Genet. Res., Camb. (1977), **29**, *pp.* 113–121
Printed in Great Britain

On the fixation probability of a gene under random fluctuations in selection intensities in small populations*

BY PREM NARAIN† AND EDWARD POLLAK
Statistical Laboratory, Iowa State University, Ames, Iowa 50011, U.S.A.

DEDICATED TO THE MEMORY OF THEODOSIUS DOBZHANSKY

(*Received* 22 *June* 1976)

SUMMARY

A population with N monoecious individuals, and having two alleles, is considered. The problem of calculating the fixation probability of a particular allele under random fluctuation of selection intensities is re-examined, employing finite Markov chain methods. An approximate but general expression for this probability is obtained and the results obtained by previous workers are shown to be special cases of this result.

1. INTRODUCTION

A problem in population genetics, which has received considerable attention in recent times, is that of the computation of the fixation probability of a gene and, along with it, the average time until its fixation. This is usually dealt with in two different ways. On the one hand, there is the diffusion approximation method, wherein the gene frequency is treated as a continuous random variable, between 0 and 1, and the time parameter of the underlying Markovian process is also taken as continuously varying. This method involves the use of the diffusion equations of Kolmogorov (1931), a mode of attack initiated for genetical problems by Fisher (1922), and the conditioned diffusion equations dealt with by Ewens (1973) and Narain (1974). On the other hand, a more exact treatment is that by a finite Markov chain involving the use of transition matrices. This method is discussed extensively in Kemeny & Snell (1960). Whichever of the two approaches is adopted, there is a basic question of whether the selection coefficients attached to the genes are constant over time or are fluctuating randomly. In regard to the latter, some results have recently been published by Jensen & Pollak (1969), Ohta (1972), Gillespie (1973) and Jensen (1973) as well as by Karlin & Levikson (1974). By and large, these results are based on the method of diffusion approximations and iteration procedures on the computer using transition matrices. According to Ohta (1972), if the ratio of the mean and variance of the selection coefficient is small, a mutant, even if

* Journal Paper No. J-8525 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project 1669.

† On leave from the Institute of Agricultural Research Statistics, Library Avenue, New Delhi, India.

selected against, becomes fixed in the population like a selectively neutral mutant. Contrary to this result, Jensen (1973) showed that the variability in the selection coefficient increases the chance of fixation of a rare gene. The problem was, however, attacked in a much more comprehensive manner by Karlin & Levikson (1974). In particular, they formulated a haploid model allowing for variability in the selection coefficients of both the types as well as for correlation between the two and showed that the variance in selection expression reduces and mitigates the mean effects of selection differentials, so that the fixation probability of the abundant allele is diminished. It seems that these different results on the fixation probability are probably due to the difference in the forms of the mean as well as variance functions for the change in gene frequency used in the diffusion approximation approach. However, if one does exact computations on the finite Markov chain, it seems that the choice of the appropriate mean and variance functions could be easily, as well as accurately, resolved and that an algebraic expression for the fixation probability could be obtained. According to Karlin & Levikson (1974), the work of Ohta (1972) suffers from an incorrect mean function. In this paper, it is shown that not only the mean but the variance function also needs correction, particularly for extremely small populations. An approximate but general expression for the fixation probability of a gene, in the haploid case, allowing for the variability in the selection coefficients of both the types as well as for the correlation between the two, is also derived, and the results obtained by the previous workers are shown to be special cases of this result.

2. THE MODEL

Consider a haploid population of $2N$ genes, corresponding to a monoecious population of individuals, of constant size N, reproducing in discrete generations. Let there be two alleles A and a with fitness coefficients in generation n as follows:

$$\begin{array}{cc} A & a \\ 1+s_1 & 1+s_2. \end{array} \tag{2.1}$$

The selection intensities, s_1 and s_2, are assumed to fluctuate over time in a random manner with identical distribution functions in all generations and independence between generations. The means, variances and covariances of these variables are

$$E(s_1) = \bar{s}_1, \quad E(s_2) = \bar{s}_2, \tag{2.2}$$

$$\text{var}\,(s_1) = v_1, \quad \text{var}\,(s_2) = v_2, \tag{2.3}$$

$$\text{cov}\,(s_1, s_2) = r, \quad \text{where} \quad |r| \leqslant \sqrt{(v_1 v_2)}. \tag{2.4}$$

We are virtually considering the haploid model of Karlin & Levikson (1974) but with the difference that means of the selection effects and their variances and covariances are not taken to be of the order of magnitude of $1/2N$.

The frequency of the A-gene in generation n, given its frequency before selection was $p_i = i/2N$ and given the selection parameters (s_1, s_2) is

$$p_i^{(n)} = p_i + p_i(1-p_i)(s_1 - s_2)/[1 + s_2 + (s_1 - s_2)p_i]. \tag{2.5}$$

https://doi.org/10.1017/S0016672300017183 Published online by Cambridge University Press

According to the standard Wright–Fisher Markov chain process, the distribution of the proportion of A-genes in generation n follows the binomial distribution with parameters $(2N, p_i^{(n)})$. In other words, the transition probability P_{ij} representing the conditional probability that there are j A-genes out of $2N$ genes, given that there were i A-genes in the population in the previous generation, is given by

$$P_{ij} = \binom{2N}{j} [p_i^{(n)}]^j [1-p_i^{(n)}]^{2N-j}, \tag{2.6}$$

where $i, j = 0, 1, 2, \ldots, 2N$.

The finite Markov chain, thus generated, could be studied by using results of Kemeny & Snell (1960). However, Narain & Robertson (1969) gave an analytical treatment of such processes in a genetic context and we follow the same procedure here.

3. EXPANSION OF THE TRANSITION PROBABILITY

Treating P_{ij} as a function of (s_1, s_2), we expand it, by Taylor's expansion, as a series in two variables. Neglecting terms involving powers and products of s_1 and s_2 greater than 2, we get

$$P_{ij} = P_{ij}(0)\,[1 + s_1 a_1 + s_2 a_2 + s_1^2 b_{11} + s_1 s_2 b_{12} + s_2^2 b_{22}], \tag{3.1}$$

where

$$P_{ij}(0) = \binom{2N}{j} p_i^j (1-p_i)^{2N-j}, \tag{3.2}$$

$$a_1 = 2N(p_j - p_i), \tag{3.3}$$

$$a_2 = -2N(p_j - p_i), \tag{3.4}$$

$$b_{11} = N[(2N-1)(p_j-p_i)^2 - p_j(1-p_j) - 2p_i(p_j-p_i)], \tag{3.5}$$

$$b_{12} = 2N[-(2N-1)(p_j-p_i)^2 + p_j(1-p_j) - (1-2p_i)(p_j-p_i)], \tag{3.6}$$

$$b_{22} = N[(2N-1)(p_j-p_i)^2 - p_j(1-p_j) + 2(1-p_i)(p_j-p_i)]. \tag{3.7}$$

If we now compute the expectation of (3.1) with respect to the distribution of s_1 and s_2, we have, in view of (2.2), (2.3) and (2.4),

$$\begin{aligned} E(P_{ij}) = P_{ij}(0)\,[1 &+ 2N\{(\bar{s}_1 - \bar{s}_2)(1-\bar{s}_2) + v_2 - r\}(p_j - p_i) \\ &+ N\{(\bar{s}_1 - \bar{s}_2)^2 + v_1 + v_2 - 2r\} \\ &\times \{(2N-1)(p_j-p_i)^2 - p_j(1-p_j) - 2p_i(p_j-p_i)\}], \end{aligned} \tag{3.8}$$

where the expectation is taken over the distribution of s_1 and s_2. Because the selection intensities are independent between generations and have the same distribution in each generation, the expression (3.8) holds as a one-step transition probability for any n. In other words, a Markov chain characterized by (3.8) is homogeneous in time.

4. MEAN AND VARIANCE FUNCTION FOR THE FREQUENCY OF A

In order to obtain the various moments of the frequency of the A-gene in the next generation, we proceed to obtain expressions for the expected values of expressions $j(j-1)\dots(j-k+1)$, given $p_i = i/2N$. Such expectations will be denoted by $E_i(j(j-1)\dots(j-k+1))$.

$$E_i[j(j-1)\dots(j-k+1)]$$

$$= \binom{2N}{j} k!\, p_i^k[1 + k(1-p_i)\{1 - \tfrac{1}{2}(\bar{s}_1+\bar{s}_2) + \tfrac{1}{2}(k(1-p_i)-p_i)(\bar{s}_1-\bar{s}_2)\}(\bar{s}_1-\bar{s}_2)$$

$$+ \tfrac{1}{2}k(1-p_i)(k(1-p_i)-p_i)(v_1+v_2-2r) - \tfrac{1}{2}k(1-p_i)(v_1-v_2)]. \quad (4.1)$$

The gene frequency expected in the next generation on the basis of binomial sampling is therefore obtained by putting $k = 1$ in (4.2). Denoting it by p_i^*, we get

$$p_i^* = p_i + E_i(\Delta p) = p_i + p_i(1-p_i)$$

$$\times \left[(\bar{s}_1-\bar{s}_2)[1 - \{\bar{s}_2 + (\bar{s}_1-\bar{s}_2)p_i\}] + \frac{v_1+v_2-2r}{2}(1-2p_i) - \frac{v_1-v_2}{2}\right], \quad (4.2)$$

where $E_i(\Delta p)$ is the expected mean change in the gene frequency in one generation, given that the frequency is $p_i = i/2N$.

Putting $k = 2$ in (4.1) and using (4.2), we get the variance of the change in gene frequency per generation, given that $p_i = i/(2N)$, $V_i(\Delta p_i)$, to the same degree of approximation, as

$$V_i[\Delta p] = \frac{p_i(1-p_i)}{2N}[1 + (1-\bar{s}_2)(\bar{s}_1-\bar{s}_2)(1-2p_i) - (\bar{s}_1-\bar{s}_2)^2 p_i(2-3p_i)$$

$$-(v_1+v_2-2r)p_i(2-3p_i) + (v_2-r)(1-2p_i)] + (v_1+v_2-2r)p_i^2(1-p_i)^2. \quad (4.3)$$

Expressions (4.2) and (4.3) can also be derived directly by manipulating expectations. An alternative expression for $V_i[\Delta p]$, in terms of p_i given by (4.2), can be written as

$$V_i[\Delta p] = \frac{p_i^*(1-p_i^*)}{2N} + \left(1 - \frac{1}{2N}\right)(v_1+v_2-2r)p_i^2(1-p_i)^2. \quad (4.4)$$

It is evident from the expressions for $E_i(\Delta p)$ and $V_i[\Delta p]$ derived above that random fluctuations in selection intensities affects both of them. These expressions can be compared with (3.7) and (3.8) on page 392 in the paper by Karlin & Levikson (1974) which may be written in the notation of this paper as

$$E_i(\Delta p) = p_i(1-p_i)[\bar{s}_1 - \bar{s}_2 - \tfrac{1}{2}(v_1-v_2) + \tfrac{1}{2}(v_1+v_2-2r)(1-2p_i)]$$

and

$$V_i(\Delta p) = \frac{p_i(1-p_i)}{2N} + (v_1+v_2-2r)p_i^2(1-p_i)^2$$

respectively. It is clear from (4.2) that our expression for $E_i(\Delta p)$ has an extra term: $-(\bar{s}_1-\bar{s}_2)(\bar{s}_2+(\bar{s}_1-\bar{s}_2)p_i)\,p_i(1-p_i)$. In expression (4.4) there are two types of terms associated with the non-additivity of the parts of $V_i(\Delta p)$ that arise if we consider

separately random changes due to (i) random sampling of gametes and (ii) random fluctuations in selection. In all the studies made so far, i.e. by Kimura (1962), Ohta (1972), Jensen (1973), as well as Karlin & Levikson (1974), this non-additivity is ignored.

Apart from this, the expression for $V_i(\Delta p)$ needs to be reconsidered, even when there is only random sampling of gametes and no random fluctuations in selection. With non-random selection ($v_1 = v_2 = r = 0$) the expressions for $E_i(\Delta p)$ and $V_i(\Delta p)$ become

$$E_i[\Delta p] = (\bar{s}_1 - \bar{s}_2)(1 - \bar{s}_2 - (\bar{s}_1 - \bar{s}_2)p_i)p_i(1-p_i), \tag{4.5}$$

$$V_i[\Delta p] = \frac{p_i(1-p_i)}{2N} + (1 - 2p_i - E_i[\Delta p])\frac{E_i[\Delta p]}{2N}. \tag{4.6}$$

It may be observed that the usual variance due to binomial sampling is strictly true only when $E_i[\Delta p] = 0$, i.e. the selectively neutral case. Even when the changes in gene frequency per generation due to selection are very small so that squares of $E_i[\Delta p]$ can be neglected, the binomial sampling variance holds only when $p_i = \frac{1}{2}$.

5. FIXATION PROBABILITY OF THE *A*-GENE

Let $u(p_i)$ be the fixation probability of the A-gene, given that initially at $t = 0$, it had frequency p_i, so that the total expected change in the frequency at the limit is

$$L_i = u(p_i) - p_i. \tag{5.1}$$

Narain & Robertson (1969) showed that the vector $\mathbf{L} = (L_1, L_2, \ldots, L_{2N-1})'$ is obtained by operating the matrix $\mathbf{T} = (\mathbf{I} - \mathbf{Q})^{-1}$ on to the vector

$$\mathbf{E}(\Delta p) = [E_1(\Delta p), E_2(\Delta p), \ldots, E_{2N-1}(\Delta p)]',$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{Q}$ is the matrix of transition probabilities when we consider transitions between the transient states only; i.e. $i, j = 1, 2, \ldots, 2N-1$. Expressed in powers and products of s_1 and s_2 up to terms involving $\{s_1\}^2$, $\{s_2\}^2$ and $\{s_1 s_2\}$, $\mathbf{T}$ is given by

$$\begin{aligned}\mathbf{T} = \mathbf{T}_0 &+ s_1\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0 + s_2\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0 + s_1^2[\mathbf{T}_0\mathbf{B}_{11}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_1\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0] \\ &+ s_1 s_2[\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0 + \mathbf{T}_0\mathbf{B}_{12}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_1\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0] \\ &+ s_2^2[\mathbf{T}_0\mathbf{B}_{22}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_2\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0],\end{aligned}$$

where $\mathbf{T}_0 = (\mathbf{I} - \mathbf{Q})^{-1}$ and $\mathbf{Q}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_{11}$, $\mathbf{B}_{12}$ and $\mathbf{B}_{22}$ are $(2N-1)\times(2N-1)$ matrices with $i-j$th elements $P_{ij}(0)$, $a_1 P_{ij}(0)$, $a_2 P_{ij}(0)$, $b_{11} P_{ij}(0)$, $b_{12} P_{ij}(0)$ and $b_{22} P_{ij}(0)$ respectively. Allowing for random variations in s_1 and s_2, as before, we get

$$\begin{aligned}E_{s_1, s_2}(\mathbf{T}) = \mathbf{T}_0 &+ (\bar{s}_1 - \bar{s}_2)(\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0) + (v_1 + \bar{s}_1^2)[\mathbf{T}_0\mathbf{B}_{11}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_1\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0] \\ &+ (r + \bar{s}_1\bar{s}_2)[\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0\mathbf{A}_1\mathbf{T}_0 + \mathbf{T}_0\mathbf{B}_{12}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_1\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0] \\ &+ (v_2 + \bar{s}_2^2)[\mathbf{T}_0\mathbf{B}_{22}\mathbf{T}_0 + \mathbf{T}_0\mathbf{A}_2\mathbf{T}_0\mathbf{A}_2\mathbf{T}_0].\end{aligned} \tag{5.2}$$

The expression $E(\Delta p)$ in powers and products of s_1 and s_2 up to terms involving

$\{s_1\}^2$, $\{s_2\}^2$ and $\{s_1 s_2\}$ and allowing for random variations in s_1 and s_2, in the same manner as before, is obtained by converting (4.2) into vector notation as

$$E[\Delta p] = \left(1 - \frac{\bar{s}_1 + \bar{s}_2}{2}\right)(\bar{s}_1 - \bar{s}_2)\,\mathbf{x}_1 + \tfrac{1}{2}(\bar{s}_1 - \bar{s}_2)^2\,\mathbf{x}_2 + \tfrac{1}{2}(v_1 + v_2 - 2r)\,\mathbf{x}_2 - \tfrac{1}{2}(v_1 - v_2)\,\mathbf{x}_1, \tag{5.3}$$

where vectors

$$\mathbf{x}_1 = [p_1(1-p_1), p_2(1-p_2), \ldots, p_{2N-1}(1-p_{2N-1})]', \tag{5.4}$$

$$\mathbf{x}_2 = [p_1(1-p_1)(1-2p_1), p_2(1-p_2)(1-2p_2), \ldots, p_{2N-1}(1-p_{2N-1})(1-2p_{2N-1})]', \tag{5.5}$$

correspond to the two eigenvalues of $\mathbf{Q}_0$ given by

$$\lambda_1 = (1 - 1/2N), \tag{5.6}$$

$$\lambda_2 = (1 - 1/2N)(1 - 2/2N). \tag{5.7}$$

Operating (5.2) on to (5.3) then gives, to the same degree of approximation,

$$\mathbf{L} = (\bar{s}_1 - \bar{s}_2)(\mathbf{T}_0 \mathbf{x}_1) + (\bar{s}_1 - \bar{s}_2)^2 \mathbf{T}_0 \mathbf{A}_1 \mathbf{T}_0 \mathbf{x}_1) - \tfrac{1}{2}(v_1 + \bar{s}_1^2)\{\mathbf{T}_0(\mathbf{x}_1 - \mathbf{x}_2)\} - (r + \bar{s}_1 \bar{s}_2)(\mathbf{T}_0 \mathbf{x}_2) + \tfrac{1}{2}(v_2 + \bar{s}_2^2)\{\mathbf{T}_0(\mathbf{x}_1 + \mathbf{x}_2)\}. \tag{5.8}$$

Using Table 1 of Narain & Robertson (1969), while noting that $\mathbf{A}_1$ is twice as large as $2\mathbf{Q}_0'$, we obtain

$$\mathbf{T}_0 \mathbf{x}_1 = 2N\mathbf{x}_1, \tag{5.9}$$

$$\mathbf{T}_0 \mathbf{A}_1 \mathbf{T}_0 \mathbf{x}_1 = \left(\frac{4N^3}{3N-1}\right)\lambda_1 \mathbf{x}_2, \tag{5.10}$$

$$\mathbf{T}_0 \mathbf{x}_2 = \left(\frac{2N^2}{3N-1}\right)\mathbf{x}_2. \tag{5.11}$$

With use of (5.9) to (5.11), the expression (5.8) becomes

$$\mathbf{L} = 2N(\bar{s}_1 - \bar{s}_2)\left(1 - \frac{\bar{s}_1 + \bar{s}_2}{2}\right)\mathbf{x}_1 + \frac{N^2(4N-1)}{3N-1}(\bar{s}_1 - \bar{s}_2)^2 \mathbf{x}_2 + \frac{N^2}{3N-1}(v_1 + v_2 - 2r)\,\mathbf{x}_2 - N(v_1 - v_2)\,\mathbf{x}_1. \tag{5.12}$$

With the non-random selection model, $v_1 = v_2 = r = 0$, so that if $\mathbf{L}_0$ denotes the corresponding vector of the expected change in the frequency of the A-gene at the limit, we have

$$\mathbf{L}_0 = 2N(\bar{s}_1 - \bar{s}_2)\left(1 - \frac{\bar{s}_1 + \bar{s}_2}{2}\right)\mathbf{x}_1 + \frac{N^2(4N-1)}{3N-1}(\bar{s}_1 - \bar{s}_2)^2 \mathbf{x}_2. \tag{5.13}$$

We can then express (5.12) as

$$\mathbf{L} = \mathbf{L}_0 - N(v_1 - v_2)\,\mathbf{x}_1 + \frac{N^2}{3N-1}(v_1 + v_2 - 2r)\,\mathbf{x}_2. \tag{5.14}$$

We can now examine $\mathbf{L}$ for three cases considered by previous workers on the problem of random variations in selection intensities.

(*a*) *Asymmetric case*

$\bar{s}_2 = v_2 = r = 0$, $v_1 > 0$. The changing environment allows variability only in the selection coefficient attached to the A-gene while its counterpart a-gene maintains the same constant selective value independent of the environmental background. We have

$$\mathbf{L}_0 = 2N\bar{s}_1(1-\bar{s}_1/2)\,\mathbf{x}_1 + \frac{N^2(4N-1)}{3N-1}\,\bar{s}_1^2\,\mathbf{x}_2, \tag{5.15}$$

$$\mathbf{L} = \mathbf{L}_0 - Nv_1\left[\mathbf{x}_1 - \left(\frac{N}{3N-1}\right)\mathbf{x}_2\right]. \tag{5.16}$$

If the initial frequency of **A** is $p_i = i/2N$, these expressions reduce to

$$u_0(p_i) = p_i + 2N\bar{s}_1(1-\bar{s}_1/2)\,p_i(1-p_i) + N^2\bar{s}_1^2\left(\frac{4N-1}{3N-1}\right)p_i(1-p_i)\,(1-2p_i), \tag{5.17}$$

$$u(p_i) = u_0(p_i) - Nv_1\left[\frac{2N(1+p_i)-1}{3N-1}\right]p_i(1-p_i). \tag{5.18}$$

This shows that the fixation probability $u(p_i)$ under random fluctuations in s_1 is always smaller than its value $u_0(p_i)$ under non-random selection. This is in conformity with the result of Karlin & Levikson (1974) obtained from the diffusion approach. However, they do not give any explicit formula for this case. We can, however, give such a formula by assuming N to be very large while $N\bar{s}_1$ and Nv_1 remain constant. The expressions (5.17) and (5.18) then give

$$u(p_i) = p_i + 2(N\bar{s}_1)\,p_i(1-p_i) + \tfrac{4}{3}(N\bar{s}_1)^2\,p_i(1-p_i)\,(1-2p_i) - \tfrac{2}{3}(Nv_1)\,p_i(1-p_i^2), \tag{5.19}$$

showing that the fixation probability is now a function of p_i, $N\bar{s}_1$ and Nv_1. For extremely small populations we have to use (5.18) in conjunction with (5.17).

(*b*) *Symmetric case in the sense of Karlin & Levikson (1974)*

In this case $\bar{s}_1 = \bar{s}_2$, $v_1 = v_2 = v > 0$, $r = 0$. Thus (5.13) and (5.14) respectively reduce to

$$\mathbf{L}_0 = \mathbf{0}, \tag{5.20}$$

$$\mathbf{L} = \frac{2N^2}{3N-1}\,v\mathbf{x}_2. \tag{5.21}$$

For the initial frequency p_i, we get

$$u_0(p_i) = p_i, \tag{5.22}$$

$$u(p_i) = p_i + \frac{2N^2v}{3N-1}\,p_i(1-p_i)\,(1-2p_i). \tag{5.23}$$

This shows that the fixation probability is smaller than p_i for $p_i > \frac{1}{2}$ and greater than p_i if $p_i < \frac{1}{2}$, which is consistent with what was found by Karlin & Levikson (1974). For large N and small v such that Nv remains constant, we get the approximation

$$u(p_i) = p_i + \tfrac{2}{3}(Nv)\,p_i(1-p_i)\,(1-2p_i) \tag{5.24}$$

giving that the fixation probability is a function of p_i and Nv.

(c) *Symmetric case in the sense of Jensen & Pollak (1969)*

We now have $\bar{s}_1 = \bar{s}_2 = \bar{s}$, $v_1 = v_2 = v > 0$, $r = -\bar{s}_1\bar{s}_2 = -\bar{s}^2 \geqslant -v$, which would result if either $s_1 > 0$, $s_2 = 0$ or $s_1 = 0$, $s_2 > 0$ in every generation. Expressions (5.13) and (5.14) now take the form

$$\mathbf{L}_0 = \mathbf{0}, \tag{5.25}$$

$$\mathbf{L} = \left(\frac{2N^2}{3N-1}\right)(v+\bar{s}_2)\,\mathbf{x}_2. \tag{5.26}$$

For initial frequency p_i, we get $u_0(p_i) = p_i$ as in Case (*b*) but $u(p_i)$ becomes

$$u(p_i) = p_i + \frac{2N^2(v+\bar{s}^2)}{3N-1}\,p_i(1-p_i)(1-2p_i), \tag{5.27}$$

which is the same as (5.23) except that v is replaced by $(v+\bar{s}^2)$. Thus we have essentially the same result as in case (*b*). Once again $u(p_i) > p_i$ if $p_i < \frac{1}{2}$ and $u(p_i) < p_i$ for $p_i > \frac{1}{2}$. This agrees with what was found by Jensen (1973). For large N and small v as well as $\bar{s}$ such that Nv and $N\bar{s}$ remain constant, the fixation probability is now a function of p_i, $N\bar{s}^2$ and Nv.

Going back to the general expressions (5.13) and (5.14) we can derive results which cover all the three cases mentioned above.

If the initial frequency of the A-gene is $p_i = i/2N$, (5.13) and (5.14) reduce to

$$u_0(p_i) = p_i + 2N(\bar{s}_1-\bar{s}_2)\left(1-\frac{\bar{s}_1+\bar{s}_2}{2}\right)p_i(1-p_i) + N^2(\bar{s}_1-\bar{s}_2)^2\left(\frac{4N-1}{3N-1}\right)p_i(1-p_i)(1-2p_i), \tag{5.28}$$

$$u(p_i) = u_0(p_i) - \frac{N}{3N-1}\,[(3N-1)(v_1-v_2) - N(v_1+v_2-2r)(1-2p_i)]\,p_i(1-p_i). \tag{5.29}$$

For large values of N and small values of $\bar{s}_1$, $\bar{s}_2$, v_1, v_2 and r such that $N(\bar{s}_1-\bar{s}_2)$, Nv_1, Nv_2 and Nr remain constant, we get the approximations:

$$u_0(p_i) = p_i + 2N(\bar{s}_1-\bar{s}_2)\,p_i(1-p_i) + \tfrac{4}{3}\{N(\bar{s}_1-\bar{s}_2)\}^2\,p_i(1-p_i)(1-2p), \tag{5.30}$$

$$u(p_i) = u_0(p_i) + \tfrac{2}{3}[(Nv_1+Nv_2-2Nr)\,p_i - (2Nv_2-Nv_1-Nr)]\,p_i(1-p_i). \tag{5.31}$$

We thus find that $u(p_i)$ is smaller or larger than $u_0(p_i)$ depending upon whether p_i is less than or greater than $(2v_2-v_1-r)/(v_1+v_2-2r)$ because $(v_1+v_2-2r) \geqslant 0$ in view of (2.3) and (2.4). If we define

$$\alpha = (2v_2-v_1-r)/(v_1+v_2-2r) \tag{5.32}$$

we find that, in general, the fixation probability of the A-gene, under random fluctuation in selection intensities, is determined solely by p_i, $N(\bar{s}_1-\bar{s}_2)$,

$N(v_1+v_2-2r)$ and α. It is thus given by

$$u(p_i) = p_i + 2\{N(\bar{s}_1-\bar{s}_2)\}p_i(1-p_i) + \tfrac{4}{3}\{N(\bar{s}_1-\bar{s}_2)\}^2 p_i(1-p_i)(1-2p_i) - \tfrac{2}{3}\{N(v_1+v_2-2r)\}p_i(1-p_i)(\alpha-p_i), \quad (5.33)$$

where α is given by (5.32).

While expression (5.33) is only approximate, it is a generally applicable explicit expression that holds regardless of the nature of the joint distribution of s_1 and s_2 in a generation. Previous authors have only been able to obtain an explicit solution in the symmetric cases. Moreover, expression (5.29) is valid for any population size, however small, with the restriction that we neglect moments of the third and higher orders in the selection intensities. This appears to be a new result in the literature.

The authors are grateful to Professor Oscar Kempthorne for helpful advice and valuable discussions. They are also grateful to Professor Alan Robertson for editorial suggestions which resulted in a considerable improvement in presentation.

This work received partial support from the National Institute of Health, Grant GM 13827.

REFERENCES

EWENS, W. J. (1973). Conditional diffusion processes in population genetics. *Theoretical Population Biology* **4**, 21–30.

FISHER, R. A. (1922). On the dominance ratio. *Proceedings of the Royal Society of Edinburgh* **42**, 321–341.

GILLESPIE, J. H. (1973). Natural selection with varying selection coefficients: a haploid model. *Genetical Research* **21**, 115–120.

JENSEN, L. (1973). Random selective advantages of genes and their probabilities of fixation. *Genetical Research* **21**, 215–219.

JENSEN, L. & POLLAK, E. (1969). Random selective advantages of a gene in a finite population. *Journal of Applied Probability* **6**, 19–37.

KARLIN, S. & LEVIKSON, B. (1974). Temporal fluctuations in selection intensities: Case of small population size. *Theoretical Population Biology* **6**, 383–412.

KEMENY, J. G. & SNELL, J. L. (1960). *Finite Markov Chains*. New York: D. van Nostrand.

KIMURA, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* **47**, 713–719.

KOLMOGOROV, A. (1931). Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen* **104**, 415–458.

NARAIN, P. & ROBERTSON, A. (1969). Limits and duration of response to selection in finite populations: The use of transition probability matrices. *Indian Journal of Heredity* **1**, 1–19.

NARAIN, P. (1974). The conditional diffusion equation and its use in population genetics. *Journal of the Royal Statistical Society* B **36**, 258–266.

OHTA, T. (1972). Fixation probability of a mutant influenced by random fluctuation of selection intensity. *Genetical Research* **19**, 33–38.