

INBREEDING IN POPULATIONS WITH OVERLAPPING GENERATIONS

D. L. JOHNSON*

Institute of Animal Genetics, West Mains Road, Edinburgh, Scotland

Manuscript received November 22, 1976

Revised copy received July 6, 1977

ABSTRACT

An inbreeding matrix is defined for populations with overlapping generations. In the short term it can be expressed in terms of a matrix specifying the passage of genes between the different age groups (and sexes) and a diagonal matrix whose elements depend on the number of individuals in each age group. Formulae for the inbreeding effective number are derived using matrix theory. A comparison is made between the inbreeding coefficients predicted by this theory and those obtained by assuming a uniform rate of inbreeding from the outset, and these in turn are compared with the exact inbreeding coefficients.

THE concept of effective population number was introduced by WRIGHT (1931, 1938) and is useful for predicting inbreeding or random genetic drift. For a population in which the generations overlap, the effective population number N_e has been determined, for specific models, under the assumption that the population has reached the equilibrium state (see HILL 1972). In particular, FELSENSTEIN (1971) considered monocious populations with age-specific birth and death rates and obtained formulas for the inbreeding and variance effective numbers.

In a population with discrete generations, the rate of inbreeding per generation is given approximately by $1/2N_e$. With overlapping generations, this rate of inbreeding is attained only in the later generations. The purpose of this study is to determine the pattern of inbreeding over an initial period of time in a newly established population and to check the usefulness of the assumption of a uniform rate of inbreeding from the outset. This is of interest for animal breeding work where one is usually concerned with the early linear phase of increase in the inbreeding coefficient.

We first consider a finite random-mating population of monocious diploids of constant size and age distribution. We assume that deaths occur at random and there is a random distribution of family size from surviving parents. FELSENSTEIN (1971) derived the inbreeding effective number for this model by considering rates of change of probabilities of identity by descent. We define a matrix \mathbf{F}_t giving probabilities of identity of two randomly chosen genes from the various

* Present address: Ruakura Agricultural Research Centre, Private Bag, Hamilton, New Zealand.

age groups at time t , and show that if second-order terms are ignored, we have approximately

$$\mathbf{F}_{t+1} = \mathbf{P} \mathbf{F}_t \mathbf{P}' + \mathbf{D}$$

where \mathbf{P} is a matrix specifying the passage of genes between the different age groups and \mathbf{D} is a diagonal matrix whose elements depend on the number of individuals in each age group. Using properties of the stochastic matrix \mathbf{P} , FELSSENSTEIN's formula for the effective number is obtained. We then extend these results to the case of separate sexes. The matrix \mathbf{P} is more general in this case, in that it is composed of blocks representing the four pathways of genes from male and female parents to male and female progeny. We also note that under our assumption of random births and deaths, the formula for the effective number quoted by HILL (1972) reduces to that obtained in this paper.

From this theory we show that, in the short term, the difference between the inbreeding coefficient and the uniform prediction reaches a constant value. Finally, we illustrate these results with an example of a farm livestock species.

PROBABILITIES OF IDENTITY BY DESCENT

Consider a population of monocious diploids with n age classes and a constant age distribution such that N_1 individuals enter the population each specified period of time. We let N_i denote the number of individuals of age i ($1 \leq i \leq n$) so that the probability of survival to age i is N_i/N_1 . We assume that the N_i values are reasonably large so that both births and deaths may be considered as independent random events.

Let p_i be the probability that a gene in a newborn individual came from a parent of age i . We define a matrix \mathbf{P} of the type discussed by LESLIE (1945) by

$$\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & \dots & \dots & p_{n-1} & p_n \\ 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 1 & 0 \end{bmatrix}$$

That is, the first row of \mathbf{P} consists of the p_1, p_2, \dots, p_n , all the sub-diagonal elements are equal to unity and all other elements are zero. The stochastic matrix \mathbf{P} has a single eigenvalue of unity and all others are of smaller absolute value. The vector $\mathbf{1}$ (with all elements equal to unity) is a right eigenvector for \mathbf{P} . We set

$$q_i = p_i + p_{i+1} + \dots + p_n \quad 1 \leq i \leq n$$

then the generation interval L is given by

$$L = \sum_{i=1}^n i p_i = \sum_{i=1}^n q_i .$$

The row vector \mathbf{q}' is a left eigenvector for \mathbf{P} with corresponding eigenvalue of unity. Using the theory of stochastic matrices (see, *e.g.*, KEMENY and SNELL 1960), it follows that the matrix \mathbf{A} where

$$\mathbf{A} = \lim_{t \rightarrow \infty} \mathbf{P}^t$$

is given by

$$\mathbf{A} = \mathbf{1} \mathbf{q}' / L .$$

We now consider probabilities of identity by descent. At a particular locus there are $2N_i$ genes among individuals of age i . Let $f_{ij}(t)$ be the probability that a gene chosen at random from among individuals of age i is identical by descent to a gene chosen at random from among individuals of age j at time t . We assume in the case $i=j$ that the genes are sampled with replacement from the gene pool of size $2N_i$. Let \mathbf{F}_t be the symmetric matrix of order n with $f_{ij}(t)$ in the (i,j) position. The following argument is similar to that of FELSENSTEIN (1971), although he considered probabilities of non-identity by descent.

We express the probabilities of identity by descent at time $t+1$ in terms of the probabilities at time t . If $i=j$ there is a probability of $1-1/2N_i$ of choosing two distinct genes from individuals of age i with further probabilities of p_r and p_s that these genes were descended from parents of ages r and s respectively at time $t-i+1$, and a probability of $1/2N_i$ that the same gene is sampled twice.

Thus

$$f_{ii}(t+1) = \frac{2N_i-1}{2N_i} \sum_{r,s=1}^n p_r p_s f_{rs}(t-i+1) + \frac{1}{2N_i} . \tag{1}$$

From (1) we then get

$$f_{11}(t+1) = \frac{2N_1-1}{2N_1} \sum_{r,s=1}^n p_r p_s f_{rs}(t) + \frac{1}{2N_1} \tag{2}$$

and

$$f_{ii}(t+1) = \left(\frac{2N_i-1}{2N_i}\right) \left(\frac{2N_{i-1}}{2N_{i-1}-1}\right) \left(f_{i-1,i-1}(t) - \frac{1}{2N_{i-1}}\right) + \frac{1}{2N_i} \quad 1 < i \leq n. \tag{3}$$

If we assume N_i large enough for

$$\frac{2N_i-1}{2N_i} \approx 1 \quad 1 \leq i \leq n \tag{4}$$

then equations (2) and (3) reduce to

$$f_{11}(t+1) = \sum_{r,s=1}^n p_r p_s f_{rs}(t) + 1/2N_1 \tag{5a}$$

and

$$f_{ii}(t+1) = f_{i-1,i-1}(t) + (1/2N_i - 1/2N_{i-1}) \quad 1 < i \leq n. \tag{5b}$$

If $i>1, j=1$ then clearly

$$f_{i1}(t+1) = \sum_{r=1}^n p_r f_{i-1,r}(t) \tag{5c}$$

and finally if $i,j>1, i \neq j$, then

$$f_{ij}(t+1) = f_{i-1,j-1}(t) . \tag{5d}$$

Define the diagonal matrix \mathbf{D} by setting

$$\begin{aligned} D_{11} &= 1/2N_1 \\ D_{ii} &= 1/2N_i - 1/2N_{i-1} & 1 < i \leq n \\ D_{ij} &= 0 & i \neq j . \end{aligned}$$

Then equations (5) may be written in matrix form as

$$\mathbf{F}_{t+1} = \mathbf{P} \mathbf{F}_t \mathbf{P}' + \mathbf{D} \quad t \geq 0 \quad (6)$$

where \mathbf{F}_0 is the diagonal matrix obtained by assuming that individuals in the population at time 0 are unrelated, that is,

$$\begin{aligned} f_{ii}(0) &= 1/2N_i & 1 \leq i \leq n \\ f_{ij}(0) &= 0 & i \neq j . \end{aligned}$$

From (6) it follows that

$$\mathbf{F}_t = \mathbf{P}^t \mathbf{F}_0 \mathbf{P}'^t + \sum_{i=0}^{t-1} \mathbf{P}^i \mathbf{D} \mathbf{P}'^i \quad t \geq 1 . \quad (7)$$

We recall the approximations (4) used to obtain the recurrence relation (6). This is equivalent to ignoring terms of order $1/N_i^2$ and it is therefore important to note that this assumption limits the application of equation (7) to values of t such that t/N_i is small. From (7) we get

$$\mathbf{F}_t - \mathbf{F}_{t-1} = \mathbf{P}^t \mathbf{F}_0 \mathbf{P}'^t - \mathbf{P}^{t-1} \mathbf{F}_0 \mathbf{P}'^{t-1} + \mathbf{P}^{t-1} \mathbf{D} \mathbf{P}'^{t-1} \quad t \ll N_i .$$

Since \mathbf{A} is the matrix $\lim_{t \rightarrow \infty} \mathbf{P}^t$ it follows that for fairly large t , but still satisfying $t \ll N_i$,

$$\mathbf{F}_t - \mathbf{F}_{t-1} \approx \mathbf{A} \mathbf{D} \mathbf{A}' .$$

With random mating the inbreeding coefficient $f(t)$ for progeny born at time t is given by

$$f(t) = \mathbf{p}' \mathbf{F}_{t-1} \mathbf{p}$$

where \mathbf{p}' is the first row of the matrix \mathbf{P} . Thus

$$f(t+1) - f(t) = \mathbf{p}' (\mathbf{F}_t - \mathbf{F}_{t-1}) \mathbf{p}$$

and so for t sufficiently large we obtain

$$f(t+1) - f(t) \approx \mathbf{p}' \mathbf{A} \mathbf{D} \mathbf{A}' \mathbf{p} = (\mathbf{q}' \mathbf{D} \mathbf{q}) / L^2 \quad t \ll N_i .$$

Thus during the early linear phase of increase of the inbreeding coefficient, its increment per generation reaches a value of $1/2N_e$ where N_e , the effective population size, is given by

$$N_e = \frac{L}{2\mathbf{q}' \mathbf{D} \mathbf{q}} = \frac{L}{\frac{1}{N_1} + \sum_{i=2}^n q_i^2 \left(\frac{1}{N_i} - \frac{1}{N_{i-1}} \right)}$$

as obtained by FELSENSTEIN (1971).

SEPARATE SEXES

We consider the same conditions as before but with separate sexes. We assume M_1 males and F_1 females enter the population each unit of time and that males and females are retained in the population for h and k mating seasons respec-

tively. The number of males of age i is M_i ($1 \leq i \leq h$) and the number of females of age j is F_j ($1 \leq j \leq k$).

Let p_{1i} ($1 \leq i \leq h$) and $p_{1,h+j}$ ($1 \leq j \leq k$) be the expected proportion of genes in male progeny which are derived from sires of age i and dams of age j respectively. Similarly $p_{h+1,i}$ and $p_{h+1,h+j}$ are the expected proportions of genes in female progeny derived from the various parental age groups. Following HILL (1974), we define a more general matrix \mathbf{P} of order $h+k$ by

$$\mathbf{P} = \left[\begin{array}{ccccc|cccc} p_{11} & p_{12} & \dots & p_{1,h-1} & p_{1h} & p_{1,h+1} & \dots & p_{1,h+k-1} & p_{1,h+k} \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ : & : & & : & : & : & & : & : \\ : & : & & : & : & : & & : & : \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 \\ \hline p_{h+1,1} & p_{h+1,2} & \dots & p_{h+1,h-1} & p_{h+1,h} & p_{h+1,h+1} & \dots & p_{h+1,h+k-1} & p_{h+1,h+k} \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ : & : & & : & : & : & & : & : \\ : & : & & : & : & : & & : & : \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \end{array} \right].$$

The blocks of \mathbf{P} correspond to the alternative pathways of genes between the different age groups and sexes, that is

$$\left[\begin{array}{c|c} \text{males from males} & \text{males from females} \\ \hline \text{females from males} & \text{females from females} \end{array} \right].$$

All the sub-diagonal elements of \mathbf{P} are equal to unity and represent the passage of genes due to ageing, the elements in the first row of each block represent the passage of genes due to reproduction. All other elements of \mathbf{P} are equal to zero. Clearly

$$\sum_{j=1}^h p_{ij} = \sum_{j=1}^k p_{i,h+j} = \frac{1}{2} \quad i=1,h+1$$

so all rows of \mathbf{P} sum to unity. For $r=1,h+1$ we define the quantities

$$q_{ri} = \begin{cases} \sum_{j=i}^h p_{rj} & 1 \leq i \leq h \\ \sum_{j=i}^{h+k} p_{rj} & h+1 \leq i \leq h+k \end{cases}$$

and let \mathbf{q} be the vector with elements

$$q_i = q_{1i} + q_{h+1,i} \quad 1 \leq i \leq h+k.$$

Let $L_{mm} = 2 \sum_{i=1}^h i p_{1i}$ and $L_{fm} = 2 \sum_{j=1}^k j p_{1,h+j}$ be the average ages of sires and Dams respectively of newborn males. Similarly define L_{mf} and L_{ff} for newborn

females (see Table 1). Then L , the generation interval, is the average age of parents along the four pathways of gametes and is given by

$$L = \frac{1}{4} (L_{mm} + L_{fm} + L_{mf} + L_{ff})$$

$$= \frac{1}{2} \sum_{i=1}^{h+k} q_i \quad .$$

The matrix \mathbf{P} is stochastic and has a single eigenvalue of unity with all others of smaller absolute value. The vector $\mathbf{1}$ is a right eigenvector for \mathbf{P} and \mathbf{q}' is a left eigenvector, corresponding to the eigenvalue of unity. The matrix \mathbf{A} where

$$\mathbf{A} = \lim_{t \rightarrow \infty} \mathbf{P}^t$$

is given by

$$\mathbf{A} = \mathbf{1} \mathbf{q}' / 2L \quad .$$

In the following, an "individual of age i " will refer to a male of age i if $1 \leq i \leq h$ or a female of age $i-h$ if $h+1 \leq i \leq h+k$. The $f_{ij}(t)$ values are defined as in the previous section, and \mathbf{F}_t is the corresponding symmetric matrix of order $h+k$. Suppose we select two genes at random from among males of age i at time $t+1$. There is a probability of $1/M_i$ that the two genes are chosen from the same individual with an equal chance that the same gene is sampled twice, or that both the sire gene and dam gene of that individual have been sampled. In the latter case the genes are derived from a sire of age r and a dam of age s at time $t-i+1$ with probability $4 p_{1r} p_{1,h+s}$ and so are identical with probability $4 \sum_{r=1}^h \sum_{s=1}^k p_{1r} p_{1,h+s} f_{r,h+s}(t-i+1)$, (we extend the definition of \mathbf{F}_t so that $\mathbf{F}_t = \mathbf{0}$ when $t < 0$). On the other hand, the two genes are sampled from different individuals with probability $1-1/M_i$, with further probabilities of p_{1r} and p_{1s} that these two genes are derived from individuals of ages r and s respectively at time $t-i+1$.

Thus

$$f_{ii}(t+1) = \frac{1}{2M_i} [1 + 4 \sum_{r=1}^h \sum_{s=1}^k p_{1r} p_{1,h+s} f_{r,h+s}(t-i+1)]$$

$$+ \frac{M_i - 1}{M_i} \sum_{r,s=1}^{h+k} p_{1r} p_{1s} f_{rs}(t-i+1) \quad 1 \leq i \leq h \quad . \quad (8a)$$

We also have

$$f_{i1}(t+1) = \sum_{r=1}^{h+k} p_{1r} f_{i-1,r}(t) \quad 1 < i \leq h \quad (8b)$$

with a similar set of equations for females.

If we select two genes from a male of age 1 and a female of age 1 then there are probabilities $p_{1,r}$ and $p_{h+1,s}$ that the genes were derived from individuals of ages r and s respectively. Thus

$$f_{1,h+1}(t+1) = \sum_{r,s=1}^{h+k} p_{1r} p_{h+1,s} f_{rs}(t) \quad (8c)$$

and similarly

$$f_{1,h+j}(t+1) = \sum_{r=1}^{h+k} p_{1r} f_{r,h+j-1}(t) \quad 1 < j \leq k \quad (8d)$$

$$f_{i,h+1}(t+1) = \sum_{r=1}^{h+k} p_{h+1,r} f_{i-1,r}(t) \quad 1 < i \leq h \quad (8e)$$

Finally

$$f_{ij}(t+1) = f_{i-1,j-1}(t) \quad \begin{matrix} i \neq j \\ i, j \neq 1, h+1 \end{matrix} \quad (8f)$$

Define the diagonal matrix **D** of order $h+k$ by setting

$$\begin{aligned} D_{11} &= 1/2M_1 \\ D_{ii} &= 1/2M_i - 1/2M_{i-1} & 1 < i \leq h \\ D_{h+1,h+1} &= 1/2F_1 \\ D_{h+j,h+j} &= 1/2F_j - 1/2F_{j-1} & 1 < j \leq k \\ D_{ij} &= 0 & i \neq j \end{aligned}$$

Then, ignoring terms of order $1/M_i^2$ and $1/F_j^2$, equations (8) may be written in the form

$$\mathbf{F}_{t+1} = \mathbf{P} \mathbf{F}_t \mathbf{P}' + \mathbf{D} \quad t \geq 0 \quad (9)$$

where **F**₀ is the diagonal matrix with elements

$$\begin{aligned} f_{ii}(0) &= 1/2M_i & 1 \leq i \leq h \\ f_{h+j,h+j}(0) &= 1/2F_j & 1 \leq j \leq k \\ f_{ij}(0) &= 0 & i \neq j \end{aligned}$$

Using the same reasoning as in the last section, it follows that for moderately large values of t satisfying $t \ll M_i, F_j$ we have

$$\mathbf{F}_t - \mathbf{F}_{t-1} \approx \mathbf{A} \mathbf{D} \mathbf{A}' = \frac{1}{4L^2} (\mathbf{q}' \mathbf{D} \mathbf{q}) \mathbf{1} \mathbf{1}'$$

Thus, during the early linear phase of increase in inbreeding, the increment per generation reaches the value of $1/2N_e$ where N_e , the effective population size, is given by

$$\begin{aligned} \frac{1}{N_e} &= \mathbf{q}' \mathbf{D} \mathbf{q} / 2L \\ &= \frac{1}{4L} \left[\frac{1}{M_1} + \sum_{i=2}^h q_i^2 \left(\frac{1}{M_i} - \frac{1}{M_{i-1}} \right) \right] \\ &\quad + \frac{1}{4L} \left[\frac{1}{F_1} + \sum_{j=2}^k q_{h+j}^2 \left(\frac{1}{F_j} - \frac{1}{F_{j-1}} \right) \right] \quad (10) \end{aligned}$$

Under more general conditions the effective size of a random mating population, with constant age distribution and with M_1 males and F_1 females entering the population each time period, is given by

$$\begin{aligned} \frac{1}{N_e} &= \frac{1}{16M_1L} [2 + \sigma_{mm}^2 + (M_1/F_1)^2 \sigma_{mf}^2 + 2(M_1/F_1) \text{cov}(mm,mf)] \\ &\quad + \frac{1}{16F_1L} [2 + \sigma_{ff}^2 + (F_1/M_1)^2 \sigma_{fm}^2 + 2(F_1/M_1) \text{cov}(fm,ff)] \quad (11) \end{aligned}$$

where the variances of lifetime family size are defined in Table 1 and $\text{cov}(mm,mf)$ and $\text{cov}(fm,ff)$ are the covariances between the lifetime number of male and female progeny from male and female parents respectively. Formula

TABLE 1

Means and variances of lifetime family size and mean age of parents

Pathway for gametes	Family size		Mean age of parents
	Mean	Variance	
Male parents having male progeny	1	σ^2_{mm}	L_{mm}
Male parents having female progeny	F_1/M_1	σ^2_{mf}	L_{mf}
Female parents having male progeny	M_1/E_1	σ^2_{fm}	L_{fm}
Female parents having female progeny	1	σ^2_{ff}	L_{ff}

(11) is based on a derivation by LATTER (1959), for discrete generation models, and a generalization to the overlapping case by HILL (1972). Under the assumption of random births and deaths one can express the above variances and covariances in terms of the quantities q_{ri} and the M_i and F_j . We get

$$\begin{aligned} \sigma^2_{mm} &= 1 + 4M_1 \sum_{i=2}^h q^2_{1i} \left(\frac{1}{M_i} - \frac{1}{M_{i-1}} \right) \\ \sigma^2_{mf} &= \frac{F_1}{M_1} + 4\frac{F_1^2}{M_1} \sum_{i=2}^h q^2_{h+1,i} \left(\frac{1}{M_i} - \frac{1}{M_{i-1}} \right) \\ \text{cov}(mm,mf) &= 4F_1 \sum_{i=2}^h q_{1i}q_{h+1,i} \left(\frac{1}{M_i} - \frac{1}{M_{i-1}} \right) \end{aligned}$$

and similarly for female parents. Substituting in (11) and recalling that $q_i = q_{1i} + q_{h+1,i}$ we obtain (10).

The inbreeding coefficient $f_m(t)$ for male progeny born at time t is defined in terms of the probability of identity of two genes, one from a male and the other from a female, chosen at time $t-1$. The probability that a male born at time t has a sire of age i and a dam of age j is $4p_{1i}p_{1,h+j}$ and so

$$f_m(t) = 4 \sum_{i=1}^h \sum_{j=1}^k p_{1i}p_{1,h+j} f_{i,h+j}(t-1)$$

with a similar definition for female progeny.

If we define vectors \mathbf{p}_{mm} and \mathbf{p}_{fm} of length $h+k$ by

$$\begin{aligned} \mathbf{p}'_{mm} &= (p_{11}, p_{12}, \dots, p_{1h}, 0, 0, \dots, 0) \\ \mathbf{p}'_{fm} &= (0, 0, \dots, 0, p_{1,h+1}, p_{1,h+2}, \dots, p_{1,h+k}) \end{aligned}$$

then we may write

$$f_m(t) = 4\mathbf{p}'_{mm} \mathbf{F}_{t-1} \mathbf{p}_{fm} .$$

We have shown that for moderate values of t these inbreeding coefficients attain a rate of increase of $1/2N_e$ per generation. Assuming a uniform rate of inbreeding from the outset, the inbreeding coefficient for both males and females born at time t would be $(t-1)/(2N_eL)$. Thus, the difference between $f_m(t)$ and this estimate is given by

$$\begin{aligned} f_m(t) - \frac{t-1}{2N_eL} &= 4\mathbf{p}'_{mm} \mathbf{F}_{t-1} \mathbf{p}_{fm} - (t-1)4\mathbf{p}'_{mm} \mathbf{A} \mathbf{D} \mathbf{A}' \mathbf{p}_{fm} \\ &= 4\mathbf{p}'_{mm} [\mathbf{F}_{t-1} - (t-1) \mathbf{A} \mathbf{D} \mathbf{A}'] \mathbf{p}_{fm} . \end{aligned}$$

TABLE 2

Age distribution of females in the sheep population

Age of ewe (<i>j</i>)	F_j	Lambing %	$p_{1,2+j} = p_{3,2+j}$
1	64	0	0.000
2	58	80	0.112
3	52	100	0.126
4	47	120	0.137
5	43	120	0.125

In the APPENDIX we show that the matrix $\mathbf{F}_t - t \mathbf{A} \mathbf{D} \mathbf{A}'$ is approximately constant for moderately large values of t , and so in the short term the difference between $f_m(t)$ and $(t-1)/2N_eL$ reaches a constant value.

EXAMPLE

We consider a sheep breeding program in which 20 rams are mated to a flock of 200 mixed-age ewes each year. Both rams and ewes have their first progeny at 2 years of age, but the rams are used only once, while the ewes are retained for 4 breeding seasons. Male and female replacements are chosen at random from the lamb drop each year.

We take $M_1 = 22$ so that a 10% death rate for both sexes and all ages will result in $M_2 = 20$. Since rams have no progeny at 1 year of age, we have $p_{11} = p_{31} = 0$ and $p_{12} = p_{32} = 1/2$. The age structure for females is given in Table 2. In Table 3 we present the inbreeding coefficient $f_m(t) = 4\mathbf{p}'_{mm} \mathbf{F}_{t-1} \mathbf{p}_{fm}$ where \mathbf{F}_{t-1} is calculated using the recurrence relation (9), for the first 12 years of the breeding program. As a comparison, the matrix \mathbf{F}_{t-1} is calculated using the exact equations (8) and the corresponding exact inbreeding coefficient is presented.

TABLE 3

Inbreeding coefficients for the sheep population

t	$f_m(t)$		$\frac{t-1}{2N_eL}$
	Using exact equations (8)	Using recurrence relation (9)	
1	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0011
3	0.0021	0.0021	0.0022
4	0.0022	0.0022	0.0034
5	0.0039	0.0042	0.0045
6	0.0043	0.0048	0.0056
7	0.0054	0.0061	0.0067
8	0.0060	0.0071	0.0079
9	0.0070	0.0083	0.0090
10	0.0080	0.0094	0.0101
11	0.0090	0.0105	0.0112
12	0.0100	0.0116	0.0123

The inbreeding coefficient calculated using recurrence relation (9) overestimates the exact value by approximately 16% after 12 years. These values are compared with the uniform inbreeding coefficient $(t-1)/2N_eL$, and this value overestimates the exact value by 23% after 12 years.

The effective population size is $N_e = 164$ and the generation interval $L = 2.775$. The difference between $f_m(t)$ and $(t-1)/2N_eL$ reaches a constant value of 0.0007.

DISCUSSION

In a population with discrete generations, inbreeding at any time can be described by a scalar quantity. With overlapping generations, a matrix formulation is necessary in order to describe the probability of identity of two genes chosen from the various age groups. A uniform rate of inbreeding of $1/2N_e$ per generation is attained only in the later stages of the early linear phase of increase of the inbreeding coefficient in a population with overlapping generations.

In the sheep example, computation of the short-term inbreeding using the inbreeding matrix F_t shows that initially the inbreeding coefficient differs considerably from that predicted by assuming a uniform rate of inbreeding from the outset, but that the difference between these two estimates settles down to a constant difference after about 9 years. These results are analogous to those of HILL (1974), who developed a theory for predicting the short-term response to selection in populations with overlapping generations, and this was compared with the classical theory of uniform rates of response.

Although in the short term the inbreeding calculated using the recurrence relation (9) is a fairly good approximation to the exact value, in practice it would seem preferable to compute the exact value, since there is very little extra effort in computation involved.

We emphasize some of the assumptions made in this paper. The recurrence relations (9) for the inbreeding matrix ignore second order terms. The population size is constant, as well as the parental age distribution. It is also assumed that deaths occur at random and that the distribution of family size from surviving parents is random at any one time period.

In another paper (JOHNSON 1976), a theory for predicting genetic drift in populations with overlapping generations has been developed. Formulae are given for the variance-covariance structure of group means in successive years of a selection program.

Thanks are due to DR. W. G. HILL at the Institute of Animal Genetics, Edinburgh, for his comments and suggestions and for the use of facilities at the Institute. Many helpful comments from DR. B. S. WEIR and the referees are also much appreciated. I am grateful to the Ministry of Agriculture and Fisheries, New Zealand, for financial support while on study leave in Edinburgh.

LITERATURE CITED

- FELSENSTEIN, J., 1971 Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics* **68**: 581-597.

- HILL, W. G., 1972 Effective size of populations with overlapping generations. *Theor. Pop. Biol.* **3**: 278-289. —, 1974 Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.* **18**: 117-139.
- JOHNSON, D. L., 1976 Variance-covariance structure of group means with overlapping generations. *Proceedings of the International Conference on Quantitative Genetics, August 1976, Iowa State University, Ames, Iowa.* (In press).
- KEMENY, I.G., and SNELL, J. L., 1960 *Finite Markov Chains.* Van Nostrand, Princeton, New Jersey.
- LATTER, B.D.H., 1959 Genetic sampling in a random mating population of constant size and sex ratio. *Aust. J. Biol. Sci.* **12**: 500-505.
- LESLIE, P. H., 1945 On the use of matrices in certain population mathematics. *Biometrika* **33**: 183-212.
- WRIGHT, S., 1931 Evolution in Mendelian populations. *Genetics* **16**: 97-159. —, 1938 Size of population and breeding structure in relation to evolution. *Science* **87**: 430-431.

Corresponding editor: W. J. EWENS

APPENDIX

We consider the asymptotic behavior of the matrix $F_t - t A D A'$. Following HILL (1974) we define a matrix B by setting

$$B = P - A .$$

Then $AB = BA = 0$ which implies that $P^t = A + B^t$ for $t \geq 1$, and hence $\lim_{t \rightarrow \infty} B^t = 0$. From (6) or (9) we have

$$\begin{aligned} F_t &= P^t F_0 P'^t + \sum_{i=0}^{t-1} P^i D P'^i \\ &= P^t F_0 P'^t + \sum_{i=1}^{t-1} (A + B^i) D (A' + B'^i) + D \\ &= P^t F_0 P'^t + (t-1) A D A' + \sum_{i=1}^{t-1} (B^i D A' + A D B'^i) + \sum_{i=0}^{t-1} B^i D B'^i . \end{aligned}$$

Now

$$\sum_{i=0}^{t-1} B^i = (I - B^t)(I - B)^{-1}$$

where I is the identity matrix. Hence we get

$$\begin{aligned} F_t - t A D A' &= P^t F_0 P'^t - A D A' + [(I - B^t)(I - B)^{-1} - I] D A' \\ &\quad + A D [(I - B^t)(I - B')^{-1} - I] + \sum_{i=0}^{t-1} B^i D B'^i . \end{aligned}$$

Thus for moderately large values of t we get approximately

$$\begin{aligned} F_t - t A D A' &\approx A(F_0 - D)A' + [(I - B)^{-1} - I] D A' \\ &\quad + A D [(I - B')^{-1} - I] + \sum_{i=0}^{\infty} B^i D B'^i \end{aligned}$$

the last term representing the sum of a convergent series. The matrix expression on the right hand side is constant.