A Mendelian Markov Process with Multinomial Transition Probabilities[1] II. Multiple Alleles and Multiple Loci.

by

R. G. Khazanie[2] and H. E. McKean

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series Number 36

February 1965

[1]Part of a Ph.D. thesis completed at Purdue University by the senior author January, 1965.

[2]Present address: Department of Mathematics, Western Illinois University, Macomb, Illinois.

A Mendelian Markov Process with Multinomial Transition Probabilities II. Multiple Alleles and Multiple Loci

by

R. G. Khazanie and H. E. McKean

Purdue University

1. Introduction

The utility of the moments of the binomial distribution in finding the n-stage transition probabilities was indicated by Khazanie and McKean (1965) In this paper, maintaining the same breeding structure as in the two-allele-case, it is proposed to extend the technique to cover the multiallelic case in general with specific emphasis on the triallelic case. The existing literature on the multiple allelic case is at best scanty (see for example Kimura (1955)). Section 6 will briefly touch on a further extension to multiple loci which segregate independently.

2. The Triallelic Case.

2.1. Formulation of the Process. In the triallelic case the $M(= 2N)$ gametes which form the population are made up of three kinds of alleles which may be denoted by A_1, A_2 and A_3. If at any time the population consists of x A_1-alleles, y A_2-alleles (and consequently, $M-x-y$ A_3-alleles) the population will be said to be in state (x, y). The evolutionary pattern is thus characterized by a Markovian process whose states are given by such pairs and the transition probabilities specified by a trinomial distribution. Since in every generation M gametes are picked and since there are three alleles the total number of states is $\binom{M+2}{2}$ as this corresponds to the number of distinguishable distributions of M objects into 3 cells.

The set $S = \left\{ (s_1,s_2): \begin{array}{l} s_1 = 0,1,\ldots,M \\ \\ s_2 = 0,1,\ldots,M \end{array} \text{ such that } s_1 + s_2 \leq M \right\}$ consisting of $\binom{M+2}{2}$ states thus constitutes the entire state space of the process. These states in S can be broken down into the following distinct classes, depending upon the kinds of alleles present in the population:

$$\text{Class I: } \begin{cases} (0,0) \text{ ——— } A_1,A_2 \text{ lost i.e. } A_3 \text{ is fixed} \\ (0,M) \text{ ——— } A_1,A_3 \text{ lost i.e. } A_2 \text{ is fixed} \\ (M,0) \text{ ——— } A_2,A_3 \text{ lost i.e. } A_1 \text{ is fixed} \end{cases}$$

$$\text{Class II: } \begin{cases} (s_1,s_2):\ s_1,s_2 > 0 \text{ such that } s_1+s_2=M, \text{ i.e. } A_3 \text{ is lost} \\ (0,\ s_2):\ s_2 \neq 0 \text{ or } M, \quad \text{i.e. } A_1 \text{ is lost} \\ (s_1,0):\ s_1 \neq 0 \text{ or } M, \quad \text{i.e. } A_2 \text{ is lost} \\ \text{Thus, in all there are } 3(M-1) \text{ states such that} \\ \text{exactly one of the three alleles is lost.} \end{cases}$$

$$\text{Class III: } \begin{cases} (s_1,s_2):\ s_1 \neq 0,\ s_2 \neq 0 \text{ and } s_1 + s_2 < M. \\ \\ \text{i.e. all the alleles coexist.} \end{cases}$$

Schematically the state space of the process can be represented by means of triangular coordinates (Fig. 1).

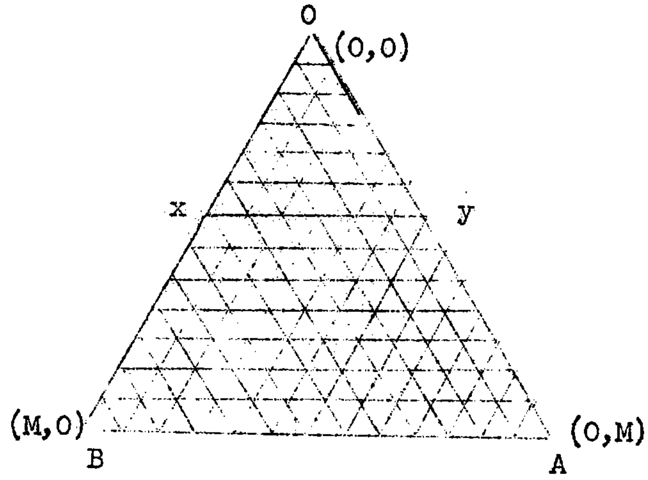

Fig. 1 A Geometric Representation of the State Space.

In the figure, the states in Class I are represented by the vertices of the triangle, whereas the states of Class II are interior points of the sides of the triangle. Class III states are interior points of the triangle. Conceptually, one can envision a point moving randomly about the triangle until it reaches one of the sides (which it must do with probability one in a finite member of generations), after which it must oscillate on that side until it is finally absorbed at one of the two vertices. Note that a transition to Class II reduces the problem to the binomial case considered in the first paper.

2.2. n-Stage Transition Probabilities. Let $\bar{X}_n$ denote the number of A_1 alleles and $\bar{Y}_n$ the number of A_2 alleles in generation n. Let $x, y > 0$. Then,

$$(1) \qquad P_{(x,y)(i,j)} = P(\bar{X}_n = i, \bar{Y}_n = j \mid \bar{X}_{n-1} = x, \bar{Y}_{n-1} = y)$$

$$= \frac{M!}{i!j!(M-i-j)!} \left(\frac{x}{M}\right)^i \left(\frac{y}{M}\right)^j \left(1 - \frac{x+y}{M}\right)^{M-i-j}$$

Expression (1) gives us the one-stage transition probabilities. The n-stage transition probabilities are obtained as follows. By the Chapman-Kolmogorov equation we know that

$$p^{n}_{(x,y)(i,j)} = \sum_{(u,v)} p^{n-1}_{(x,y)(u,v)}\, p_{(u,v)(i,j)},$$

where the summation is carried over all the possible states (u,v). Now $p_{(u,v)}$ is given by (1). Hence,

$$p^{n}_{(x,y)(i,j)} = \sum_{u=0}^{M}\sum_{v=0}^{M-u} p^{n-1}_{(x,y)(u,v)} \frac{M!}{i!j!(M-i-j)!} \left(\frac{u}{M}\right)^{i} \left(\frac{v}{M}\right)^{j} \left(1-\frac{u+v}{M}\right)^{M-i-j}$$

$$= \sum_{u=0}^{M}\sum_{v=0}^{M-u} p^{n-1}_{(x,y)(u,v)} \frac{M!}{i!j!(M-i-j)!} \left(\frac{u}{M}\right)^{i} \left(\frac{v}{M}\right)^{j} \cdot \sum_{r=0}^{M-i-j} (-1)^{r} \binom{M-i-j}{r} (u+v)^{r} M^{-r}$$

$$= \frac{M!\, M^{-i-j}}{i!j!(M-i-j)!} \sum_{r=0}^{(M-i-j)} (-1)^{r} \binom{M-i-j}{r} M^{-r} \cdot \sum_{u=0}^{M}\sum_{v=0}^{M-u} p^{n-1}_{(x,y)(u,v)}\, u^{i} v^{j} (u+v)^{r}$$

$$= \frac{M!\, M^{-i-j}}{i!j!(M-i-j)!} \sum_{r=0}^{(M-i-j)} (-1)^{r} \binom{M-i-j}{r} M^{-r} \cdot \sum_{u=0}^{M}\sum_{v=0}^{M-u} p^{n-1}_{(x,y)(u,v)}\, u^{i} v^{j} \sum_{\alpha=0}^{r} \binom{r}{\alpha} u^{\alpha} v^{r-\alpha}$$

$$= \frac{M!\, M^{-i-j}}{i!j!(M-i-j)!} \sum_{r=0}^{(M-i-j)} (-1)^r \binom{M-i-j}{r} M^{-r} \cdot$$

$$\cdot \sum_{\alpha=0}^{r} \binom{r}{\alpha} \sum_{u=0}^{M} \sum_{v=0}^{M-u} u^{i+\alpha} v^{j+r-\alpha} P^{n-1}_{(x,y)(u,v)} \cdot$$

Whence

$$\text{(2)} \qquad P^{n}_{(x,y)(u,v)} = \frac{M!\, M^{-i-j}}{i!j!(M-i-j)!} \sum_{r=0}^{M-i-j} (-1)^r M^{-r} \binom{M-i-j}{r} \cdot$$

$$\cdot \sum_{\alpha=0}^{r} \binom{r}{\alpha} E(\overline{X}^{i+\alpha}_{n-1} \overline{Y}^{j+r-\alpha}_{n-1} | \overline{X}_0 = x, \overline{Y}_0 = y)$$

Eq. (2) suggests that we can utilize the product moments of $\overline{X}_{n-1}$ and $\overline{Y}_{n-1}$ in order to find the nth stage transition probabilities. Thus the problem amounts to that of finding the product moments of the type $E(\overline{X}^t_{n-1} \overline{Y}^s_{n-1} | \overline{X}_0 = x, \overline{Y}_0 = y)$ for all t and s such that $t + s \leq M$.

These product moments can be obtained from the following recurrence relation (Khazanie, 1965)

$$\text{(3)} \qquad E(\overline{X}^t_n \overline{Y}^s_n | \overline{X}_0 = x, \overline{Y}_0 = y) = \sum_{\mu=1}^{t} \sum_{\nu=1}^{s} a_{t\mu}\, a_{s\nu} \frac{M^{(\mu+\nu)}}{M^{\mu+\nu}} E(\overline{X}^{\mu}_{n-1} \overline{Y}^{\nu}_{n-1} | \overline{X}_0 = x, \overline{Y}_0 = y).$$

where the a_{ij}'s are Stirling's numbers of the second kind and $M^{(i)}$, the ith factorial power of M, given by $M^{(i)} = \begin{cases} M(M-1)\ldots(M-i+1), & i > 0 \\ 1, & i = 0 \end{cases}$.

Put in matrix form (3) yields,

$$\text{(4)} \qquad {}_{(x,y)}E_n = C_1 \cdot {}_{(x,y)}E_{n-1}$$

where

(i) writing $E(\bar{X}_n^{\mu}\bar{Y}_n^{\nu}|X_0 = x,\ Y_0 = y) = {}_{(x,y)}\eta_n^{\mu,\nu}$,

$$
{}_{(x,y)}E_n = ({}_{(x,y)}\eta_n^{1,1},\ {}_{(x,y)}\eta_n^{2,1},\ldots,{}_{(x,y)}\eta_n^{M-1,1},{}_{(x,y)}\eta_n^{1,2},{}_{(x,y)}\eta_n^{22},\ldots,{}_{(x,y)}\eta_n^{M-2,2},\ldots,{}_{(x,y)}\eta_n^{1,M-1})',
$$

the prime indicating the transpose.

(ii) C_1 is the matrix of the coefficients which can be partitioned as

$$
C_1 = \begin{bmatrix}
B_{11} & & & & \\
B_{21} & B_{22} & & & \\
\vdots & & & & \\
B_{k1} & B_{k2} & \ldots, & B_{kk} & \\
B_{(M-1),1} & B_{(M-1),2} & \ldots & B_{(M-1),k}, & \ldots, \ B_{(M-1),(M-1)}
\end{bmatrix}
$$

where $B_{ij} = (E_{ij}|0_{ij})$ are $(M-i) \times (M-j)$ matrices with

$$
E_{ij} = a_{ij}\begin{bmatrix}
a_{11}\dfrac{M^{(j+1)}}{M^{j+1}} & 0 & 0 \ldots\ldots 0 \\
a_{21}\dfrac{M^{(j+1)}}{M^{j+1}} & a_{22}\dfrac{M^{(j+2)}}{M^{j+2}} & 0 \ldots\ldots 0 \\
\vdots & \vdots & \\
a_{M-i,1}\dfrac{M^{(j+1)}}{M^{j+1}} & a_{M-i,2}\dfrac{M^{(j+2)}}{M^{j+2}} \ \ldots & a_{M-i,M-i}\dfrac{M^{(M+j-i)}}{M^{M+j-i}}
\end{bmatrix}
$$

and 0_{ij} an $(M-i) \times (i-j)$ zero matrix.

From (4) it follows that ${}_{(x,y)}E_n = C_1^n \cdot {}_{(x,y)}E_0$, and, since (x,y) is the initial starting point,

$$(x,y)^{E_0} = (xy, x^2y, \ldots, x^{M-1}y, xy^2, \ldots, x^{M-2}y^2, \ldots, xy^{M-2}, x^2y^{M-2}, xy^{M-1})'.$$

We will now proceed to obtain the spectral resolution of C_1.

The distinct eigenvalues of C_1 are $M^{(r)}/M^r$, $r=2,\ldots,M$, $M^{(r)}/M^r$ being repeated $r-1$ times. In general, a matrix with repeated eigenvalues need not be diagonable. We will establish in the following section that the matrix C_1 because of its nature is diagonable and will show how to get the similarity transformation.

3. Diagonability of C_1.

Lemma 1. If T, $\overline{U}$, V are given $p \times p$ lower triangular matrices such that $t_{ss} \neq u_{kk}$ for $1 \leq s \leq k \leq p$, then there exists a triangular matrix Ω such that

$$\Omega T - \overline{U}\Omega = V. \tag{5}$$

Proof: The rows of Ω may be chosen step by step working to the left from the diagonal in each successive row.

We will show by induction on k that for every k there exists a pxp matrix $\Omega^{(k)}$ which makes (5) true in rows $1,2,\ldots,k$. For $k = 1$ we have to solve

$$(\omega_{11}t_{11},0,\ldots,0) - (u_{11}\omega_{11},0,\ldots,0) = (v_{11},0,\ldots,0).$$

That is, $(t_{11}-u_{11})\omega_{11} = v_{11}$. Since $t_{11} \neq u_{11}$, we get $\omega_{11} = v_{11}/(t_{11}-u_{11})$. Whence the first row of Ω is given by $(v_{11}/(t_{11}-u_{11}),0,0,\ldots,0)$.

Suppose then there exists a pxp matrix Ω which makes (5) true in rows $1,2,\ldots,k-1$. Since all matrices under consideration are triangular, only the first (k-1) rows of Ω are involved in the first (k-1) rows of $\Omega T - \bar{U}\Omega$. Therefore, we can construct $\Omega^{(k)}$ by taking as its first (k-1) rows those of $\Omega^{(k-1)}$ and by then choosing the kth row of $\Omega^{(k)}$ so as to make (5) true in the kth row. That is, suppose the set $\{\omega_{rs},\ r=1,2,\ldots,k-1\}$ is known with $\omega_{rs} = 0$ if $r < s$. We must now choose $\omega_{ks}(1 \le s \le k)$ with $\omega_{ks} = 0$ for $k < s$ to satisfy,

$$\sum_{r=s}^{k} (\omega_{kr} t_{rs} - u_{kr}\omega_{rs}) = v_{ks}, \quad (s = 1,2,\ldots,k)$$

that is,

$$\sum_{r=s}^{k} \omega_{kr} t_{rs} - u_{kk}\omega_{ks} = v_{ks} + \sum_{r=s}^{k-1} u_{kr}\omega_{rs}.$$

Whence

(6) $$\sum_{r=s}^{k} \omega_{kr} t_{rs} - u_{kk}\omega_{ks} = \eta_{ks}, \text{ say} \quad (s=1,2,\ldots,k).$$

By the induction hypothesis, the numbers $\omega_{rs}(s \le r \le k-1)$ are known. Therefore the values η_{ks} are known. Thus from (6), given $\eta_{ks}, t_{rs}, u_{kk}$ we have only to solve

(7) $$\omega_{ks}(t_{ss} - u_{kk}) = \eta_{ks} - \sum_{r=s+1}^{k} \omega_{kr} t_{rs}, \quad (s=1,2,\ldots,k).$$

Now each $t_{ss} - u_{kk}$ is non-zero, whence, in particular

(8) $$\omega_{kk} = v_{kk}/(t_{kk} - u_{kk}).$$

By downward induction on s we can now successively obtain $\omega_{kk}, \omega_{k,k-1},\ldots,\omega_{k1}$.

<u>Note</u>: In the special case when T is actually diagonal (as in our application of the lemma) (7) simply reduces to $\omega_{ks}(t_{ss} - u_{kk}) = \eta_{ks}$. Therefore $\omega_{ks} = \eta_{ks}/(t_{ss} - u_{kk})$, $(s = 1,2,\ldots,k)$.

<u>Theorem 1</u>: Let $m > n$ and $P = \left(\begin{array}{c|c} A & 0 \\ \hline B & C \end{array}\right)$ where,

i) A_{mxm} is triangular and diagonable; thus, there exists a non singular matrix M_{mxm} (triangular) such that

$$M^{-1}AM = F = \begin{pmatrix} T & 0 \\ 0 & R \end{pmatrix}$$

partitioned so that T, R are diagonal with T an nxn matrix and R an $(m-n) \times (m-n)$ matrix.

ii) C_{nxn} is a triangular matrix which is diagonable with $t_{ss} \neq c_{kk}$ for $1 \leq s \leq k \leq n$.

iii) $B = (E|0)$ where 0 is an $n \times (m-n)$ zero matrix and E is nxn triangular matrix so that B is $n \times m$.

iv) 0_{mxn} is an mxn zero matrix,

then the matrix P is diagonable.

<u>Proof</u>: Since C is diagonable there exists a non-singular matrix N such that $N^{-1}CN = G$ where G is a diagonal matrix whose diagonal elements are the eigenvalues of C.

Let $L = \left(\begin{array}{c|c} M & 0 \\ \hline K & N \end{array}\right)$ where k is an nxm matrix to be suitably chosen.

Since M and N are non-singular so is L. Hence L^{-1} exists. Consider $L^{-1}PL$.

$$L^{-1}PL = \begin{pmatrix} M^{-1} & 0 \\ -N^{-1}KM^{-1} & N^{-1} \end{pmatrix} \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} \begin{pmatrix} M & 0 \\ K & N \end{pmatrix}$$

$$= \begin{pmatrix} M^{-1}AM & 0 \\ -N^{-1}KM^{-1}AM + N^{-1}(BM+CK) & N^{-1}CN \end{pmatrix}$$

$$= \begin{pmatrix} F & 0 \\ N^{-1}(-KF + BM + CK) & G \end{pmatrix}$$

This shows that if L is to diagonalize P it suffices to choose K such that $N^{-1}(-KF+BM+CK) = 0$. That is, choose K such that $KF-CK = BM$. Since M is triangular we can partition it as

$$M = \begin{array}{c} \begin{array}{cc} \overbrace{\quad}^{n} & \overbrace{\quad}^{m-n} \end{array} \\ \left(\begin{array}{c|c} H & 0 \\ \hline Q & S \end{array}\right) \end{array} \begin{array}{l} \\ \}\, n \\ \}\, m-n \end{array}$$

with H_{nxn} triangular. If we now choose $K = (\overset{n}{\widetilde{\Omega}}, \overset{m-n}{\widetilde{0}})\}n$, we need only find Ω to satisfy

$$(\Omega, 0)\begin{pmatrix} T & 0 \\ 0 & R \end{pmatrix} - C(\Omega, 0) = (E, 0)\left(\begin{array}{c|c} H & 0 \\ \hline Q & S \end{array}\right)$$

so that $(\Omega T, 0) - (C\Omega, 0) = (EH, 0)$ and consequently,

$$\Omega T - C\Omega = EH = V, \quad \text{say.} \tag{9}$$

Now E and H are triangular. Hence EH is triangular. The result now follows from Lemma 1 with $\bar{U} = C$

Remark: (9) indicates that it is not essential that E be triangular. It suffices to have E such that EH is triangular.

In the following let $B_{ij}(t, \ell)$ represent the (t, ℓ)th element of the matrix B_{ij}.

Theorem 2: Let $$D^*_k = \begin{pmatrix} B_{11} & & & \\ B_{21} & B_{22} & & \\ \vdots & & & \\ B_{k1} & B_{k2} & \cdots & B_{kk} \end{pmatrix}$$

be a triangular matrix where $B_{ij} = (E_{ij}|O_{ij})$, $(i,j=1,2,\ldots,k)$, is an $(M-i) \times (M-j)$ matrix with E_{ij} an $(M-i) \times (M-i)$ triangular matrix and O_{ij} an $(M-i) \times (i-j)$ zero matrix. Also let B_{ii}, $(i = 1,2,\ldots,k)$ be diagonable, and if $\mu > \nu$ let $B_{\nu\nu}(s,s) \neq B_{\mu\mu}(k,k)$, $1 \leq s \leq k \leq M-\mu$. Then for every $k < M$, D^*_k can be diagonalized by a trinagular matrix P^*_k which can be written as

$$P^*_k = \begin{pmatrix} S_{11} & & & \\ S_{21} & S_{22} & & \\ \vdots & & & \\ S_{k1} & S_{k2} & \cdots & S_{kk} \end{pmatrix}$$

where $S_{ij} = (T_{ij}|O_{ij})$, $(i,j=1,2,\ldots,k)$ is an $(M-i) \times (M-j)$ matrix with T_{ij} an $(M-i) \times (M-i)$ triangular matrix and O_{ij} and $(M-i) \times (i-j)$ zero matrix. Furthermore, in the process of diagonalization of D^*_k by P^*_k, S_{ii}

diagonalizes B_{ii} for every i so that $S_{ii}^{-1} B_{ii} S_{ii} = F_{ii}$.

Proof: The proof will be by induction on k. B_{11}, B_{21}, B_{22} satisfy the conditions of Theorem 1. Therefore $D_2^* = \begin{pmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{pmatrix}$ can be diagonalized by a matrix of the type $P_2^* = \begin{pmatrix} S_{11} & 0 \\ S_{21} & S_{22} \end{pmatrix}$ so that

$$P_2^{*-1} D_2^* P_2^* = \begin{pmatrix} S_{11}^{-1} B_{11} S_{11} & 0 \\ 0 & S_{22}^{-1} B_{22} S_{22} \end{pmatrix} = \begin{pmatrix} F_{11} & 0 \\ 0 & F_{22} \end{pmatrix}.$$

Hence the theorem is true for $k = 2$. Next let us assume that it is true for $k-1$; that is, D_{k-1}^* can be diagonalized by a matrix P_{k-1}^* so that $P_{k-1}^{*-1} D_{k-1}^* P_{k-1}^* = \mathrm{Diag}\,\{F_{11}, F_{22}, \ldots, F_{k-1,k-1}\} = F_{k-1}$, say.

We will now establish that, indeed, $D_k^* = \left(\begin{array}{c|c} D_{k-1}^* & 0 \\ \hline R & B_{kk} \end{array}\right)$

where $R = (B_{k1}, \ldots, B_{k,k-1})$, can be diagonalized by $P_k^* = \left(\begin{array}{c|c} P_{k-1}^* & 0 \\ \hline Q & S_{kk} \end{array}\right)$

where $Q = (S_{k1}, S_{k2}, \ldots, S_{k,k-1})$.

If P_k^* diagonalizes D_k^* then following the argument in Theorem 1,

$$P_k^{*-1} D_k^* P_k^* = \left(\begin{array}{c|c} F_{k-1} & 0 \\ \hline S_{kk}^{-1}(-QF_{k-1}+RP_{k-1}^*+B_{kk}Q) & F_{kk} \end{array}\right) .$$

Thus it is sufficient to choose Q such that $S_{kk}^{-1}(-QF_{k-1}+RP_{k-1}^*+B_{kk}Q) = 0$, which yields $QF_{k-1}-B_{kk}Q = RP_{k-1}^*$; that is, it suffices to choose $S_{k1},S_{k2},\dots,S_{k,k-1}$ such that

$$(S_{k1},\dots,S_{k,k-1})\ \mathrm{Diag}\ \{F_{11},F_{22},\dots,F_{k-1,k-1}\} - B_{kk}(S_{k1},\dots,S_{k,k-1}) =$$

$$(B_{k1},\dots,B_{k,k-1}) \begin{pmatrix} S_{11} & & & \\ S_{21} & S_{22} & & \\ \vdots & & & \\ S_{k-1,1} & S_{k-1,2} & \cdots & S_{k-1,k-1} \end{pmatrix};$$

whence, $(S_{k1}F_{11}-B_{kk}S_{k1},\dots,S_{ki}F_{ii}-B_{kk}S_{ki},\dots,S_{k,k-1}F_{k-1,k-1}-B_{kk}S_{k,k-1})$

$$= \left(\sum_{t=1}^{k-1} B_{kt}S_{t},\dots, \sum_{t=i}^{k-1} B_{kt}S_{ti},\dots,B_{k,k-1}S_{k-1,k-1}\right).$$

Thus we get

$$S_{ki}F_{ii} - B_{kk}S_{ki} = \sum_{t=i}^{k-1} B_{kt}S_{ti}, \quad i = 1,2,\dots,k-1. \tag{10}$$

In equation (10) we know that $B_{kt} = \{(\overbrace{E_{kt}}^{M-k} \mid \overbrace{O_{kt}}^{k-t})\}\ M-k$. Also S_{ti} can be written as

$$S_{ti} = \begin{array}{cc} \overbrace{\quad}^{M-k} & \overbrace{\quad}^{k-i} \end{array}$$
$$\left(\begin{array}{c|c} \psi_{ti} & 0 \\ \hline \chi_{ti} & \xi_{ti} \end{array}\right) \begin{array}{l} \}\ M-k \\ \}\ k-t \end{array}$$

with ψ_{ti} an (M-k) x (M-k) triangular matrix. Hence

$$B_{kt}\, S_{ti} = (\overbrace{E_{kt}\ \psi_{ti}}^{M-k} \mid \overbrace{0}^{k-i})\}\ M-k.$$

where $E_{kt}\psi_{ti}$ is an (M-k) x (M-k) triangular matrix since E_{kt} and ψ_{ti} are triangular. Consequently,

$$\sum_{t=i}^{k-1} B_{kt}S_{ti} = (\overbrace{\sum_{t=i}^{k-1} E_{kt}\ \psi_{ti}}^{M-k} \mid \overbrace{0}^{k-i})\}\ M-k$$

with $\sum_{t=i}^{k-1} E_{kt}\ \psi_{ti}$ triangular.

Next, let us write

$$F_{ii} = \begin{array}{cc} \overbrace{\quad}^{M-k} & \overbrace{\quad}^{k-i} \end{array}$$
$$\left(\begin{array}{c|c} \Phi^{(i,i)} & 0 \\ \hline 0 & \zeta^{(i,i)} \end{array}\right) \begin{array}{l} \}\ M-k \\ \}\ k-i \end{array}$$

If we now choose $S_{ki} = (\overbrace{\Gamma_{ki}}^{M-k} \mid \overbrace{0}^{k-i})\}\ M-k$, then from (10) we need find T_{ki} to satisfy

$$(T_{ki}|0)\left(\begin{array}{c|c} \Phi^{(i,i)} & 0 \\ \hline 0 & \zeta^{(i,i)} \end{array}\right) - B_{kk}(T_{ki}|0) = (\sum_{t=i}^{k-1} E_{kt}\,\psi_{ti}|0);$$

that is, $(T_{ki}\,\Phi^{(i,i)}|0) - (B_{kk}\,T_{ki}|0) = (\sum_{t=i}^{k-1} E_{kt}\,\psi_{ti}|0)$. In other words we must find T_{ki} such that

$$T_{ki}\,\Phi^{(i,i)} - B_{kk}\,T_{ki} = \sum_{t=i}^{k-1} E_{kt}\,\psi_{ti}\,. \tag{11}$$

Now since $\Phi^{(i,i)}$, T_{ki} and $\sum_{t=i}^{k-1} E_{kt}\,\psi_{ti}$ satisfy the condition of Lemma 1, the fact that we can find a triangular matrix T_{ki} is an immediate consequence of the Lemma.

4. The Distribution of Time to Homozygosity

Let us suppose that the population starts with x A_1-alleles and y A_2-alleles, and let $\tau_{(x,y)}$ denote the time taken to reach one of the homozygous conditions for such a population. In order to obtain the distribution of $\tau_{(x,y)}$ we will state without proof the following lemma which is basic to the distribution. (Khazanie, 1965).

Lemma 2: The probability of fixation for any given allele is independent of the initial gene frequency of the other two alleles. In fact,

$$p^n_{(x,y)(M,0)} = p^n_{xM} \qquad \text{for every } y \tag{12}$$

where, p^n_{xM} is the probability of fixation by the nth generation, as defined in the two allele case, when the initial frequency is x. (Khazanie and McKean, 1965).

We will now derive the distribution of $\tau_{(x,y)}$. The population can become homozygous either by absorption in the state (M,0), or in (0,M) or in (0,0). Following the argument in the two allele-case then,

$$P(\tau_{(x,y)} = 1) = (\tfrac{x}{M})^M + (\tfrac{y}{M})^M + (1 - \tfrac{x+y}{M})^M , \tag{13a}$$

and, if $n > 1$,

$$P(\tau_{(x,y)} = n) = p^n_{(x,y)(M,0)} - p^{n-1}_{(x,y)(M,0)} + p^n_{(x,y)(0,M)} - p^{n-1}_{(x,y)(0,M)}$$

$$+ p^n_{(x,y)(0,0)} - p^{n-1}_{(x,y)(0,0)} .$$

Applying lemma 2 and noting that

$$p^n_{xM} = \frac{x}{M} + M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} (M^{(s)}/M^s)^{n-1} u_{Ms}\, v_{s\beta}\, x^\beta ,$$

(see Khazanie and McKean, 1965) it is a matter of simple algebra to show that

$$P(\tau_{(x,y)} = n) = M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} (M^{(s)}/M^s)^{n-2} \left[\frac{M^{(s)}}{M^s} - 1\right] \tag{13b}$$

$$\left[x^\beta + y^\beta + (M-x-y)^\beta\right] u_{Ms}\, v_{s\beta} .$$

Comparison of the distribution of $\tau_{(x,y)}$ with that of τ_x in the two allele case shows that the functional form of both is the same. Hence, writing $C(x,y,s,\beta)$ for $\left[\frac{M^{(s)}}{M^s} - 1\right]\left[x^\beta + y^\beta + (M-x-y)^\beta\right] u_{Ms} v_{s\beta}$ we get:

1. The probability generating function $G_{(x,y)}(z)$ of $\tau_{(x,y)}$ is

$$(14) \qquad G_{(x,y)}(z) = \left[\left(\frac{x}{M}\right)^M + \left(\frac{y}{M}\right)^M + \left(1-\frac{x+y}{M}\right)^M\right] z + M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} C(x,y,s,\beta) \frac{z^2}{\left(1-\frac{M^{(s)}}{M^s} z\right)}$$

2. The expected value of $\tau_{(x,y)}$ is

$$(15) \qquad E(\tau_{(x,y)}) = \left(\frac{x}{M}\right)^M + \left(\frac{y}{M}\right)^M + \left(1-\frac{x+y}{M}\right)^M + M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} C(x,y,s,\beta) \frac{2-\frac{M^{(s)}}{M^s}}{\left(1-\frac{M^{(s)}}{M^s}\right)^2} .$$

3. The variance of $\tau_{(x,y)}$ is

$$E(\tau_{(x,y)}(\tau_{(x,y)}-1)) + E(\tau_{(x,y)}) - (E(\tau_{(x,y)}))^2$$

where

$$(16) \qquad E(\tau_{(x,y)}(\tau_{(x,y)}-1)) = M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} C(x,y,s,\beta) \frac{2}{\left(1-\frac{M^{(s)}}{M^s}\right)^3} .$$

5. Multiple Alleles.

The extension of the above treatment to the multiallelic case is now straight forward. Let us suppose that there are $k+1$ alleles which we will denote by $A_1, A_2, \ldots, A_{k+1}$. If the population consists of i_ν genes of type A_ν, $\nu=1,2,\ldots,k+1$ such that $\sum_{\nu=1}^{k+1} i_\nu = M$, we will say that the population is in the state $(i_1, i_2, \ldots, i_k)$ and call this particular state S^i. Geometrically, the representation can be accomplished by means of a regular k-dimensional simplex where the states correspond to the points of intersection of the (k-1) dimensional hyperplanes parallel to the faces of the simplex.

Let $\underset{\nu}{\overline{X}}{}_n$, $\nu=1,2,\ldots,k+1$ denote the number of A_ν alleles in generation n. The transition probabilities from any state $S^i=(i_1, i_2, \ldots, i_k)$ to any other state $S^j = (j_1, j_2, \ldots, j_k)$, given by

$$(17) \qquad p_{S^iS^j} = \frac{m!}{j_1!j_2!\ldots j_k!(M-j_1-\ldots-j_k)!} \left(\frac{i_1}{M}\right)^{j_1} \ldots \left(\frac{i_k}{M}\right)^{j_k} \cdot$$

$$\cdot \left(1-\frac{i_1+\ldots+i_k}{M}\right)^{M-j_1-\ldots-j_k},$$

completely specify the underlying Markovian process.

In analogy with the triallelic case it is a matter of straight forward generalization to show that

$$(18) \qquad p^n_{S^xS^j} = \frac{M!M^{-j_1-\ldots-j_k}}{j_1!j_2!\ldots j_k!j_{k+1}!} \sum_{r=0}^{j_{k+1}} (-1)^r \binom{j_{k+1}}{r} M^{-r}.$$

$$\cdot \sum_{\alpha_1=0}^{r} \cdots \sum_{\alpha_{k-1}=0}^{(r-\alpha_1-\cdots-\alpha_{k-2})} \frac{r!}{\alpha_1!\cdots\alpha_{k-1}!(r-\alpha_1-\cdots-\alpha_{k-1})!} \cdot$$

$$\cdot E({}_1\bar{X}_{n-1}^{j_1+\alpha_1} \cdots {}_k\bar{X}^{j_k+r-\alpha_1-\cdots-\alpha_{k-1}} \Big| {}_\nu\bar{X}_0 = x_\nu,\ \nu=1,\ldots k).$$

and the recurrence relation between the product moments is,

$$(19) \quad E({}_1\bar{X}_n^{r_1} \cdots {}_k\bar{X}_n^{r_k} \Big| {}_\nu\bar{X}_0 = x_\nu,\ \nu=1,2,\ldots,k)$$

$$= \sum_{\mu_1}^{r_1} \cdots \sum_{\mu_k}^{r_k} a_{r_1\mu_1} \cdots a_{r_k\mu_k} \frac{M^{(\mu_1+\cdots+\mu_k)}}{M^{\mu_1+\cdots+\mu_k}} \cdot$$

$$\cdot E({}_1\bar{X}_{n-1}^{\mu_1} \cdots {}_k\bar{X}_{n-1}^{\mu_k} \Big| {}_\nu\bar{X}_0 = x_\nu,\ \nu=1,2,\ldots k)$$

Put in matrix notation, this gives

$$(20) \qquad (x_1,\ldots,x_k)^{E_n} = C_k \cdot (x_1,\ldots,x_k)^{E_{n-1}}$$

where, $(x_1,\ldots,x_k)^{E_n}$ is the vector of the conditional product moments in generation n given ${}_\nu\bar{X}_0 = x_\nu$, $\nu=1,2,\ldots,k$ and C_k is the matrix of coefficients analogous to C_1 in the triallelic case. Like C_1, C_k is also triangular. The procedure for diagonalizing it is similar to that of diagonalizing C_1.

As regards the distribution to time to homozygosity, if ${}_S\tau_x$ represents such a random variable with $S_x = (x_1,\ldots,x_k)$ as the initial state of the process, adopting the argument in the triallelic case, it immediately follows that

$$\text{(21a)} \qquad P(\tau_{S^x} = 1) = \sum_{\nu=1}^{k+1} \left(\frac{x_\nu}{M}\right)^M$$

and if $n > 1$,

$$\text{(21b)} \qquad P(\tau_{S^x} = n) = M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} \left(\frac{M^{(s)}}{M^s}\right)^{n-2} \left[\frac{M^{(s)}}{M^s} - 1\right] \left\{\sum_{\nu=1}^{k+1} x_\nu^\beta\right\} u_{Ms} \, v_{s\beta} \, ,$$

where, of course, $\sum_{\nu=1}^{k+1} x_\nu = M$.

The functional form of the probability function of τ_{S^x} being similar to that in the triallelic case, the probability generating function, the expected value and the variance of τ_{S^x} can be obtained in a like manner.

6. Multiple Loci.

The investigation will now be extended to a more general case involving an arbitrary number of alleles at an arbitrary number of loci but assuming that the loci segregate independently. Let there be L loci with k_ℓ alleles at the ℓth locus. Further, let $\tau^\ell_{S^x_\ell}$ represent the time taken by the population to reach homozygosity at locus ℓ when it initially starts from the state $S^x_\ell = (x_1^{(\ell)}, \ldots, x_{k_\ell - 1}^{(\ell)})$ at that locus. The distribution of τ_{S^x} has already been found in (21a), (21b). What we are interested in here is the distribution of the time taken to reach the homozygous condition for all the L loci concerned. This random variable which we will denote by τ^* is obviously

$$\tau^* = \max_{\ell} (\tau^{\ell}_{S_\ell^x} : \ell = 1,2,\ldots,L).$$

The distribution function of τ^* is given by,

$$P(\tau^* \leq n) = P(\tau^{1}_{S_1^x} \leq n, \ldots, \tau^{\ell}_{S_\ell^x} \leq n, \ldots, \tau^{L}_{S_L^x} \leq n)$$

$$= \prod_{\ell=1}^{L} P(\tau^{\ell}_{S_\ell^x} \leq n),$$

since the loci are assumed to be independent. The probability function of τ^* is therefore,

(22) $$P(\tau^* = n) = \prod_{\ell=1}^{L} P(\tau^{\ell}_{S_\ell^x} \leq n) - \prod_{\ell=1}^{L} P(\tau^{\ell}_{S_\ell^x} \leq n-1),$$

which requires that we know the distribution function of $\tau^{\ell}_{S_\ell^x}$ at every locus. Now it has been shown in (21a), (21b) that

$$P(\tau^{\ell}_{S_\ell^x} = 1) = \sum_{\nu=1}^{k_\ell} \left(\frac{x_\nu^{(\ell)}}{M}\right)^M$$

and, if $m > 1$,

$$P(\tau^{\ell}_{S_\ell^x} = m) = M^{-M} \sum_{s=2}^{M} \sum_{\beta=1}^{s} \left(\frac{M^{(s)}}{M^s}\right)^{m-2} \left[\frac{M^{(s)}}{M^s} - 1\right] \left\{ \sum_{\nu=1}^{k} [x_\nu^{(\ell)}]^\beta \right\} u_{Ms} \, v_{s\beta} \,.$$

Therefore, clearly

$$P(\tau^{\ell}_{S_{x_\ell}} \leq n) = \sum_{\nu=1}^{k_\ell}\left[\frac{x_\nu^{(\ell)}}{M}\right]^M + M^{-M}\sum_{s=2}^{M}\sum_{\beta=1}^{s}\left[\frac{M^{(s)}}{M^s}-1\right]\left\{\sum_{\nu=1}^{k_\ell}\left[x_\nu^{(\ell)}\right]^\beta\right\} u_{Ms}\, v_{s\beta} \cdot$$

$$\cdot \frac{1-\left[\frac{M^{(s)}}{M^s}\right]^{n-1}}{1-\frac{M^{(s)}}{M^s}} .$$

Whence, from (22) we get

$$P(\tau^* = n)$$

$$= \prod_{\ell=1}^{L}\left[\sum_{\nu=1}^{k_\ell}\left[\frac{x_\nu^{(\ell)}}{M}\right]^M + M^{-M}\sum_{s=2}^{M}\sum_{\beta=1}^{s}\left[\frac{M^{(s)}}{M^s}-1\right]\left\{\sum_{\nu=1}^{k_\ell}\left[x_\nu^{(\ell)}\right]^\beta u_{Ms}\, v_{s\beta}\left[\frac{1-\left[\frac{M^{(s)}}{M^s}\right]^{n-1}}{1-\frac{M^{(s)}}{M^s}}\right]\right\}\right]$$

$$- \prod_{\ell=1}^{L}\left[\sum_{\nu=1}^{k_\ell}\left[\frac{x_\nu^{(\ell)}}{M}\right]^M + M^{-M}\sum_{s=2}^{M}\sum_{\beta=1}^{s}\left[\frac{M^{(s)}}{M^s}-1\right]\left\{\sum_{\nu=1}^{k_\ell}\left[x_\nu^{(\ell)}\right]^\beta u_{Ms}\, v_{s\beta}\left[\frac{1-\left[\frac{M^{(s)}}{M^s}\right]^{n-2}}{1-\frac{M^{(s)}}{M^s}}\right]\right\}\right] .$$

7. Numerical Illustrations of the General Technique.

From lemma 2 and its obvious extension to the case of k alleles, irrespective of the number of allelomorphs at a locus, basic to the distribution of the time to homozygosity when a single locus is involved are the probabilities of fixation p_{iM}^{n}. For example, with three alleles at a locus in a population of twelve gametes and initial frequencies of 3,4, and 5 the distribution can be obtained from Table 7 of Khazanie and McKean (1965) by considering columns corresponding to $i=3$, $i=4$, and $i=5$. For example, for n=15, we have $P(\tau = 15) = (.11843 - .10783) + (.17497 - .16166) + (.24077 - .22540) = .03928$.

Table 1 gives such distributions for 3 allelic, 4-allelic and 5-allelic cases with respective initial states, (3,4), (2,3,6) and (4,2,2,2). From this is obtained in Table 2 the distribution when the above mentioned loci are involved simultaneously in the population but segregate independently.

To illustrate the calculations for Table 2, let τ_1, τ_2, and τ_3 denote the times to homozygosity of loci 1,2, and 3, respectively. From Table 1, $P(\tau_1 \leq 5) = .07132$ (the sum of the first 5 entries in column (3, 4). Similarly, $P(\tau_2 \leq 5) = .08888$, and $P(\tau_3 \leq 5) = .03289$. Whence, since $P(\tau^* \leq 5) = P(\tau_1 \leq 5)\ P(\tau_2 \leq 5)\ P(\tau_3 \leq 5)$ (for $\tau^* = \max(\tau_1,\tau_2,\tau_3)$), we have $P(\tau^* \leq 5) = .000209$, which is the fifth entry in Table 2.

8. Summary and Conclusion

By the methods outlined in the series of two papers, it is possible to obtain directly the exact distribution of time-to-homozygosity of a diploid, monoecious, random mating population of any finite size and with arbitrarily many independently segregating loci each with an arbitrary number of alleles. The technique does not depend upon continuity assumption and further is numerically feasible even for large population sizes. The main results of the papers have numerical verification in the paper of Ewans (1963) who obtained results identical to ours (to four significant figures) by the direct approach of powering the transition matrix.

Extension of this procedure to allow for mutation appears feasible and is presently under investigation .

Acknowledgment

The authors wish to thank Professors Michael Drazin and Marcel Neuts of the Division of Mathematical Sciences, Purdue University, for some very useful ideas during the period of this investigation.

References

1. Ewans, W.J. (1963). Numerical results and diffusion approximations in a genetic process. Biometrika 50: 241-249.

2. Khazanie, R.G. (1965). On a mendelian markov process with multinomial transition probabilities. Ph.D. thesis, Purdue University Libraries.

3. Khazanie, R.G., and McKean, H.E. (1965). A mendelian markov process with multinomial transition probabilities, I. The Binomial Case. Mimeo series number 33, Dept. of Statistics, Purdue University.

4. Kimura, M. (1955). Random genetic drift in multi-allelic locus. Evolution 9: 419-435.

Table 1. Exact Distribution of τ_{S^x}, Time to Homozygosity (M = 12) for 40 generations

Initial State	(3,4)	(2,3,6)	(4,2,2,2)
Generation			
1	0.00003	0.00024	0.00000
2	0.00218	0.00590	0.00042
3	0.01106	0.01760	0.00345
4	0.02347	0.02846	0.01013
5	0.03461	0.03668	0.01889
6	0.04268	0.04242	0.02768
7	0.04767	0.04606	0.03519
8	0.05018	0.04799	0.04089
9	0.05086	0.04858	0.04473
10	0.05026	0.04815	0.04692
11	0.04880	0.04696	0.04775
12	0.04678	0.04525	0.04751
13	0.04442	0.04318	0.04648
14	0.04189	0.04089	0.04489
15	0.03928	0.03849	0.04291
16	0.03668	0.03606	0.04070
17	0.03413	0.03365	0.03835
18	0.03167	0.03130	0.03596
19	0.02933	0.02904	0.03358
20	0.02711	0.02689	0.03126
21	0.02502	0.02485	0.02901
22	0.02307	0.02294	0.02687
23	0.02125	0.02115	0.02485
24	0.01955	0.01948	0.02294
25	0.01798	0.01792	0.02115
26	0.01653	0.01648	0.01948
27	0.01518	0.01515	0.01793
28	0.01394	0.01392	0.01649
29	0.01280	0.01278	0.01516
30	0.01175	0.01174	0.01393
31	0.01078	0.01077	0.01279
32	0.00989	0.00988	0.01175
33	0.00908	0.00907	0.01078
34	0.00832	0.00832	0.00989
35	0.00763	0.00763	0.00908
36	0.00700	0.00700	0.00833
37	0.00642	0.00642	0.00764
38	0.00589	0.00589	0.00701
39	0.00540	0.00540	0.00642
40	0.00495	0.00495	0.00589
Expected Value of τ_{S^x}	17.79313	17.64650	19.59614
Variance of τ_{S^x}	145.92499	148.64218	150.10495

Table 2. Cumulative Distribution of τ^*, Time to Homozygosity ($M = 12$), for 40 Generations

Initial State / Generation	Loc 1 (3,4) Loc 2 (2,3,6) Loc 3 (4,2,2,2)
1	.00000
2	.00000
3	.00000
4	.00003
5	.00021
6	.00091
7	.00275
8	.00653
9	.01306
10	.02302
11	.03686
12	.05477
13	.07669
14	.10234
15	.13130
16	.16305
17	.19702
18	.23264
19	.26936
20	.30666
21	.34408
22	.38124
23	.41780
24	.45349
25	.48808
26	.52142
27	.55337
28	.58386
29	.61283
30	.64025
31	.66614
32	.69050
33	.71336
34	.73478
35	.75480
36	.77347
37	.79086
38	.80704
39	.82207
40	.83600