

## The effect of random selection coefficients on populations of finite size – some particular models

BY P. J. AVERY\*

*Department of Biomathematics, University of Oxford,  
Pusey St., Oxford OX1 2JZ*

(Received 4 June 1976)

### SUMMARY

The model of random selection coefficients is considered in the context of a finite population of diploids. The selection coefficients of the homozygotes are allowed to vary with equal variance while the fitness of the heterozygote is kept fixed. Steady-state solutions are found in the case of equal two-way mutation rates with particular reference to the expected heterozygosity. Increasing the variance of the selection coefficients of the homozygotes is found to uniformly increase the heterozygosity for all values of the average selection coefficients and its effect is largest when the selection coefficients of the homozygotes are fully correlated. The fate of mutant genes is also considered in the case of random selection coefficients by looking at the probability of ultimate fixation and the mean times to fixation and extinction. The errors in previous calculations (e.g. Kimura, 1954; Ohta, 1972) are pointed out. It is found that a small average heterozygote advantage together with a reasonable degree of variance in the coefficients can cause an unexpectedly large amount of heterozygosity to be maintained. It is also seen that probabilities of fixation and mean times to boundaries are usually increased by increasing the variance showing that it in fact helps to keep the population heterozygous for much longer than the non-random case. This is in contradiction to some conclusions of Karlin & Levikson (1974) because their haploid results are not easily extendable to the consideration of this sort of diploid model.

### 1. INTRODUCTION AND BASIC MODEL

Fluctuation of selection coefficients at a locus can be caused by random environmental changes. Most species inhabit an environment which is subject to change, sometimes very drastic change, and thus for an understanding of the effect of this variability, the model of random selection coefficients seems to be an important one. Kimura first worked on this topic some time ago (e.g. Kimura, 1954, 1964). However, it has been pointed out recently that his results, and later works based upon them, are in fact, incorrect. Jensen (1973) and Gillespie (1973) examined simple haploid models to exemplify the error, and Karlin & Levikson (1974) gave a general review of the effects of variability, concentrating more or less entirely on haploid models. Hartl & Cook (1973, 1975, 1976) have also examined this topic, particularly

\* Present address: Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN.

concentrating on the case when the alleles are on average neutral, which they call 'quasi-neutrality'. They have, however, ignored the effects of mutation and finite population size which are considered here. Takahata, Ishii & Matsuda (1975) attempt to extend the analysis of selection coefficients to include a correlation from generation to generation. However, the method in which they do so seems somewhat obscure. By mathematics, which this author finds difficult to follow, they produce a Diffusion Equation in which the term, due to the variance of selection coefficients, in the infinitesimal variance term,  $V_{\delta x}$ , is double that found with the independence approach (i.e. Karlin & Levikson, 1974). Otherwise the equation is unchanged. This seems to be a rather unexpected result and is not discussed by them.

Most species have diploid loci, rather than haploid ones. Thus the haploid results have been extended in this paper to diploids. With all selection schemes, one is predominantly concerned with relative fitnesses. This is particularly true if a diffusion equation approach is to be used as the selection coefficients are by necessity considered to be small. Thus, in this paper, the heterozygote fitness will be considered to be fixed. The inclusion of variability for all fitnesses as suggested by Karlin & Levikson (1974) seems an unnecessary complication. Because of this set-up of genotypic fitnesses, the results from haploid models cannot be easily extended to diploid models and thus a new appraisal is needed. In this paper, models of average heterozygote superiority (and inferiority) and average no-dominance selection will be considered. In the analysis of this paper, the magnitude of the errors of Ohta & Kimura (1972) and Ohta (1972) are pointed out. The number of quantities considered for the models is extended so that a greater appreciation of the effect of variability of selection coefficients can be obtained.

Two lines of approach will be taken. In both, a colony of finite size will be considered. First, equal forward and backward mutation rates will be allowed and the steady-state gene frequency distribution and the steady-state heterozygosity will be calculated. The aim of this approach is to see how the parameters of this model affect the genetic variability maintained in a population. Secondly, in order to see how random selection coefficients affect the rate of change of populations, the fate of a small number of mutant genes will be examined by considering the probabilities of ultimate fixation and extinction, and the mean times to each of these events, assuming that they will happen.

The basic model, used in both approaches is that of two alleles at a diploid locus with fitnesses at a generation,  $i$ , given by

Genotype	$AA$	$Aa$	$aa$
Fitness	$1 + \sigma$	$1$	$1 + \Upsilon$

where the selection intensities ( $\sigma, \Upsilon$ ) are assumed to fluctuate over time in a random manner, with identical distributions in all generations and independence between generations. In the following sections, the following symbols will be used:

$$E(\sigma) = s; \quad E(\Upsilon) = t; \quad \text{var}(\sigma) = \text{var}(\Upsilon) = V_{\sigma} \quad \text{and} \quad \text{cov}(\sigma, \Upsilon) = rV_{\sigma}.$$

2. MODEL WITH TWO-WAY MUTATION RATES

For simplicity, let us assume equal forward and backward mutation rates,  $v$  say. Then from the formulae produced by Karlin & Levikson (1974), we obtain the following expressions for the moments of the small change in gene frequency,  $x$ , of allele  $A$  per generation:

$$E(\delta x) = M_{\delta x} = -x(1-x)(t - (s+t)x) + V_{\sigma}x(1-x)(1-2x)(1 - (1+r)x(1-x)) + v(1-2x), \tag{2.1}$$

$$\text{var}(\delta x) = V_{\delta x} = V_{\sigma}x^2(1-x)^2(1 - 2(1+r)x(1-x)) + x(1-x)/2N_e,$$

under the assumption that  $s, t, V_{\sigma}$  and  $v$  are small and of the same order as  $(2N_e)^{-1}$ .

The steady-state gene frequency distribution  $\phi(x)$  is then found by using Wright's solution of Kolmogorov's Forward Diffusion Equation (Wright, 1938, 1949), i.e.

$$\phi(x) = \frac{C}{V_{\delta x}} \exp\left(2 \int \frac{M_{\delta x}}{V_{\delta x}} dx\right), \tag{2.2}$$

where  $C$  is a normalizing constant such that  $\int_0^1 \phi(x) dx = 1$ . Thus one has to evaluate  $I = \int (M_{\delta x}/V_{\delta x}) dx$ . In order to demonstrate the qualitative effects of the different parameters, we shall take the simplifying assumption that  $\alpha = N_e s = N_e t$  (i.e. the selection coefficients of the homozygotes have, on average, the same value). Later in this section, it will be shown that similar results are obtained in the non-symmetric case.

One can simplify the integrand of  $I$  by substitution and the subsequent use of partial fractions. The resultant functions can be integrated easily. Details are given in Appendix (i). On substituting the answer for  $I$ , which one obtains, in equation (2.2), one finds that for  $r \neq -1$  and  $V_{\sigma} \neq 0$

$$\phi(x) = C(\alpha_1 - x(1-x))^{\gamma_1} (x(1-x) - \alpha_2)^{\gamma_2} (x(1-x))^{\beta-1}, \tag{2.3}$$

where

$$\alpha_i = \frac{1}{4(1+r)} \left(1 \pm \sqrt{1 + \frac{8(1+r)}{V}}\right) \quad (i = 1, 2),$$

$$\gamma_i = -1 \mp \frac{(d - \alpha_i)}{\alpha_1 - \alpha_2} - \frac{\beta}{V\alpha_i + 2} \quad (i = 1, 2),$$

$$d = \frac{(1 - 2\alpha/v)}{1+r}, \quad V = 2N_e V_{\sigma}, \quad \text{and} \quad \beta = 4N_e v$$

If  $r = -1$ ,

$$\phi(x) = C(x(1-x) + V^{-1})^{1-\beta-4\alpha v} (x(1-x))^{\beta-1}. \tag{2.4}$$

Fig. 1 gives the steady state gene frequency distribution for some values of the parameters. One sees that, as  $V$  increases, a peak is produced around  $x = 0.5$ , which is typical of systems which maintain a high degree of heterozygosity. One further sees that the peak is most pronounced if  $r = 1$ , i.e. the two sets of random selection coefficients are completely correlated.

The moments of  $\phi(x)$ , and in particular  $H$ , the steady state heterozygosity, can

be found by use of numerical integration. In order to evaluate  $C$  accurately, a few manipulations need to be done (see Appendix (ii)). Having found  $C^{-1}$ ,  $H$  is easily found by using

$$H = \int_0^1 2x(1-x)\phi(x) dx.$$

Figs. 2 and 3 show the effect of the different parameters on the steady-state heterozygosity. Fig. 2 gives the effect of  $r$ , when  $s$  is zero, and one sees that, as one would

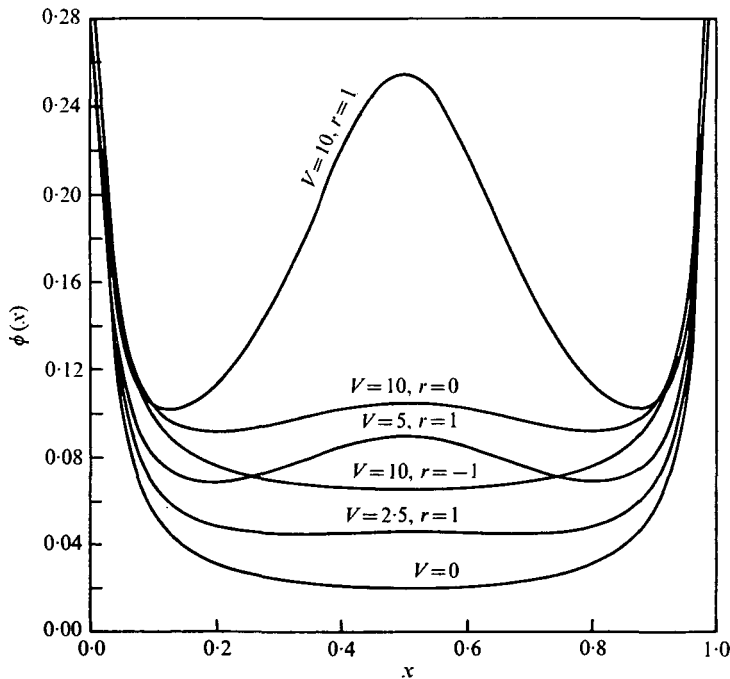


Fig. 1. The steady-state gene frequency distribution,  $\phi(x)$ , is plotted against the gene frequency,  $x$ , in the case of quasi-neutrality for various values of the variance and the correlation coefficient of the fitnesses.  $s = 0$ ,  $\beta = 4N_e v = 0.01$ ,  $V = 2N_e V_\sigma$ .

expect from Fig. 1, that  $H$  increases quite substantially with  $V$ , the effect being largest when  $r$  is 1, and decreasing with  $r$  for fixed values of  $V$ . The graphs marked 'UC' are the results which one obtains if one neglects the variance term in  $M_{\delta x}$  as was done by Kimura (1954) and in other later papers. The solutions are very similar algebraically but, as can be seen from the graphs, are very different when actual parameter values are inserted. For example, if  $r \in (-1, 1]$ , the uncorrected formula is

$$\phi(x) = C(\alpha_1 - x(1-x))^{\gamma_1} (x(1-x) - \alpha_2)^{\gamma_2} (x(1-x))^{\beta-1},$$

where

$$\gamma_i = -1 \pm \frac{2\alpha}{V(1+r)(\alpha_1 - \alpha_2)} - \frac{\beta}{V\alpha_i + 2} \quad (i = 1, 2). \tag{2.5}$$

Fig. 3 gives the effect of a non-zero mean,  $r$  being kept constant at one. As one

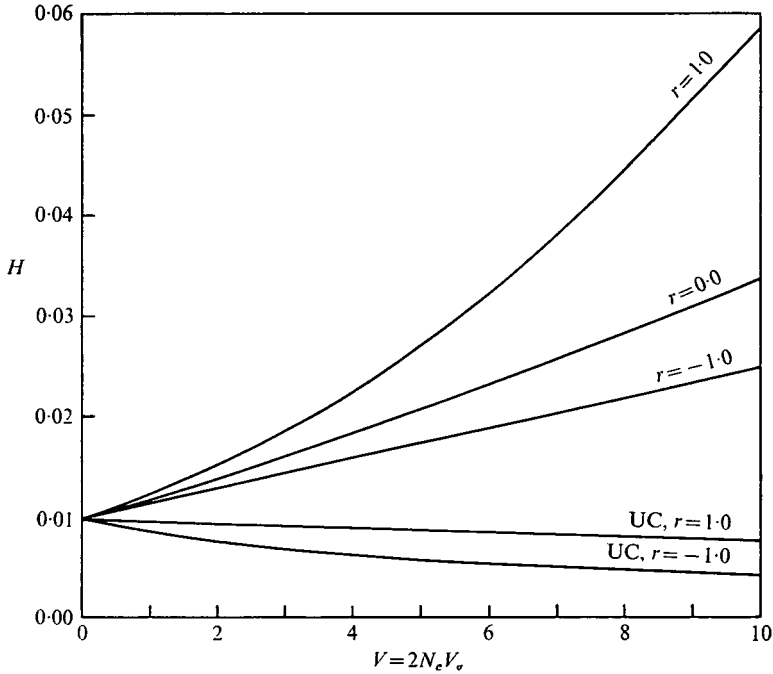


Fig. 2. The steady-state heterozygosity,  $H$ , is plotted against the variance factor,  $V$ , in the case of quasi-neutrality for various values of the correlation coefficient,  $r$ . The graphs marked 'UC' give the uncorrected solutions.  $\alpha = N_e s = 0$ ,  $\beta = 4N_e v = 0.01$ .

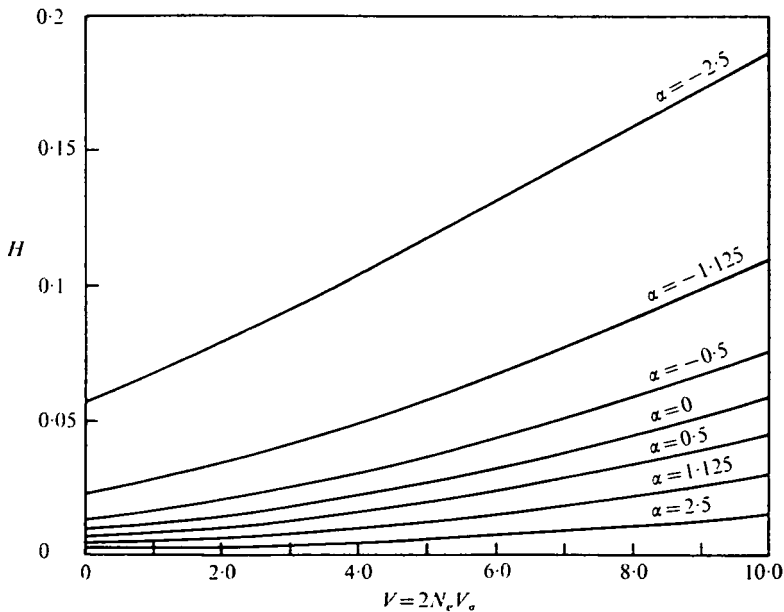


Fig. 3. The steady-state heterozygosity,  $H$ , is plotted against the variance factor,  $V$ , in the case of fully correlated coefficients (i.e.  $r = 1$ ), for various values of the average selection factor,  $\alpha$  ( $= N_e s$ ).  $\alpha = N_e s$ ,  $\beta = 4N_e v = 0.01$ .

might expect, increasing the mean,  $s$ , uniformly decreases  $H$  for all  $V$  as it creates average heterozygote disadvantage. The fact that average heterozygote superiority and a reasonable amount of variability produce much larger degrees of heterozygosity than one would otherwise expect is certainly important and perhaps is an explanation to some genetic variability.

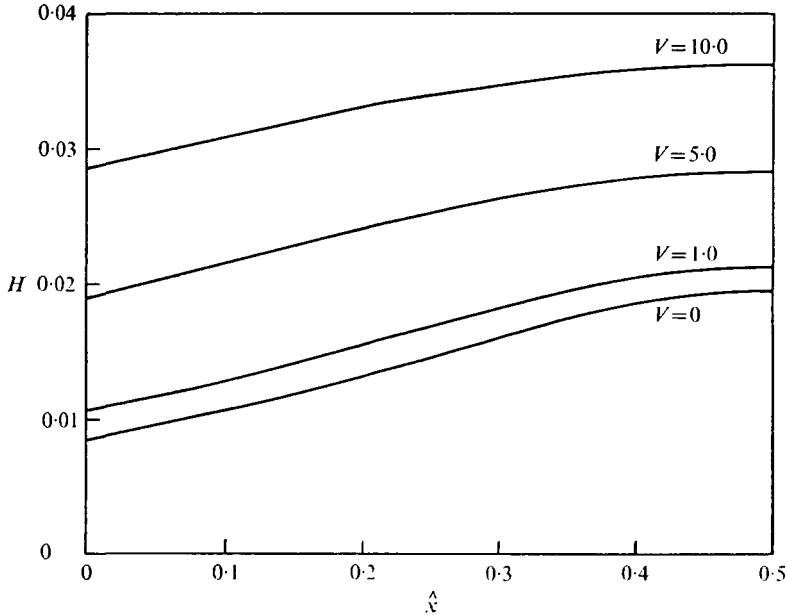


Fig. 4. The steady-state heterozygosity,  $H$ , is plotted against the degree of asymmetry for fixed average selection in the case of random heterozygote advantage ( $r = -1$ ).  $(\frac{1}{2}N_e)(s+t) = 1.0$ ,  $\hat{x} = t/s+t$ .

In order to show that these results are valid, when the heterozygote superiority is asymmetric, I have taken the case which is simplest, mathematically, i.e.  $r = -1$ . Already in this section, I have shown that this value of  $r$  is the least efficient value for maintaining a high value of  $H$ . Thus if we find that variability still leads to increased heterozygosity in the asymmetric case,  $r = -1$ , we can reasonably assume that the same is true for other values of  $r$ .

If  $r = -1$ , then proceeding as we did above, we obtain:

$$\phi(x) = C(x(1-x))^{\beta-1} (E_1-x)^{\lambda_1} (x-E_2)^{\lambda_2},$$

where

$$E_i = \frac{1}{2}[1 \pm \sqrt{(1+4V^{-1})}] \quad \text{and} \quad \lambda_i = 1 - \beta - 2\frac{(\delta_1 + \delta_2)}{V} \pm 2\frac{(\delta_2 - \delta_1)}{V\sqrt{(1+4V^{-1})}} \quad (i = 1, 2).$$

This simplifies to (2.4) if  $\delta_1 = \delta_2 = \alpha$  as  $(E_1-x)(x-E_2) = x(1-x) + V^{-1}$ . The effect of asymmetry is shown in Fig. 4 by plotting the heterozygosity,  $H$ , against

$$\hat{x} = \frac{t}{s+t} = \frac{\delta_2}{\delta_1 + \delta_2} = \frac{1}{2} \left[ \frac{\delta_2 - \delta_1}{\delta_1 + \delta_2} + 1 \right]$$

keeping  $\frac{1}{2}(\delta_1 + \delta_2) = \frac{1}{2}N_e(s+t)$  constant. One can see that  $V$  increases  $H$  for all values of  $\hat{x}$ .

3. FATE OF SMALL NUMBERS OF MUTANT GENES

In this section I shall consider the fate of mutant genes by looking at their probabilities of ultimate fixation and extinction, and the mean times to these events assuming that they take place. It is assumed that no further mutation occurs after the initial production of the mutant. The methods which I shall use depend only on  $p$ , the initial frequency of mutants but, in order that the diffusion approximation is valid, the colony size must not be too small. I shall look here at the fate of single mutant genes. Similar results are found if larger numbers of initial mutants are considered.

The formula for  $u(p)$ , the probability of ultimate fixation is derived from Kolmogorov's Backward Diffusion Equation (Kimura, 1962, 1964) and is given by

$$u(p) = \int_0^p G(x) dx / \int_0^1 G(x) dx, \tag{3.1}$$

where

$$G(x) = \exp\left[-2 \int_0^x \frac{M_{\delta x}}{V_{\delta x'}} dx'\right].$$

The mean times to fixation and extinction are given in general form by Kimura & Ohta (1969*a, b*). These times are conditional on the particular event happening. The two equations are:

$$\begin{aligned} \text{mean time to fixation} = \bar{t}_f &= \int_p^1 \psi(x) u(x) (1 - u(x)) dx + \frac{1 - u(p)}{u(p)} \\ &\int_0^p \psi(x) [u(x)]^2 dx \tag{3.2} \end{aligned}$$

and

$$\begin{aligned} \text{mean time to extinction} = \bar{t}_e &= \int_0^p \psi(x) u(x) (1 - u(x)) dx - \frac{u(p)}{1 - u(p)} \\ &\int_p^1 \psi(x) (1 - u(x))^2 dx, \tag{3.3} \end{aligned}$$

where  $\psi(x) = 2g(0, 1)/(V_{\delta x} G(x))$  and  $g(x, y) = \int_x^y G(z) dz$ .

I shall now consider two modelling schemes: that introduced above and a model of no dominance (i.e.  $\Upsilon = -\sigma$ ).

(i) Model 1

Genotype	<i>AA</i>	<i>Aa</i>	<i>aa</i>
Fitness	$1 + s + \sigma'$	1	$1 + s + \Upsilon'$

where  $\sigma' = \sigma - s$ ,  $\Upsilon' = \Upsilon - s$ ,  $E(\sigma') = E(\Upsilon') = 0$  and correlation coefficient of  $(\sigma', \Upsilon') = r$ .

Thus the equations for  $M_{\delta x}$  and  $V_{\delta x}$  are the same as equation (2.1) with  $s = t$ , except that the term in  $v(1 - 2x)$  has been removed from  $M_{\delta x}$ .

If the selection coefficients are fixed, i.e.  $V_\sigma = 0$ , then one obtains that

$$G(x) = \exp[4\alpha x(1-x)]. \tag{3.4}$$

If  $V_\sigma \neq 0$ , then  $I (= \int (M_{\delta x}/V_{\delta x}) dx)$ , calculated in section 2, can be used to evaluate  $G(x)$ . Thus one finds that, for  $r \neq -1$ .

$$G(x) = (1-x(1-x)/\alpha_1)^{\gamma_1^2} (1+x(1-x)/(-\alpha_2))^{\gamma_2^2}, \tag{3.5}$$

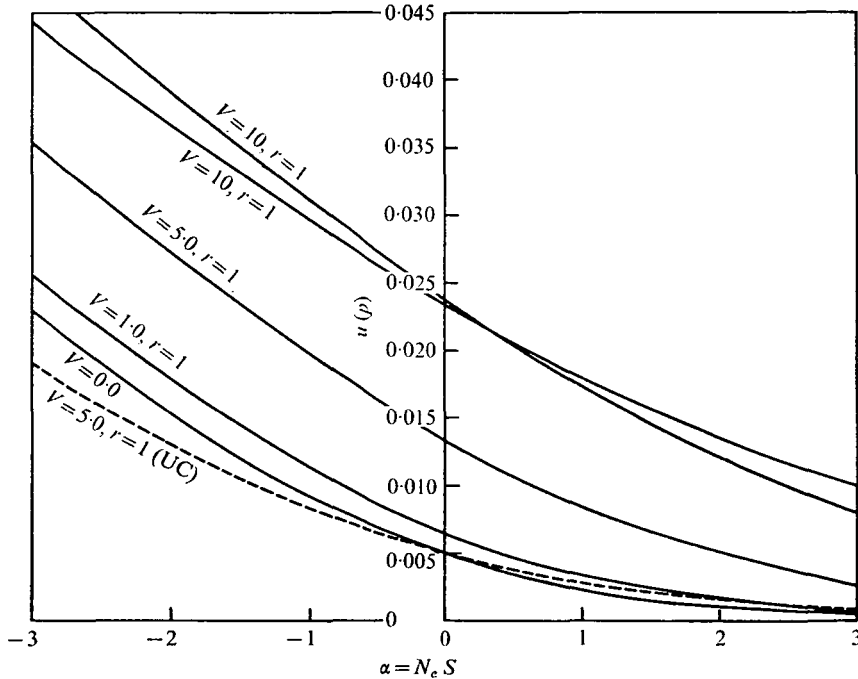


Fig. 5. The probability of ultimate fixation,  $u(p)$ , is plotted against the selection parameter,  $\alpha$ , for model (i), with various values of the variance and the correlation coefficient. The graph, marked 'UC', gives the uncorrected solution.  $V = 2N_e V_\sigma$ ,  $N_e = N = 100$ .

where

$$\begin{aligned} \gamma_i^2 &= \pm \frac{d - \alpha_i}{\alpha_1 - \alpha_2} \quad (i = 1, 2) \quad (d \text{ and } \alpha_i \text{'s are as defined in section 2)} \\ &= \pm \frac{1\frac{1}{2} - 4\alpha/v}{\sqrt{(1+8(1+r)/V)}} - \frac{1}{2} \quad (i = 1, 2) \end{aligned}$$

and for  $r = -1$ ,  $G(x) = (1 + Vx(1-x))^{4\alpha/V-2}$ .  $u(p)$ ,  $\bar{t}_e$  and  $\bar{t}_f$  can now be found using numerical integration. Some values for them are shown in Figs. 5-7. From Fig. 5, one can see that  $u(p)$  is approximately independent of  $r$  by the two graphs for  $V = 10$ . Intermediate values of  $r$  give intermediate graphs. This relative invariance is due to the fact that  $G(x)$  is approximately independent of  $r$  for small  $x$ :

$$G(x) = \left(1 - \frac{\gamma_1^2}{\alpha_1} x(1-x) + \frac{\gamma_2^2}{-\alpha_2} x(1-x)\right) = 1 + 2V \left(\frac{2\alpha}{V} - 1\right) x(1-x) \quad \text{for all } r. \tag{3.6}$$



Also from Fig. 5, one can see that variability increases the chance of fixation of a mutant gene for all values of  $\alpha$ , as one would expect from the heterozygosity curves, and that, for all values of  $V$ ,  $u(p)$  increases as  $\alpha$  decreases, which again is reasonable as decreasing  $\alpha$  produces heterozygote superiority. The mutant initially, will be completely in heterozygous form and thus will benefit more from heterozygote superiority.

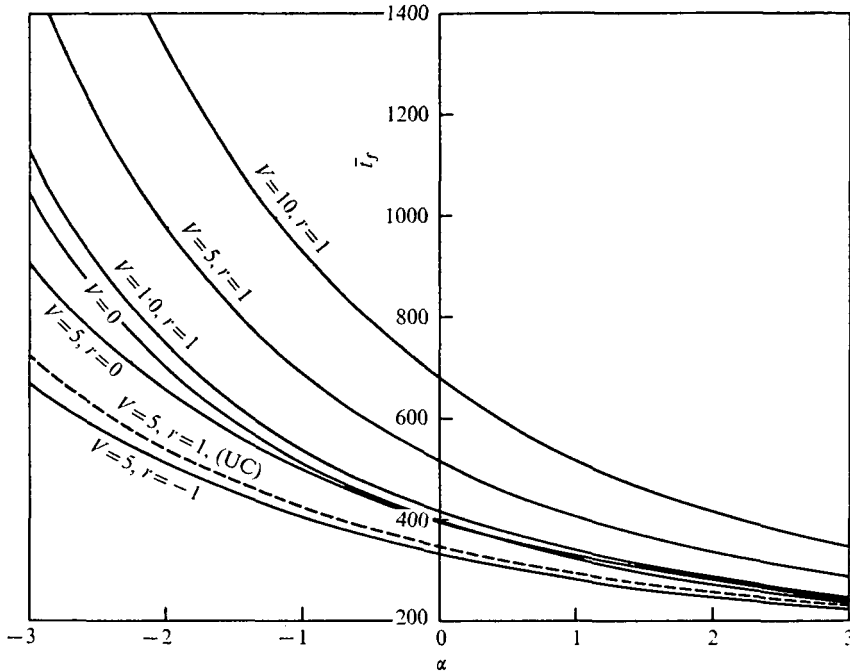


Fig. 6. The time to fixation,  $\bar{t}_f$ , is plotted against the selection parameter,  $\alpha$ , for model (i) for various values of the variance and the correlation coefficient, the dotted graph being the uncorrected solution.  $V = 2N_e V_\sigma$ ,  $\alpha = N_e s$ ,  $N_e = N = 100$ .

From the graphs of  $\bar{t}_f$ , one sees that variance uniformly increases the mean time to fixation, if  $r = 1$ . However, if  $r$  decreases,  $\bar{t}_f$  decreases for all  $\alpha$  until at  $r = -1$ ,  $\bar{t}_f$  is always less than the graph for  $V = 0$  for all the values of  $\alpha$  of interest. This trend is exemplified by the three graphs for  $V = 5.0$ . The graphs for  $\bar{t}_e$  have a similar qualitative form but decreasing  $r$  has a smaller effect, particularly for positive values of  $\alpha$ . For both  $\bar{t}_e$  and  $\bar{t}_f$ , increasing  $\alpha$  causes a decrease for all values of  $V$ .

The effect of ignoring the variance term in  $M_{\delta z}$  is again given by the dotted graphs, marked 'UC', in Figs. 5-7.  $V = 5.0$  and  $r = 1$  are the parameter values used for these graphs.

(ii) Model 2

Genotype	$AA$	$Aa$	$aa$
Fitness	$1 - \sigma$	1	$1 + \sigma$

where the  $\sigma$ 's obey the same properties as defined previously and  $a$  is the mutant

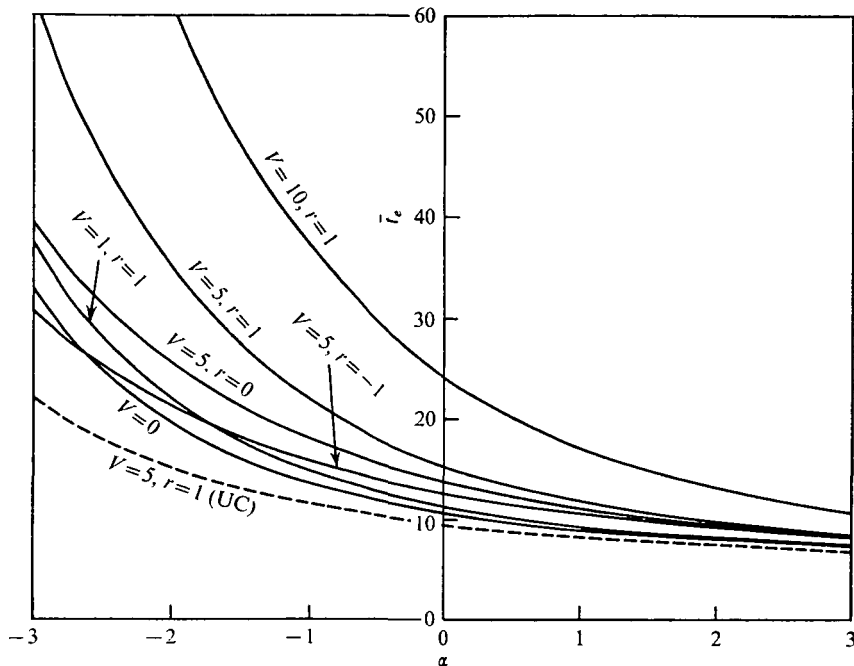


Fig. 7. The time to extinction,  $t_e$ , is plotted against the selection parameter,  $\alpha$ , for model (i) for various values of the variance and the correlation coefficient, the dotted graph being the uncorrected solution.  $V = N_e V_\sigma$ ,  $\alpha = N_e s$ ,  $N_e = N = 100$ .

Table 1. Mean times to extinction in non-random no dominance model ( $R = 1$ )

$s (= E(\sigma))$	$t_{e1}$	$t_{e2}$	$t$				
			$N = 25$	$N = 50$	$N = 75$	$N = 100$	$N = 200$
0.001	13.2748	13.2891	7.9825	9.2981	10.0765	10.628	11.9273
0.005	10.0560	10.1114	7.9523	9.1758	9.8081	10.1723	10.5667
0.01	8.6697	8.7670	7.86097	8.8367	9.1553	9.2237	9.0561
0.02	7.2834	7.4513	7.5347	7.9105	7.8348	7.7408	7.5856
0.03	6.4724	6.7012	7.0994	7.0858	6.9560	6.8857	6.7888
0.05	5.4508	5.7843	6.2229	6.0120	5.9297	5.8913	5.8368
0.1	4.0645	4.6061	4.8392	4.7174	4.6806	4.6627	4.6366
0.2	2.6782	3.5192	3.6505	3.5953	3.5778	3.5692	3.5664

gene. One can obtain expressions for  $M_{\delta x}$  and  $V_{\delta x}$  from (2.1) ( $x$  now being the frequency of  $a$ ) by substituting  $Y = -\sigma$  and  $r = -1$ .

If the  $\sigma$ 's are fixed, then substitution of these values in equation (3.1) yields

$$G(x) = e^{-4ax},$$

$$u(p) = \frac{e^{-4ap} - 1}{e^{-4a} - 1}, \tag{3.7}$$

and

$$\psi(x) = -\frac{(e^{-4x} - 1)e^{4ax}}{sx(1-x)}.$$

$\bar{t}_e$  and  $\bar{t}_f$  are again then found by numerical integration using equations (3.2) and (3.3). The effect of  $\alpha$  on  $u(p)$ ,  $\bar{t}_e$  and  $\bar{t}_f$  for  $V = 0$  is given in Figures 8, 9 and 10. One finds that  $\bar{t}_e$  and  $\bar{t}_f$  are independent of the sign of  $s$  which can easily be proved algebraically using the above formulae for  $u(p)$  and  $\psi(x)$ . Changing  $N_e$  or  $R$  (the original number of mutants) merely shifts the graphs up and down, and does not alter the qualitative trends. In producing these graphs, I have assumed that  $N_e = N$ . If  $N_e = cN$  ( $0 < c < 1$ ), then the results obtained for the mean times to extinction and fixation would be  $c\bar{t}_e$  and  $c\bar{t}_f$  for all fixed values of  $\alpha$ .

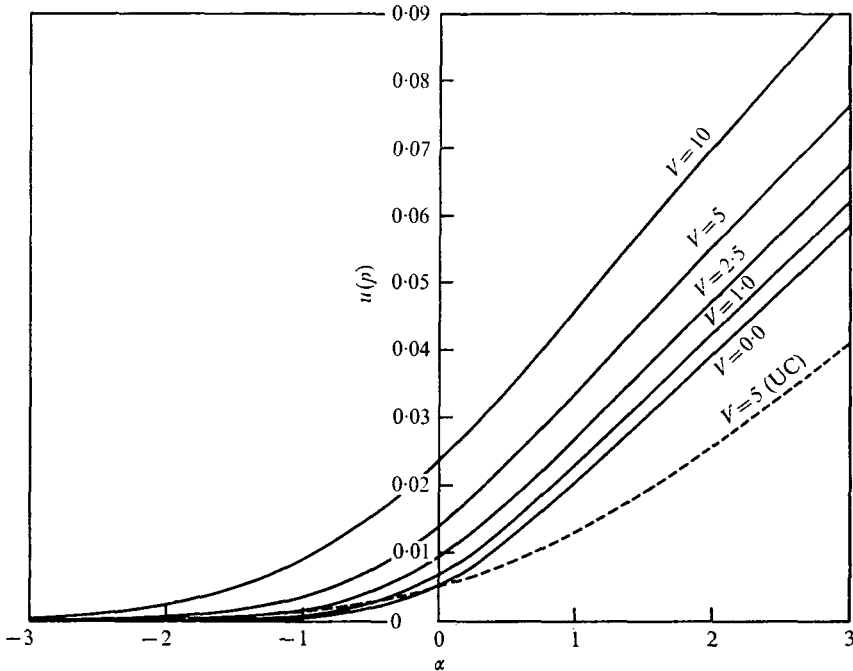


Fig. 8. The probability of ultimate fixation,  $u(p)$ , is plotted against the selection parameter,  $\alpha$ , for model (ii) for various values of the variance. The graph, marked 'UC' gives the uncorrected solution.  $\alpha = N_e s$ ,  $V = 2N_e V_e$ ,  $N_e/N = 1$ ,  $N = 100$ ,  $R = 1$ .

Nei (1971) put forward two approximations for  $\bar{t}_e$  in this model. In deriving both, he assumes that  $\alpha \gg 1$ . From Table 1, one can see that the approximation is reasonable as long as  $\alpha$  is not too small. However, the numerical integrations can be done easily by use of a computer and thus an approximate solution is not necessary. The two approximations of Nei are:

$$\bar{t}_{e1} = \frac{2N_e}{N} [\log_e(N/2N_e s) - \gamma + 1] \quad (R = 1 \text{ only}) \tag{3.8}$$

and

$$\bar{t}_{e2} = \left( e^{4\alpha p} \int_0^{4\alpha p} \frac{1 - e^{-t}}{t} dt - (e^{4\alpha p} - 1) (\log_e(4\alpha p) + \gamma) \right) / s, \tag{3.9}$$

where  $\gamma =$  Euler's constant.

If  $V \neq 0$ , one obtains that,

$$I = \int_0^{x'} \frac{M_{\delta x}}{V_{\delta x}} dx = \int_0^{x'} \frac{2\alpha + V(1-2x)}{Vx(1-x) + 1} dx, \tag{3.10}$$

where  $\alpha = N_e s$  and  $V = 2N_e V_\sigma$ .

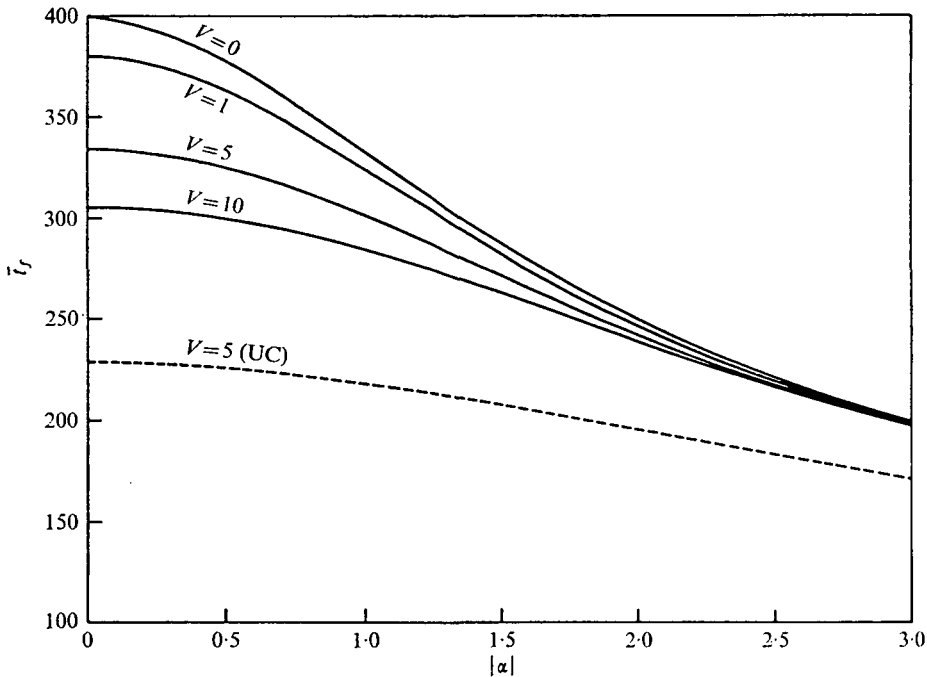


Fig. 9. The time to fixation,  $\bar{t}_f$ , is plotted against the selection parameter,  $\alpha$ , for model (ii), for various values of the variance, 'UC' being the uncorrected solution.  $\alpha = N_e s$ ,  $V = 2N_e V_\sigma$ ,  $N_e = N = 100$ ,  $R = 1$ ,  $\alpha$ .

This integral is again easily evaluated by use of partial fractions and one thus obtains:

$$G(x) = \frac{(1 - x/\alpha_1)^{\lambda_1}}{(1 + x/(-\alpha_2))^{\lambda_2}},$$

where

$$\lambda_i = \frac{4\alpha/V}{\sqrt{(1+4V^{-1})} \mp 2} \quad \text{and} \quad \alpha_i = \frac{1}{2}(1 \pm \sqrt{(1+4V^{-1})}) \quad (i = 1, 2).$$

$u(p)$ ,  $\bar{t}_e$  and  $\bar{t}_f$  are then found from equations (3.1)–(3.3) by use of numerical integration, as before, and some results for them are shown in Figs. 8–10. As in model (i), increasing the variance increases  $u(p)$  for all values of  $\alpha$ . Also in a similar manner to the results of model (i) with  $r = -1$ ,  $\bar{t}_f$  decreases uniformly with  $V$ , and  $\bar{t}_e$  increases uniformly with  $V$ .  $\bar{t}_f$  maintains its independence from the sign of  $s$  for non-zero  $V$ , but  $\bar{t}_e$  shifts its maximum towards negative values of  $\alpha$  as  $V$  increases, the value of  $\alpha$  giving maximum  $\bar{t}_e$  decreasing as  $V$  increases. In both Figs. 9 and 10, one sees that,

as  $\alpha$  becomes large, the effect of variability of  $\sigma$  becomes negligible. As previously in this paper, the dotted curves give the uncorrected results, where the variance has been neglected from  $M_{sz}$  (e.g. Ohta, 1972).

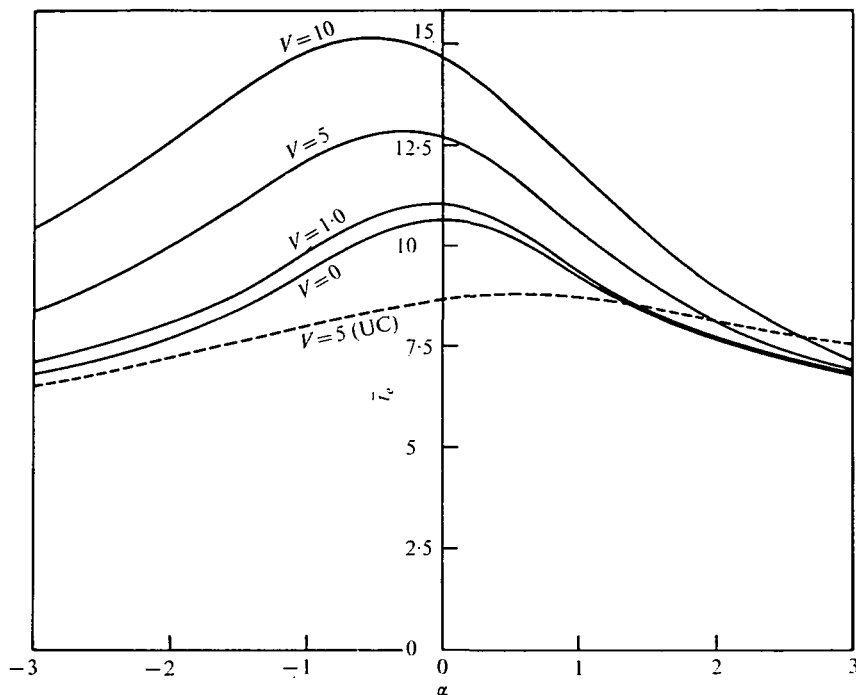


Fig. 10. The time to extinction,  $\bar{t}_e$  is plotted against the selection parameter,  $\alpha$ , for model (ii), for various values of the variance, 'UC' being the uncorrected solution.  $\alpha = N_e s$ ,  $V = 2N_e V_\sigma$ ,  $N_e = N = 100$ ,  $R = 1$ .

#### 4. DISCUSSION

In the previous sections, it has been shown that, under this diploid model, variability of selection coefficients can, generally, markedly increase the genetic variability of the population, expressed either by the steady-state heterozygosity, or by the times to fixation and extinction of a small number of mutant genes. Thus the comment of Karlin & Levikson (1974), that one can easily conclude diploid trends from haploid ones, appears to be untrue, as most of the results contradict the trends advocated by Karlin & Levikson (1974), which are mostly based on haploid results. (Karlin & Levikson use the time to absorption  $T(x)$  which can be expressed thus:  $T(x) = u(x)\bar{t}_f + (1 - u(x))\bar{t}_e$ .) The results of this paper show that increasing the variance of the homozygote selection coefficient decreases the spread of the steady-state gene frequency distribution and increases the steady-state heterozygosity. Also, if  $r$  is not near minus one, it is found that the probabilities of ultimate fixation and mean times to fixation and extinction are increased by increasing variability. The figures, for all the models, also clearly show that the uncorrected solutions

based upon Kimura (1954) give results which are both qualitatively and quantitatively very different.

In this paper the variance of the selection coefficients of the two homozygotes has been kept equal. One could relatively easily change this but it would increase the mathematical complexity. In the asymmetric case, trends are likely to change but not overconsiderably. It seems likely, following Karlin & Levikson (1974), that, for example, if only the fitness of the homozygotic mutant fluctuates, increasing variance will reduce the probability of ultimate fixation. However, in my opinion the assumption of equal variance is not an unreasonable one.

The limitation of the present paper to two alleles is a distinct drawback. However, the more likely situation in which only two alleles are viable, all others being highly deleterious, is likely to give similar qualitative answers. Extensions to include a time autocorrelation, attempted by Takahata *et al.* (1975), would be instructive and desirable also. However, the complexity of the mathematics seems, at present, to be prohibitive.

Proof of the validity of random selection coefficients as a force maintaining genetic variability is as yet relatively small, though claims have been made of its importance in some insect populations (e.g. Fisher & Ford, 1947; Powell, 1971; Smith, 1975). Most species, however, live in environments which are subject to considerable fluctuations, and thus it seems very reasonable that we should take this into account in our calculations.

I would like to thank Dr M. G. Bulmer for helpful comments and encouragement during the work for this paper. Many thanks also to Prof. M. S. Bartlett for his help, while I was in Oxford, to Dr W. G. Hill and Prof. A. Robertson for their many useful comments upon this paper, and to the referees for their very helpful advice.

#### APPENDIX

##### (i) Derivation of equation (2.3)

On substitution of the expressions for  $M_{\delta x}$  and  $V_{\delta x}$  (2.1) in the expression  $I = \int M_{\delta x}/V_{\delta x} dx$ , one obtains:

$$I = \int \frac{(Vx(1-x)(1-(1+r)x(1-x)) + \beta/2)(1-2x) - 2x(1-x)(\delta_2 - (\delta_1 + \delta_2)x)}{Vx^2(1-x)^2(1-2(1+r)x(1-x)) + x(1-x)} dx \quad (\text{A } 1)$$

where  $V = 2N_eV_s$ ,  $\delta_1 = N_e s$ ,  $\delta_2 = N_e t$  and  $\beta = 4N_e v$ . Taking the case  $\delta_1 = \delta_2 = \alpha$ , and substituting  $y = x(1-x)$ , (A 1) simplifies to

$$I = \int \frac{Vy(1-(1+r)y) - 2\alpha y + \beta/2}{Vy^2(1-2(1+r)y) + y} dy. \quad (\text{A } 2)$$

The denominator can now be easily factorized, if  $r \neq -1$ , to

$$2V(1+r)y(\alpha_1 - y)(y - \alpha_2) \quad (\text{A } 3)$$

where

$$\alpha_i = \frac{1}{4(1+r)} \left[ 1 \pm \sqrt{\left(1 + \frac{8(1+r)}{V}\right)} \right] \quad (i = 1, 2).$$

As  $0 < y \leq \frac{1}{4}$ ,  $\alpha_1 > \frac{1}{4}$  and  $\alpha_2 < 0$  for all values of  $r \in (-1, 1]$ , for  $V \neq 0$ , the denominator is always non zero. The integrand of (A 2) can now be easily broken up into partial fractions, all of which can be integrated easily. Thus

$$I = \frac{1}{2} \left[ - \left( \frac{d-\alpha_1}{\alpha_1-\alpha_2} + \frac{\beta}{V\alpha_1+2} \right) \log(\alpha_1-y) + \left( \frac{d-\alpha_2}{\alpha_1-\alpha_2} - \frac{\beta}{V\alpha_2+2} \right) \log(y-\alpha_2) + \beta \log_e y \right] \quad (\text{A } 4)$$

which on substitution in equation (2.2), yields (2.3).

(ii) Evaluation of constant, C, in equation (2.3)

C can be evaluated by using the fact that

$$\int_0^1 \phi(x) dx = 1 \quad (\text{by the definition of density functions}).$$

Let  $\bar{\phi}(x) = \phi(x)/C$ . Then

$$C^{-1} = \int_0^1 \bar{\phi}(x) dx.$$

$\bar{\phi}(x)$ , however, tends to infinity at  $x = 0$  and 1, and thus numerical integrals are subject to great inaccuracies due to rounding error. If, however, one substitutes  $Z = |2x - 1|$  and integrates by part one finds that

$$\begin{aligned} C^{-1} &= 4^{1-\beta} \int_0^1 f(Z) (1-Z^2)^{\beta-1} dZ \\ &= \frac{4^{1-\beta}}{\beta} \left[ f(0) - \int_0^1 (1-Z)^\beta \frac{\partial}{\partial Z} [f(Z) (1+Z)^{\beta-1}] dZ \right], \end{aligned} \quad (\text{A } 5)$$

where

$$\begin{aligned} f(Z) &= (\alpha_1 - (1-Z^2)/4)^{\gamma_1} ((1-Z^2)/4 - \alpha_2)^{\gamma_2} \quad \text{if } r \neq -1. \\ &= ((1-Z^2)/4 + V^{-1})^{1-\beta-4\alpha/V} \quad \text{if } r = -1. \end{aligned}$$

The integrand in equation (A 5) is finite at the end points (i.e.  $Z = 0, 1$ ) and thus presents no problem when integrated numerically.

REFERENCES

FISHER, R. A. & FORD, E. B. (1947). The spread of a gene in natural conditions in a colony of the moth *Panmixia dominula* L. *Heredity* **1**, 143-174.  
 GILLESPIE, J. H. (1973). Natural selection with varying selection coefficients - a haploid model. *Genetical Research* **21**, 115-120.  
 HARTL, D. L. & COOK, R. D. (1973). Balanced polymorphisms of quasi-neutral alleles. *Theoretical Population Biology* **4**, 163-172.

- HARTL, D. L. & COOK, R. D. (1974). Stochastic selection in large and small populations. *Theoretical Population Biology* **7**, 55–64.
- HARTL, D. L. & COOK, R. D. (1975). Stochastic selection and the maintenance of genetic variation. In *Population Genetics and Ecology* (ed. S. Karlin and E. Nevo). Academic Press, London and New York.
- JENSEN, L. (1973). Random selection advantages of genes and their probability of fixation. *Genetical Research* **21**, 215–219.
- KARLIN, S. & LEVIKSON, B. (1974). Temporal fluctuations in selection intensities: case of small population size. *Theoretical Population Biology* **6**, 383–412.
- KIMURA, M. (1954). Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics* **39**, 280–295.
- KIMURA, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* **47**, 713–719.
- KIMURA, M. (1964). Diffusion models in Population Genetics. *Journal of Applied Probability* **1**, 177–232.
- KIMURA, M. & OHTA, T. (1969*a*). The average number of generations until fixation of a mutant gene in a finite population. *Genetics* **61**, 763–771.
- KIMURA, M. & OHTA, T. (1969*b*). The average number of generations until extinction of an individual mutant gene in a finite population. *Genetics* **63**, 701–709.
- NEI, M. (1971). Extinction times of deleterious mutant genes in large populations. *Theoretical Population Biology* **2**, 419–425.
- OHTA, T. (1972). Fixation probability of a mutant influenced by random fluctuation of selection intensity. *Genetical Research* **19**, 33–38.
- OHTA, T. & KIMURA, M. (1972). Fixation time of overdominant alleles influenced by random fluctuation of selection intensity. *Genetical Research* **20**, 1–7.
- POWELL, J. R. (1971). Genetic polymorphisms in varied environments. *Science* **174**, 1035–1036.
- SMITH, D. A. S. (1975). Sexual selection in a wild population of the butterfly *Danaus chrysippus* L. *Science* **187**, 664–665.
- TAKAHATA, N., ISHII, K. & MATSUDA, H. (1975). Effect of temporal fluctuation of selection coefficient on gene frequency in a population. *Proceedings of the National Academy of Sciences* **72**, 4541–4545.
- WRIGHT, S. (1938). The distribution of gene frequencies under irreversible mutation. *Proceedings of the National Academy of Sciences* **24**, 253–259.
- WRIGHT, S. (1949). Adaptation and selection. In *Genetics, Palaeontology and Evolution* (ed. Jepson *et al.*). Princeton University Press, Princeton, New Jersey, U.S.A.