventually it concentrates on C_n corresponds to the observation that in such a process every individual can eventually be traced back to a common ancestor. If we suppose that $X_0 \in C_n$, then $\{X_m, m \geq 0\}$ is an irreducible C_n -valued Markov chain, and the method of the proof of Theorem 1 applies immediately to show that the stationary distribution $\{r_q\}$ of the chain is given by

$$r_q = 1/(n-1)!, \quad q \in C_n. \quad (15)$$

Thus the stationary measure is uniform on C_n , reflecting the reproductive symmetry in the model.

Permutations have been used, implicitly or explicitly, to describe genealogy in several other places. Of particular relevance in the present context is the work of Harding (1971) and Tajima (1983), who implicitly used (15) to study the genealogy of neutral models. The method of this section may also be thought of as a forward-in-time version of Kingman's coalescent (Kingman, 1982a), although the information contained in the two processes is somewhat different.

6. The birth–death–immigration process

Throughout this paper we have viewed a permutation as a representation of the history of the individuals in the population. In this section we look at the permutation-valued description of the (linear) birth, death and immigration process. Like

the Moran model, a birth is represented by adding a branch to one of the remaining trees, an immigration is represented by starting a new tree, and a death is represented by removing one of the branches from an existing tree.

Let $\{\Pi(t), t \geq 0\}$ be the permutation-valued description of the birth and death process with immigration with birth rate 1, death rate μ and immigration rate θ . Let $I(t)$ be total number of individuals alive at time t . We will show that $P[\Pi(t) = p | I(t) = n] = \theta^k / \theta_{(n)}$, whenever $|p| = k$. The introduction of deaths makes the process quite a bit more complicated. In the birth with immigration process at the time of the n th transition there were n individuals alive in the process, and the entire history of the process was recorded in a single permutation. This is no longer the case with the birth-death process with immigration.

It is well known that

$$P[I(t) = n] = \frac{\theta_{(n)}}{n!} (1 - b_t)^\theta b_t^n, \quad (16)$$

where

$$b_t = \frac{\exp(1 - \mu)t - 1}{\exp(1 - \mu)t - \mu}.$$

Let $r(p, q)$ be the transition rate from p to q . Note that if $p \in S_{n+1}$, $q \in S_n$, then it follows from Lemma 8 of the appendix that $\sum_p D_{pq} = (n+1)^2$ where $D_{pq} = \#\{k: h_k^{(n+1)}(p) = q\}$. If $|q| = k$, then by Lemma 9 of the appendix,

$$\sum_{|p|=k+1} D_{pq} = n+1. \quad (17)$$

So from the above

$$\sum_{|p|=k} D_{pq} = n(n+1).$$

Theorem 3. Let $P_p(t) = P[\Pi(t) = p]$. Then

$$P_p(t) = \frac{\theta^k}{n!} (1 - b_t)^\theta b_t^n \quad \text{if } |p| = k.$$

Proof. Our method is to check the Chapman-Kolmogorov equations.

$$\begin{aligned} \sum_p r(p, q) P_p(t) &= -[n(\mu + 1) + \theta] \frac{\theta^k}{n!} (1 - b_t)^\theta b_t^n \\ &\quad + \sum_{|p|=k+1} \mu D_{pq} \frac{\theta^{k+1}}{(n+1)!} (1 - b_t)^\theta b_t^{n+1} \\ &\quad + \sum_{|p|=k} \mu D_{pq} \frac{\theta^k}{(n+1)!} (1 - b_t)^\theta b_t^{n+1} \\ &\quad + \frac{\theta^k}{(n-1)!} (1 - b_t)^\theta b_t^{n-1} \\ &= \frac{\theta^k}{n!} (1 - b_t)^\theta b_t^{n-1} \left[-(n(\mu + 1) + \theta) b_t \right. \\ &\quad \left. + \frac{\mu\theta}{n+1} b_t^2 \sum_{|p|=k+1} D_{pq} + \frac{\mu}{(n+1)} b_t^2 \sum_{|p|=k} D_{pq} + n \right]. \end{aligned}$$

From (17) it follows that the above is

$$\begin{aligned}
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} [-(n(\mu+1)+\theta)b_i + \mu\theta b_i^2 + \mu b_i^2 n + n] \\
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} [\mu n b_i (b_i-1) - n(b_i-1) + \theta b_i (\mu b_i-1)] \\
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} [n(b_i-1)(\mu b_i-1) + \theta b_i (\mu b_i-1)] \\
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} [(1-\mu b_i)(n(1-b_i) - \theta b_i)].
 \end{aligned}$$

But

$$\begin{aligned}
 P'_q(t) &= \frac{\theta^k}{n!} n(1-b_i)^\theta b_i^{n-1} b'_i - \frac{\theta^{k+1}}{n!} (1-b_i)^{\theta-1} b_i^n b'_i \\
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} b'_i \left[n - \frac{\theta b_i}{1-b_i} \right] \\
 &= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} b'_i \left[\frac{n(1-b_i) - \theta b_i}{1-b_i} \right].
 \end{aligned}$$

An easy calculation gives $b'_i = (1-\mu b_i)(1-b_i)$ so the above

$$= \frac{\theta^k}{n!} (1-b_i)^\theta b_i^{n-1} (1-\mu b_i) [n(1-b_i) - \theta b_i].$$

It follows that

$$P'_q(t) = \sum_p r(p, q) P_p(t),$$

and the proof is complete. \square

The following corollary follows easily from (16) and the previous theorem.

Corollary 4.

$$P[\Pi(t) = p | I(t) = n] = \theta^k / \theta_{(n)}, \quad \text{if } p \in S_n \text{ and } |p| = k. \quad \square$$

Note that conditional on the number of individuals alive at time t , the distribution of $\Pi(t)$ does not depend on time and has the same distribution as the birth process with immigration. See Joyce and Tavaré (1987) and Tavaré (1989) for related results.

It is straightforward to show that if the death rate is larger than the birth rate then the process has a limiting (and stationary) distribution, for if $\mu > 1$ then $b_t \rightarrow \mu^{-1}$ as $t \rightarrow \infty$. Thus by the last corollary and (16), if $p \in S_n$ and $|p| = k$ then

$$\lim_{t \rightarrow \infty} P[\Pi(t) = p] = \frac{\theta^k}{\theta_{(n)}} \lim_{t \rightarrow \infty} P[J(t) = n] = \frac{\theta^k}{n!} \left(1 - \frac{1}{\mu}\right)^\theta \left(\frac{1}{\mu}\right)^n.$$

We have therefore established that for $\mu > 1$:

Corollary 5.

$$\lim_{t \rightarrow \infty} P[\Pi(t) = p] = \frac{\theta^k}{n!} \left(1 - \frac{1}{\mu}\right)^\theta \left(\frac{1}{\mu}\right)^n \quad \text{if } p \in S_n, |p| = k. \quad \square$$

7. The birth–death process

We conclude with a short description of the linear birth–death process. Since there is no immigration, the natural state space for the birth–death process is the collection C of permutations containing a single cycle. It will once more be convenient to let C_n denote the set of $p \in S_n$ comprising a single cycle. First, let us formally take the limit as θ (the immigration parameter) goes to zero. Then from Corollary 4,

$$\lim_{\theta \rightarrow 0} P[\Pi(t) = p | I(t) = n] = \frac{1}{(n-1)!}, \quad \text{if } p \in C_n,$$

and is 0 otherwise. Since $(n-1)!$ is the number of cycles of length n , this indicates that given that n individuals are alive at time t , each of the $(n-1)!$ states of the process is equally likely, independent of time t or the death rate μ . While the lengths of the branches of the family tree for a birth–death process clearly depend on the birth and death rates, the symmetries in the history of the population are reflected in the fact that the phylogeny is independent of these rates (subject of course to the scalings used here).

Let $J(t)$ be the number of individuals alive at time t for a birth–death process. The distribution of $J(t)$ is

$$P[J(t) = n] = (1 - \mu b_t)(1 - b_t)b_t^{n-1}.$$

If we let $\{\Pi^*(t), t \geq 0\}$ be the C -valued description of the birth–death process, then it is intuitively clear that if $n > 0$ and $p \in C_n$,

$$P[\Pi^*(t) = p] = \frac{1}{(n-1)!} P[J(t) = n] = \frac{1}{(n-1)!} (1 - \mu b_t)(1 - b_t)b_t^{n-1}.$$

To give a precise proof of the above result one need only check the Chapman–Kolmogorov equations. Since the proof is almost identical to that of the birth–death with immigration case we omit it and just state the result.

Theorem 6. If $\{\Pi^*(t) : t \geq 0\}$ is the C -valued description of the birth-death process and $p \in C_n, n > 0$, then

$$P[\Pi^*(t) = p] = \frac{1}{(n-1)!} (1 - \mu b_t)(1 - b_t) b_t^{n-1}.$$

If $n = 0$, then

$$P[\Pi^*(t) = (0)] = \mu b_t. \quad \square$$

Appendix

This appendix contains proofs of the results in equations (10) and (11). The function $h_k(p)$ defined at (5) relabels the permutation p in a particular way. It will be convenient to reverse the relabelling. To do this we define a function which is in some sense an inverse to h_k .

Let $g_{ij} : S_n \rightarrow S_n$ be given by

$$g_{ij}(p) = \begin{cases} (ij)s_j[\hat{n}(p)]s_j^{-1}, & i \leq j, \\ s_i[\hat{n}(p)]s_i^{-1}(ij), & i \geq j. \end{cases} \quad (18)$$

Let $L_{pq} = \#\{(i, j) : g_{ij}(p) = q\}$. It is clear that $\sum_q L_{pq} = n^2$. We now relate g_{ij} to h_i .

Lemma 7. (a) $g_{p,i}[h_i(p)] = p$.

(b) $h_i[g_{p,i}(p)] = \hat{n}(p)$.

Proof. (a) By the definition of g_{ij} given by (18),

$$g_{ij}(h_i(p)) = \begin{cases} (ij)s_j[\hat{n}(h_i(p))]s_j^{-1}, & i \leq j, \\ s_i[\hat{n}(h_i(p))]s_i^{-1}(ij), & i \geq j. \end{cases}$$

By (5) and (18) and the fact that $\hat{n}(h_i(p)) = h_i(p)$ we get

$$g_{p,i}(h_i(p)) = \begin{cases} (i p_i)s_i[(p_i n)s_i^{-1}p s_i]s_i^{-1}, & p_i < i, \\ s_{p_i}[s_{p_i}^{-1}p s_{p_i}(n i)]s_{p_i}^{-1}(i p_i), & p_i > i, \\ s_i[s_i^{-1}p s_i]s_i^{-1}, & p_i = i. \end{cases}$$

Case 1. $p_i < i$:

$$g_{p,i}(h_i(p)) = (i p_i)s_i(p_i n)s_i^{-1}p = (i p_i)(i p_i)p = p.$$

Case 2. $p_i > i$:

$$g_{p,i}(h_i(p)) = p s_{p_i}(n i)s_{p_i}^{-1}(i p_i) = p(i p_i)(i p_i) = p.$$

Case 3. $p_i = i$:

$$s_i[s_i^{-1}p s_i]s_i^{-1} = p.$$

(b) By (5),

$$h_i(g_{p_i}(p)) = \begin{cases} (p_i n) s_i^{-1} [g_{p_i}(p)] s_i, & p_i < i, \\ s_{p_i}^{-1} [g_{p_i}(p)] s_{p_i}(n i), & p_i > i, \\ s_i^{-1} [g_{p_i}(p)] s_i, & p_i = i. \end{cases}$$

Using (5) and (18) and the fact $i \rightarrow p_i$ in $g_{p_i}(p)$,

$$h_i(g_{p_i}(p)) = \begin{cases} (p_i n) s_i^{-1} [(p_i i) s_i [\hat{n}(p)] s_i^{-1}] s_i, & p_i < i, \\ s_{p_i}^{-1} [s_{p_i} [\hat{n}(p)] s_{p_i}^{-1}(p_i i)] s_{p_i}(n i), & p_i > i, \\ s_i^{-1} [s_i [\hat{n}(p)] s_i^{-1}] s_i, & p_i = i. \end{cases}$$

Case 1. $p_i < i$:

$$h_i(g_{p_i}(p)) = (p_i n) s_i^{-1} (p_i i) s_i \hat{n}(p) = (p_i n)(p_i n) \hat{n}(p) = \hat{n}(p).$$

Case 2. $p_i > i$:

$$h_i(g_{p_i}(p)) = \hat{n}(p) s_{p_i}^{-1} (p_i i) s_{p_i}^{-1}(n i) = \hat{n}(p)(n i)(n i) = \hat{n}(p).$$

Case 3. $p_i = i$:

$$h_i(g_{p_i}(p)) = s_i^{-1} s_i [\hat{n}(p)] s_i^{-1} s_i = \hat{n}(p). \quad \square$$

Lemma 8. $\sum_p D_{pq} = n^2$.

Proof.

$$D_{pq} = \#\{k: h_k(p) = \hat{n}(q)\} = \#\{k: p = g_{p,k}(q)\} \quad (\text{by Lemma 7}).$$

Note that $j \rightarrow i$ in the permutation $g_{ij}(q)$, so if $g_{ij}(q) = p$ then $p_j = i$. Thus

$$D_{pq} = \#\{k: p = g_{p,k}(q)\} = \#\{(i, j): g_{ij}(q) = p\} = L_{qp},$$

and so $\sum_p D_{pq} = \sum_p L_{qp} = n^2$. \square

Lemma 9.

$$\sum_{\{p: |p|=|\hat{n}(q)\}} D_{pq} = n.$$

Proof. If $|p| = |\hat{n}(q)|$ then p must have a singleton cycle in order that $D_{pq} \neq 0$. The singleton can be any of the first n integers, and so

$$\sum_{\{p: |p|=|\hat{n}(q)\}} D_{pq} = n. \quad \square$$

Acknowledgements

The authors were supported in part by NSF grant DMS88-03284.

References

- P. Donnelly, Partition structures, Pólya urns, the Ewens sampling formula, and the ages of alleles, *Theoret. Pop. Biol.* 30 (1986) 271–288.
- P. Donnelly and S. Tavaré, The ages of alleles and a coalescent, *Adv. Appl. Probab.* 18 (1986) 1–19.
- W.J. Ewens, The sampling theory of selectively neutral alleles, *Theoret. Pop. Biol.* 3 (1972) 87–112.
- W.J. Ewens, Population genetics theory—the past and the future, in: S. Lessard, ed., *Mathematical and Statistical Developments of Evolutionary Theory*, Proc. NATO ASI Symp. (Kluwer, Dordrecht, The Netherlands, 1990), 177–228.
- E.F. Harding, The probabilities of rooted tree-shapes generated by random bifurcation, *Adv. Appl. Probab.* 3 (1971) 44–77.
- F.M. Hoppe, The sampling theory of neutral alleles and an urn model in population genetics, *J. Math. Biol.* 25 (1987) 123–159.
- P. Joyce and S. Tavaré, Cycles, permutations and the structure of the Yule process with immigration, *Stochastic Process. Appl.* 25 (1987) 309–314.
- F.P. Kelly, *Reversibility and Stochastic Networks* (Wiley, New York, 1979).
- J.F.C. Kingman, On the genealogy of large populations, *J. Appl. Probab.* 19A (1982a) 27–43.
- J.F.C. Kingman, The coalescent, *Stochastic Process. Appl.* 13 (1982b) 235–248.
- P.A.P. Moran, Random processes in genetics, *Proc. Cambridge Philos. Soc.* 54 (1958) 60–71.
- F. Tajima, Evolutionary relationship of DNA sequences in finite populations, *Genetics* 105 (1983) 437–460.
- S. Tavaré, The genealogy of the birth, death and immigration process, in: M. Feldman, ed., *Mathematical Evolutionary Theory* (Princeton Univ. Press, Princeton, NY, 1989) pp. 41–56.