Theoretical Population Genetics of Spatially Structured Populations

Ian J. Lundy

27 October 1997

Thesis submitted for the degree of

Doctor of philosophy

in

The University of Adelaide
(Department of Applied Mathematics)

Errata

p18, l 15: using a differential equation

p34: Second last sentence should read "The transition probabilities for subpopulation 1 to have k A_1 alleles at time $t+1$ given the subpopulations contained X_t and Y_t individuals respectively at time t can be derived by making the following observation."

p35, l 7: l not k on RHS in equation (3.2)

p39, l 15: qualitative

p67, l 15: mean

p70, l -3: coefficients

p75, l -0: μ_{32}

p77, l -9: although

p96, l -3: $E[X_{l,t+1}|x_t]$

p130, l 9: allele

p138: "mean absorption time" should be "mean fixation time"

p144, l 5: k in summation in formula for $T_2(j)$

Some further helpful references

Herbots, H. M. (1997). The Structured Coalescent. In Donnelly, P. and Tavaré, S., editors, *Progress in Population Genetics and Human Evolution*, IMA volumes in Mathematics and its Applications, **87**, pages 231 – 255, Springer-Verlag, Berlin.

Sved, J. A. and Latter, B. D. H. (1977). Migration and mutation in stochastic models of gene frequency change. I. The island model. *Journal of Mathematical Biology*, **5**, pages 61 – 73.

Latter, B. D. H. and Sved, J. A. (1981). Migration and mutation in stochastic models of gene frequency change. II. Stochastic migration with a finite number of islands. *Journal of Mathematical Biology*, **13**, pages 95 – 104.

Contents

5 Models for three and four subpopulations 66

6 A general model for fixation probabilities 92

Declaration

I declare that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due refereence has been made in the text.

I also give consent to this thesis being available for photocopying and loan if accepted for the award of the degree.

Ian Lundy
27 October 1997

Acknowledgements

There are many people who have helped during the preparation of this thesis and I would like to thank all of them. I would also specifically like to thank the following people.

First, my supervisor, Hugh Possingham has provided me with excellent supervision and advice during my candidature.

I would also like to thank all of my fellow students especially Jemery Day and Geoff Tuck for there help and support during my candidature.

Christine O'Keefe from the Department of Pure Mathematics provided some valuable guidance on the geometrical theory used in this thesis. Peter Taylor of the Teletraffic Research Centre provided some valuable guidance on the some aspects of Markov chains.

Summary of thesis

In this thesis I consider the question of fixation probabilities and mean absorption times for alleles when a population is divided into a number of subpopulations with asymmetric migration between the subpopulations. The emphasis of the thesis is on small populations and conservation genetics.

Initially I consider a population which is subdivided into two subpopulations. Non-symmetric migration is allowed between the two populations and there is no selection. I use a discrete time, discrete state space, probabilistic model (Wright-Fisher model) to find and prove a general formula for the fixation probability in this case. I then use the same model to consider the question of mean absorption and fixation times. I compare the results I obtain with the results of other authors and with the low migration limit approximation of Slatkin (1981).

I extend this model to the case of three and four populations and extend the results obtained for the two subpopulation case using the same methods as for the two subpopulation case.

I find a generalisation of the fixation probability result for a general number, n, of subpopulations. This result is proved using results from graph theory as well as the probability theory which was used in the two subpopulation case. The result is then interpreted using graph theory.

Following this I consider a single population with selection acting and find general matrix formulae which the fixation probability and the mean absorption time must satisfy. These equations are solved analytically for a specific case.

I then combine the selection of this model with the population structure of the two subpopulation case. Again I obtain general matrix formulae which the fixation probability and the mean absorption time must satisfy. These are solved numerically and compared with the results of other authors and the single population with selection case.

Finally I consider a population which is not of constant size. Selection is included in this model. Initially I consider a population which alternates between two different population sizes and find general matrix formulae which the fixation probability and the mean absorption time must satisfy. These are solved numerically for some simple cases and compared with a population of constant size N_e, the effective population size proposed by Wright (1931). These results are also compared with a diffusion approximation of this model.

I then consider populations which have a longer cycle and find general matrix formulae for the fixation probability and mean absorption time. I also consider a population which varies randomly according to some unspecified population dynamic process. General matrix formulae are found for the fixation probability.

Chapter 1

Literature Review

In this thesis I consider theoretical population genetical models that potentially have use in the management and conservation of genetic variability. Specifically, I will use population genetic models to consider how changes in population structure due to natural causes or changed land uses and habitat fragmentation affect the conservation of genetic variation in wild populations. The quantities considered here as measures of the extent and rate of loss of genetic diversity are the fixation probability, and the mean fixation and mean absorption times. In this chapter I review the literature and discuss models by other authors which are of relevance to the to the problems being considered in the rest of this thesis.

1.1 Background to relevant models

One of the main reasons that population genetic models have been constructed historically is to model changes in the genetic structure of a population with the purpose of understanding evolution. Although the basic processes of evolution have been understood for some time (Darwin, 1859), the details of how evolution has acted to produce the diversity of life and adaptations which exists today are still a matter of research. One class of population genetic models which are used to study evolution assume that the population is large and the question of interest is "How likely is it that a mutant individual in the population will achieve fixation and how long does this process take?" In other words, these models are intended to find the *rate* of evolution under various sets of conditions. The results from these studies can then be used to compare the feasibility of different theories of evolution by finding which theories can explain our observations of the rates of evolution. See for example Wright (1931; 1969), Moran (1962) and Barton and Whitlock (1997).

Other models which have been used to study the process of evolution include game theory models, see for example Maynard-Smith (1982) and Dawkins (1976; 1982). This type of model provides little information about the rate of evolution so will not be examined in this thesis.They have been especially useful in the study of behavioural evolution and life history "strategies", and have led to the concept of Evolutionary Stable Strategies.

Genetic drift of neutral alleles is a process which has been shown to be of less relative importance in evolution than selection of alleles which provide a benefit to the organism involved. With the exception of small populations, the times to fixation of neutral genes will be large and thus cannot account for more than a small amount of change in the genetic composition of a population. In very large populations drift has almost no effect as it is unlikely that a gene will randomly become fixed in a very large population.

In conservation genetics the populations being considered are usually small remnant populations left as a result of human land management techniques. Genetic drift could play a large part in the loss of heterogeneity in these circumstances. Kimura (1983) wrote

> Furthermore, we should not overlook the possibility that some of the neutral alleles may become advantageous under an appropriate environmental condition or a different genetic background; thus, neutral mutants have a latent potential for selection ... To regard random fixation of neutral mutants as "evolutionary noise" is inappropriate and misleading.

This means that it is important to consider what will happen to neutral genes as heterozygosity provides at least some of the variation which selection can act on when circumstances change and a neutral gene confers some selective advantage. Maintenance of heterozygosity is especially important in the case of small populations and in the case of populations which are fragmented into small subpopulations as this is where genetic diversity is at greatest risk from genetic drift.

The loss and fragmentation of large quantities of habitat in parts of the world has meant that much of the remaining natural vegetation in some areas is in small, subdivided areas and hence habitat for many species is now in small subdivided blocks. Conservation agencies have become increasingly concerned with the maintenance of genetic variability in populations under these circumstances (for example, Sampson *et al.*, 1988; Falk, 1990; Hickey *et al.*, 1991; Leary *et al.*, 1993; Campbell *et al.*, 1994), so as to avoid further loss of genetic variability which could be detrimental to the long term survival of the species.

This thesis is concerned with the genetics of these small and subdivided populations as a conservation issue. The work presented here also applies to evolutionary genetic modelling in the sense that maintenance of heterozygosity is an important factor in evolution (Kimura, 1983).

In this review of the literature I will examine the work of other authors which is relevant to the questions which I will consider in this thesis. Each of the models of other authors that I consider in this chapter is a single locus model. I will only consider the haploid version of each model. This means that diploid models with additive selection are implicitly considered but diploid models with any form of dominance are not considered. There is no mutation in any of the models used in this thesis as mutation is not generally important in small populations. Note that as the population is finite and there is no mutation, all alleles must eventually be either fixed in the population or lost from the population.

In each of the models considered there are three quantities of interest:

- The probability of ultimate fixation of an allele, here shortened to the *fixation probability* of an allele. This is used to examine the effect of structure on loss of heterozygosity.
- Mean absorption time of an allele. The mean time until the allele is either fixed in the population or lost from the population. This is a measure of the rate of loss of heterozygosity.
- Mean fixation time of an allele. The mean time to fixation in those cases where the allele is ultimately fixed in the population. This is a measure of the rate at which evolution procedes.

The aim of this chapter is to provide background information on what work has been done by other authors so that comparisons may be made between their results and the new results in this thesis.

Each of the relevant models which has been studied by previous authors is described and results which have been derived are presented. For each model the assumptions will be discussed and results given, first for the case of no selection and then for the case of selection.

1.1.1 Notation

For each of the models of other authors which I consider in this chapter, the population is of size $2N$. There are two alleles A_1 and A_2 with x representing the number of A_1 alleles and x_0 representing the initial number of A_1 alleles. Hence there are initially $2N - x_0$ A_2 alleles. The selection coefficient is s. The interpretation of s is different for each model and will be explained in the description of each model. The variable $u(x_0)$ represents the fixation probability, $T(x_0)$ represents the mean absorption time, and $T^*(x_0)$ represents the mean fixation time for an allele initially present in x_0 individuals.

1.2 Wright-Fisher models

The simplest models in population genetics are the models first used by Fisher (1922) and Wright (1931). The models assume that time is discrete and that the state space is discrete. Each time point is a generation, so generations are non-overlapping. The population contains $2N$ haploid individuals which is equaivalent to a diploid population of N individuals when there is no selection or when selection is additive. The models are stochastic and consider a single locus with two alleles A_1 and A_2. A new generation is produced by assuming that at each generation there are a very large number of juveniles born in proportion to the adults in that generation. The breeding stock for the next generation are then randomly sampled from this large pool of juveniles, giving a binomial distribution for the genotypes of the new generation. Immediately before time $t + 1$ the population has i genes of type A_1. At time $t + 1$ the existing population dies and a new generation forms where each individual is drawn randomly from an infinitely large pool of juveniles with proportion $i/2N$ of gene A_1, giving a binomial distribution for the adults in the new generation. This means that the transition probabilities are

$$p_{ij} = P(X_{t+1} = j | X_t = i) = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}$$

where X_t is the number of A_1 alleles in the population at time t.

Much of the early work in the field of population genetics used models of this type as well as various continuous approximations to obtain results on fixation probabilities, rates of loss of heterogeneity and probability distributions of gene frequency.

Continuous time and state space approximations to this model are known as diffusion models and will be discussed in section 1.4. Although the full Wright-Fisher model is simple to describe, it is generally easier to find results using other types of model, such as diffusion models, especially when the assumptions are varied to allow for further issues such as selection, dominance, non random mating or population subdivision.

Moran (1960) showed using an exact Wright-Fisher model that the fixation probability of a gene with selection acting according to the selection parameter s is approximately:

$$u(x_0) = \frac{1 - e^{-4N\theta x_0}}{1 - e^{-4\theta N}}$$

for some θ, where $s/(1+s) \leq \theta \leq s$. If s is small then $s \approx \theta$ and then an adequate approximation for the fixation probability is

$$u(x_0) = \frac{1 - e^{-4sx_0}}{1 - e^{-4sN}}.$$

The interpretation of s in the Wright-Fisher model is that allele A_1 has fitness $1+s$ and allele A_2 has fitness 1. This can be thought of as a very large (effectively infinite) number of alleles being born into a pool of juveniles in the proportion $(1+s)i$ to $N-i$. These juveniles are then sampled binomially to make the N adults in the next breeding generation.

In the case where there is no selection the fixation probability for allele A_1 when there are initially x_0 A_1 alleles present is

$$u(x_0) = x_0/2N.$$

1.2.1 Mean absorption and fixation times

For the Wright-Fisher model it is possible to accurately find bounds which the mean fixation time for a neutral mutant must lie within. Gale (1990, page 236) gives the upper bound for the mean fixation time as $4N - 2$ generations and the lower bound for the mean fixation time as $4N - \frac{1}{3}\log(2N-1) - 2.0551$. It is not possible to find a general formula for the fixation time using the Wright-Fisher model when there is selection. In general, analytic or even approximate formulas for mean fixation or mean absorption times are difficult to find for this model.

1.3 Moran models

Moran (1958) proposed an alternative model which could also be used to consider the same questions of fixation probability and mean fixation time. It also assumes that time is discrete and that the state space is discrete. Unlike the Wright-Fisher model, the Moran model assumes overlapping generations and is formulated as follows. At each time point an individual dies and is replaced by another individual. The remaining $N-1$ individuals are unaffected at this time point. The number of A_1 individuals in the population immediately before time $t+1$ is i. An individual dies at time $t+1$ and is replaced by an A_1 individual with probability $i/2N$ and an A_2 individual with probability $1-i/2N$. So the transition probabities (p_{ij} = probability of transition from state i to state j) are

$$\begin{aligned} p_{i,i+1} &= \frac{i}{2N}\left(1-\frac{i}{2N}\right) \\ p_{i,i} &= \left(\frac{i}{2N}\right)^2+\left(1-\frac{i}{2N}\right)^2 \\ p_{1,i-1} &= \frac{i}{2N}\left(1-\frac{i}{2N}\right) \end{aligned} \tag{1.1}$$

in the case where there is no selection. When there is selection acting the transition probabiities are

$$\begin{aligned} p_{i,i+1} &= \frac{(1+s)i(2N-i)}{(2N+si)2N} \\ p_{i,i-1} &= \frac{i(2N-i)}{(2N+si)2N} \\ p_{i,i} &= 1-p_{i,i+1}-p_{i,i-1} \end{aligned} \tag{1.2}$$

$$\tag{1.3}$$

and the selection coefficient can be defined as

$$1+s=\frac{p_{i,i+1}}{p_{i,i-1}}$$

so as with the Wright-Fisher model, $s=0$ means that there is no selection. Note that the interpretation of s here is essentially the same as the interpretation in the Wright-Fisher case although the mechanism is slightly different. The interpretation of s is that the likelihood of an A_1 allele occurring compared to an A_2 allele in a given generation is about $1+s$ times as great as would be expected if there was no selection.

It is also possible to approximate the Moran model by a continuous time, continuous state space model to obtain a diffusion model. This will be discussed in section 1.4

The first quantity of interest is the fixation probability of a gene with initial frequency $x_0/2N$ in the population. Moran (1962) showed that the fixation probability for a gene A_1 when it has a selective advantage over A_2 is

$$u(x_0)=\frac{1-(1+s)^{-2x_0}}{1-(1+s)^{-2N}} \tag{1.4}$$

and when s is small

$$u(x_0) \approx \frac{1 - e^{-2sx_0}}{1 - e^{-2sN}}.$$

When there is no selection

$$u(x_0) = x_0/2N.$$

That is the fixation probability of the A_1 allele is equal to the initial frequency of the A_1 allele in the population when there is no selection. This result is the same as was found for the Wright-Fisher model.

1.3.1 Mean absorption and fixation times

The mean absorption and fixation times for the Moran model can be found by summing the mean time spent in each of the states other than the absorbing states. The result in the case where there is no selection is due to Watterson (1961) who showed that the mean absorption time for an allele with initial frequency $x_0/2N$ is

$$T(x_0) = \sum_{i=1}^{x_0} \frac{2N - x_0}{2N - i} + \sum_{x_0+1}^{2N-1} \frac{x_0}{i}$$

and when $x_0 = 1$, $T(1) = \log(2N - 1) + 0.5772$.

The mean fixation time is found by considering a modified process where only those paths that result in fixation are considered. The mean absorption time for this modified process is the mean fixation time for the original process. The mean fixation time for the original process then is

$$T^*(x_0) = \sum_{i=1}^{x_0} \frac{(2N - x_0)i}{(2N - i)x_0} + 2N - 1 - x_0$$

or in the case when $x_0 = 1$, $T^*(1) = 2N - 1$.

The result in the case where there is selection is due to Ewens (1963). He showed that when the selection parameter is s, the mean time to absorption is

$$T^*(x_0) = \frac{s^{2N-x_0} - 1}{s^{2N} - 1} \sum_{i=1}^{x_0} \frac{s^i - 1}{p_{i,i-1}(s - 1)} + \frac{s^{2N} - s^{2N-x_0}}{s^{2N} - 1} \sum_{i=x_0+1}^{2N-1} \frac{s^{2N} - s^i}{p_{i,i+1}(s^{2N} - s^{2N-1})}.$$

1.4 Diffusion methods

It is often difficult to obtain analytic results using discrete time, discrete state space models except when populations are small. A simplification which can be used is to approximate a model which operates in discrete time and has a discrete state space by assuming that both gene frequency and time are continuous variables. The differential equations which arise from this type of approximation are of the same form as the differential equations for the diffusion of heat in a solid, so this type of genetic model is known as a diffusion model. By assuming the population is large it is reasonable to assume the gene frequency is a continuous random variable and by scaling the generation time it is also reasonable to use a continuous time approximation. Fisher (1922) was the first author to use a diffusion model. The model he used was a continuous time, continuous state space approximation to the Wright-Fisher model. Diffusion models can also be derived from Moran models.

Under the assumption that the population is large and with the generation time appropriately scaled, it is possible to use a diffusion approximation to the underlying genetic process. It is generally much easier to obtain results using a diffusion model than it is to obtain results with the original Wright-Fisher or Moran model. In particular, when dominance is introduced into the diploid model it is far easier to use the diffusion model approach.

The main problem with diffusion process models is that they are an approximation based on large population size. The results obtained do not necessarily apply when populations are small which is the particular problem being investigated in this thesis. For this reason I do not use diffusion models in this thesis except as comparisons to the models which I develop.

The probability of fixation using a diffusion model was first calculated by Kimura (1962) where allele A_1 has relative fitness $1+s$ and allele A_2 has relative fitness 1. As was the case with the Wright-Fisher models, the probability of fixation of allele A_1 with initial frequency x_0 is

$$u(x_0) = \frac{1-e^{-4sx_0}}{1-e^{-4sN}}$$

and again when there is no selection this is equal to

$$u(x_0) = x_0/2N.$$

Note that the results obtained here are exactly the same as the approximate solutions to the Wright-Fisher model given in section 1.2

1.4.1 Mean absorption and fixation times

The mean absorption time for an allele can be found by calculating the probability density of the mean time spent in each state and then integrating this over the appropriate range to get the total expected time spent in states other than the absorbing states. Similarly, the mean fixation time can be calculated by considering a modified process where only sample paths that finish with fixation are considered. The absorption time for this modified process is the fixation time for the original process.

Watterson (1961) showed that the mean absorption time for an allele with initial number of individuals x_0 (that is, initial frequency $x_0/2N$) can be calculated explicitly using the diffusion model with no selection and is

$$T(x_0) = -4N_e[\frac{x_0}{2N}\log\frac{x_0}{2N} + (1-\frac{x_0}{2N})\log(1-\frac{x_0}{2N})].$$

Kimura and Ohta (1969b) showed that the mean time to fixation for an allele with initial number of individuals x_0 when there is no selection is

$$T^*(x_0) = -4N_e\left(\frac{(1-\frac{x_0}{2N})\log(1-\frac{x_0}{2N})}{2N-1}\right).$$

When selection is involved it is impossible to find the mean fixation and mean absorption times analytically and they must be found numerically. The method involves finding the expected sojourn time for all values of the initial frequency between 0 and 1. This function is then integrated over $(0, 1)$ to find the mean time to absorption or fixation. Watterson (1961) derives a formula involving integrals for the mean absorption time of a process with no selection. As an example Gale (1990, page 265) shows that in the case when the initial frequency is $1/2N$ – the case of most interest from the evolutionary point of view – and $N_e s > 20$ then the mean fixation time is

$$T^*(1/2N) = \frac{3.9720 + 2\log N_e + 2\log s}{s} - \frac{2N_e}{N}.$$

1.5 Branching processes

Branching processes were first used in population genetics by Fisher (1922) to determine the fixation probability of a gene which is initially rare in a large population. They have not been used extensively in the study of population genetics, although Pollak (1966) derived results for a subdivided population using branching processes. These models differ from Wright-Fisher models in the way that a generation is produced from the preceding generation. The Wright-Fisher model assumes that A_1 individuals are produced binomially in proportion to the number of individuals in the previous generation – effectively assuming that each individual gives birth to a large number of juveniles and that the new generation is sampled from these individuals. The branching process model on the other hand assumes that each individual from the existing generation gives rise to a relatively small number of indivudals according to some distribution and the new generation is made up combining all of these juveniles. Generally the population size is not assumed to be constant although it can be assumed that the population size will on average remain constant by assuming that each individual produces on average one juvenile. As with the Wright-Fisher model, branching process models assume that time moves in discrete jumps and that the state space is also discrete. Generations are non-overlapping. The approximation is best when the population is large. The initial population size is N. In a branching process model it is assumed that each individual in the population is haploid, is individually propagating, that each indiviual has $0, 1, 2 \ldots$ offspring with probability $p_0, p_1, p_2, \ldots$ and that $\sum_i p_i = 1$. This means that the population size is not constant and can tend towards infinity in the long run.

To find fixation probabilities using this model, consider the probability generating function

$$f(x) = \sum_{i=0}^{\infty} p_i x^i.$$

If there are r individuals in the population in the first generation then the probability of having s individuals in the second generation is the coefficent of the dummy variable x^s in $(f(x))^r$.

If there was one A_1 individual in the first generation then the probability that there are s A_1 individuals in the third generation is the coefficient of x^s in $f(f(x))$. The equilibrium conditions for this process are that $f(x) = x$ and if we denote the solution for this equation as l then

$$T^*(x_0) = 1 - l^{x_0}.$$

If the population is not assumed to be effectively infinite then it can be shown (for example, Pollak, 1966) that the fixation probability is

$$T^*(x_0) = \frac{1 - l^{x_0}}{1 - l^N}$$

which is essentially the same as results derived for other models discussed previously, that is,

$$T^*(x_0) = \frac{1 - e^{-2sx_0}}{1 - e^{-2sN}}$$

as e^{-2s} is a good approximation for the solution of $x = f(x) = e^{(1+s)(x-1)}$ when s is small.

1.6 Effective population size

The concept of *effective population size*, usually denoted by N_e, was introduced by Wright (1931). It is a useful means of dealing with populations which do not fit all of the assumptions made in the previous models with regard to sex ratios, constancy of the population size, randomness of mate choice or heritability of fecundity. By considering a population with N_m males and N_f females, where N_m and N_f are not necessarily equal, Wright showed that this population lost heterozygosity at a rate equivalent to a population with equal numbers of each sex and total population size

$$N_e = \frac{4N_m N_f}{N_m + N_f}.$$

The population with N_m males and N_f females has effective population size N_e.

This means that when $N_m = N_f = \frac{1}{2}N$ the effective population size is N – the same as the actual population size. When N_m is small and N_f is very large, for example when a dominant male breeds with a herd of females, the effective population size is $4N_m$. So the population will lose heterozygosity at the same rate as a population with $2N_m$ males and $2N_m$ females breeding randomly.

There are several other methods for calculating an effective population size which are not all equivalent. The above formula can also be obtained by equating the variance of the population under consideration to the variance of a random mating equal sex ratio population of size N_e. Hence it is known as the *variance effective population size.* Another similar measure is the *inbreeding effective population size* which can be useful when it is assumed that there is a large amount of inbreeding in the population which may be reducing the genetic variability of the population. This is discussed by Wright (1969) and Ewens (1982) but is not considered further here. Other measures of effective population size are also available – see, for example, Nei and Murata (1966).

1.7 Population subdivision

In all of the models discussed so far it has been assumed that the population has occurred as one single homogeneous population. As one of the major areas of investigation in this thesis is the genetic effect of subdividing a population into small subpopulations, I will now examine what other authors have written about subdivided populations. Population subdivision has been looked at for three of the four types of model described above. It is generally easier to find an approximate expression for the fixation probability of an allele than it is to find an approximate expression for the mean absorption or mean fixation time of an allele. Upper and lower bounds have been found for the mean fixation time in a subdivided population by several authors, these results will now be discussed.

1.7.1 Moran model

A Moran model for a subdivided population was first considered by Maruyama (1974). In addition to the assumptions for the Moran model he assumes that a population consists of subpopulations between which there is migration. "Between the birth-death events, migration between colonies may occur. There is no restriction on these processes. The gene frequencies in the whole population are not changed by this, but there will be local changes." (Maruyama, 1977, page 159). He finds that for the Moran model, the probability of fixation is

$$u(x_0) = \frac{1-(1+s)^{-2x_0}}{1-(1+s)^{-2N}} \approx \frac{1-e^{-sx_0}}{1-e^{-2sN}}$$

where the total population size is $2N$. This is the same as the fixation probability for a single population. However, there are several implied assumptions in Maruyama's model. Migration between colonies is never quantified by Maruyama and it is not made clear how or when migration takes place. The way in which the model is set up essentially means that Maruyama is analysing the case where there is no long run net migration between the subpopulations (this includes symmetric migration). Under Maruyama's assumptions, migration does not change gene frequencies overall so effectively he is considering a population structure where adults are allocated to a subpopulation in a particular generation and their offspring could be allocated to any subpopulation in the next generation as there is no restriction on the migration process. It is effectively a single population with a structure in the adult phase which does not affect how genes are passed from one generation to the next.

1.7.2 Branching processes

Pollak (1966) looked at subdivided populations using a Branching process model. There are two approaches for calculating the fixation probability. The first assumes that there is no limit on the population size. Define π_i = Prob(the line descended from a gene of type A_1 becomes extinct | the ancestral gene of the line is in subpopulation i). This can be found by numerically solving the system of equations

$$\pi_i = e^{\alpha_i \sum_{j=1}^{K} \mu_{ij}(\pi_j - 1)} \tag{1.5}$$

where K is the number of subpopulations, α_i is the expected number of offspring surviving to adulthood and μ_{ij} = Prob(offspring of gene in subpopulation i is in subpopulation j), that is, the migration probability. The fixation probability is then given by the formula

$$\text{P(gene survives|there are initially } z_{j,0}\ A_1 \text{ genes in subpopulation } j) = 1-\pi_1^{z_{1,0}} \ldots \pi_K^{z_{K,0}}.$$

The second approach for calculating the fixation probability using branching processes is also due to Pollak (1966). In this approach the population is assumed to be large and constant with the N_i individuals in subpopulation i. He defines α_i to be equal to

$$\alpha_i = \frac{\text{the expected number of offspring of an } A_1 \text{ gene in subpopulation } i}{\text{the expected number of offspring of an } A_2 \text{ gene in subpopulation } i}$$

then the fixation probabilty is given by the expression

$$T^*(z_1, \ldots, z_N) = \frac{1 - \pi_1^{z_1} \ldots \pi_K^{z_K}}{1 - \pi_1^{N_1} \ldots \pi_K^{N_K}},$$

where π_i is defined as in equation 1.5. Pollak points out that in the case where there is only one population, this reduces to

$$T^*(z_1) = \frac{1 - \pi_1^{z_1}}{1 - \pi_1^{N_1}}$$

which is very similar to the result derived for other models such as the diffusion model and the Moran model. In the above notation the result is

$$T^*(z_1) = \frac{1 - e^{-2z_1(\alpha_1 - 1)}}{1 - e^{-2N_1(\alpha_1 - 1)}}$$

as $e^{-2(\alpha_1-1)}$ is a good approximation to π_i when $\alpha_1 - 1$ is small.

In the special case where in each subpopulation there is no selection, Pollak (1966) shows that the probability of fixation is equal to the initial frequency in the whole population, that is,

$$T^*(z_1, \ldots, z_N) = \frac{\sum_{i=1}^K z_i}{\sum_{i=1}^K N_i}$$

which has been shown also using other models.

1.7.3 Diffusion models

Maruyama (1970a; 1977) looked at subdivided populations using a diffusion model. Maruyama shows that the fixation probability in this case is the same as the fixation probability for a single population, that is,

$$u(x_0) = \frac{1 - e^{-4sx_0}}{1 - e^{-4sN}}.$$

The assumptions used in the first proof by Maruyama (1970a) are that the population is divided into connected subpopulations. Migration does not affect the gene frequency in the total population although it may affect the gene frequency or even the population structure at the local level. Selection is additive of magnitude s and is constant throughout the population. The total population size is constant at $2N$. It is not made clear in Maruyama's paper when in the life cycle migration takes place. Essentially a single population is being used as approximation to a subdivided population for the same reasons as in Maruyama's Moran model (Maruyama, 1974) - see section 1.7.1.

Maruyama (1977) gives a second proof of this theorem using differential equation. He assumes that selection is genic and that migration may change the local distribution of genes and colony size but not the global frequency. Once again there is no migration explicitly in the model. Migration at no stage explicitly affects the transition probabilities or population sizes in the model. Maruyama then uses a time transformation based on the heterozygosity in the population to simplify the differential equation that must be solved and obtains the same expression for the fixation probability as for the undivided population.

$$u(x_0) = \frac{1 - e^{-4sx_0}}{1 - e^{-4sN}}$$

which is the same as the solution for an undivided population. What Maruyama is actually finding here is the fixation probability under no net migration. There may be migration from one population to another but overall, population sizes will remain the same in the long run. As with Maruyama's Moran model, because migration does not effect gene frequencies in the whole population this model can be thought of as a single population which is arbitrarily divided at each generation. Breeding occurs and then the population is arbitrarily divided again at the next generation but the migration at no stage affects the gene frequencies in the population. It also does not effect the expected number of offspring per individual directly or indirectly – it only effects where each individual breeds.

1.7.4 Absorption and fixation times

The calculation of absorption and fixation times for a model involving a subdivided population is extremely difficult. It has not been attempted for the case of Wright-Fisher or Moran models and for diffusion processes there are only bounds on the times rather than exact times.

Slatkin (1981) used a diffusion model to find bounds for the fixation time for a subdivided population. Slatkin (1981) assumed that there are n subpopulations each of size N. Selection and random mating take place in each subpopulation separately and migration occurs at the gamete stage. It is also assumed that migration is symmetric, that is the fraction of gametes that go from population i to population j is equal to the proportion of gametes that goes from population j to population i for all pairs i, j. This means that migration may change gene frequencies locally but will not change the gene frequency in the whole population. An upper and lower bound for the fixation times are found as follows. The lower bound for the fixation times is the high migration limit. To calculate the high migration limit it is assumed that the subpopulations have migration between them at such a high level that the total population can be considered as a single population. The high migration limit is the absorption time for a single population with size equal to the total of the subpopulations. To calculate the low migration limit it is assumed that migration decreases towards the case of n isolated colonies. Migration is sufficiently small that selection and genetic drift within each subpopulation take place independently of genetic drift and selection in the other subpopulations. The only role of migration is to introduce new alleles to subpopulations. That is, because migration is at such a low level, the fixation or loss of an allele in a subpopulation and migration of individuals from one population to another are on different timescales. The time between migration events will be much greater than the time to fixation or loss within an individual subpopulation so these processes can be considered seperately. The separation of the timescales effectively means that all subpopulations are almost always fixed for one of the alleles. By considering a new Markov process where the populations are fixed for either one allele or the other and transition events are individual migrations with transition probabilities calculated according to which allele the source subpopulation is fixed for, it is possible to find that the low migration limit is independent of the migration rates. In the case where there is initially one subpopulation fixed for allele A_1 the fixation probability is given by the formula

$$\begin{aligned} U &= \frac{1-v(N)/u(N)}{1-(v(N)/u(N))^n} && u(N) \neq v(N) \\ &= 1/n && u(N) = v(N) \end{aligned}$$

where $u(N)$ is the probability of ultimate fixation of allele A_1 in a subpopulation of size N when it is introduced with initial frequency $1/2N$ and $v(N)$ is the probability of fixation of allele A_2 in a subpopulation of size N when it is introduced with frequency $1/2N$. In this new process the migration parameter, m, represents the fraction of each subpopulation which receive a migrant in each generation. The results of Kimura and

Ohta (1969a) are used to find approximations for the fixation times. For the case where the homozygotes have equal fitness it can be shown that the upper bound of the expected time to fixation (the low migration limit) for the island model is

$$T^*(1/2N) = \frac{n}{2Nmu(N)}$$

and for the stepping stone model is

$$T^*(1/2N) = \frac{n^2}{12Nmu(N)}.$$

These two values found using the new Markov process are for the extreme population structures and provide bounds for most other types of population structure. Slatkin also compares these results with results obtained using simulation.

1.8 Cyclical variation in population size

The simplest variation in population size to deal with is a deterministic cyclical variation. This problem is generally dealt with by finding a population with constant size N_e^* which has the same effective size as the varying population. Wright (1939) showed that this effective population size is the harmonic mean of the effective population sizes in each generation during the cycle. If the cycle is k generations long and $(N_e)_t$ is the effective population size of the varying population at generation t of the cycle then

$$\frac{1}{N_e^*} = \frac{1}{k}\sum_{t=1}^{k}\frac{1}{(N_e)_t}.$$

The results previously obtained can then be applied to this population of constant size N_e^*. The fact that the harmonic mean is used to find the effective population size will cause the effective population size to be more strongly influenced by the smallest population size than the largest.

For branching processes the analysis is more complicated and rather than just substituting a different effective population size, the analysis is redone with the equation under consideration appropriately modified. This is discussed by Ewens (1967).

1.9 Other related models

The models considered so far cover most of the classes of models that have been used to calculate fixation probabilities and mean fixation times. There are some interesting variations on these models which have been used and usually involve relaxing further assumptions of one of the model types considered so far in this chapter. In this section I will briefly discuss some other models which are of relevance to this thesis.

It was assumed in all of the models considered so far in this chapter that selection is constant throughout the population and constant through time. Ohta (1972) considered a single population model where the selection coefficient changed over time. The average level of selection was assumed to be $\bar{s}$. Ohta found that not only is $N_e\bar{s}$ important in determining the fixation probability, but also the ratio $\bar{s}/V_s$, where V_s is the variance of s, is also important in determining the fixation probability. When the ratio $\bar{s}/V_s$ is small then the allele acts as a neutral allele regardless of the magnitude of $\bar{s}$. Other authors such as Gillespie (1973) and Jensen (1973) have also considered the problem using other methods and criticised the methods of Ohta (1972). Although the details of their results differ, all of these authors have reached the same basic conclusion. The main result which comes from all of these models is that variance in fitnesses is very important in determining the fixation probability of a mutant allele sometimes even more important than mean fitness.

Tachida and Iizuka (1991) consider a model where the population is divided into two subpopulations of sizes N_1 and N_2. The level of selection is strong and different in these two subpopulations. It is assumed that migration is nonsymmetric and weak. The level of migration from subpopulation 1 to subpopulation 2 is parameterised by the variable m_2 and from subpopulation 2 to subpoulation 1 is parameterised by m_1. They find that the fixation probability for an allele which starts with a single mutant of the a allele is

$$u = \left(\frac{N_1 N_2 u_{a1} u_{a2}}{N_1 + N_2}\right)\left(\frac{m_2}{N_1 m_1 u_{A1} + N_2 m_2 u_{a2}} + \frac{m_1}{N_2 m_2 u_{A2} + N_1 m_1 u_{a1}}\right)$$

where u_{ai} and u_{Ai} are the probabilities of fixation of the a and A alleles respectively in subpopulation i starting from a single allele in subpopulation i. This is considered to be the low migration limit of the fixation probability as it is assumed in the model that migration is weak and so the mean time between migration events is much greater than the mean time for fixation of an allele in a subpopulation.

Barton (1993) considers a subdivided population with symmetric migration based upon the Wright-Fisher with the added assumption that subpopulations can randomly become extinct. He finds that the probability that an allele is fixed is independent of the population structure when there is no random extinction of subpopulations. The introduction of random extinction of subpopulations reduces the fixation probability of a favoured mutant allele, that is it reduces the effect of selection. This result is derived using a diffusion model.

1.10 Deterministic Models

Although they are not strictly relevant to this thesis, deteministic models can also be used in genetical modelling (see for example, Hill, 1974). An implicit assumption in deterministic models is that populations are large (effectively infinite). When a population is very large it is not possible to find a fixation probability as the population will go to an equilibrium rather than fixation or loss. Deterministic models can only be used to consider what happens to a population as they approach equilibrium or to determine what the equilibrium will be.

Chapter 2

General description of models used in this thesis

Throughout this thesis I use many different models of populations, however there are some underlying assumptions common to all of these models which are described in detail here. A physical interpretation and justification are given for each of these assumptions. The model assumptions are also compared to the model used by Maruyama (1970a), because this section is intended to give a clearer picture of how this model is different from models used by other authors and why it can be used to extend the results obtatined by other authors.

Assumptions

All of the models used in this thesis are based on the Wright-Fisher Model (Wright, 1931). Depending on the model, the population may be divided into a number of subpopulations, there may or may not be selection and the population may vary in size. Any variation from the basic model described here will be specified when a particular model is used.

The models are discrete in time and have a discrete state space. Each state in the state space represents the number of adult individuals present in each of the subpopulations at time t. Each point in time represents a generation so there is no overlap of generations. This is an approximation for most species as it would require that each generation breeds to create the next generation then dies before the next generation breeds. In those species where this is an approximation, it will not affect the fixation probability but will have some effect on the fixation time.

One haploid locus is considered and at this locus there are two alternative alleles, A_1 and A_2. There is no mutation as the populations considered are assumed to be small. There is a maximum of one successful migrant from a subpopulation to any other subpopulation in any generation, this assumption is standard in "classical" metapopulations, see Hanski and Gilpin (1991). This assumption is made for reasons of mathematical simplicity and can be relaxed. This makes the calculations substantially more algebraically complicated but does not change the model conceptually. The population is considered to be constant in size in the sense that the population has a constant number of adults at each generation. Where the population is subdivided the subpopulations may be different in size but each subpopulation has a constant number of adults over time.

There is migration between subpopulations and this migration is stochastic and asymmetric. The migration occurs at the juvenile stage of the life cycle. The life cycle is shown in figure 2.1.

An implicit assumption in the models is that there are many juveniles produced by each adult in the population when they breed and the next generation is randomly sampled from these juveniles. This implies a close approximation to the binomial distribution for the genotypes of the adults of the next generation. This approximation

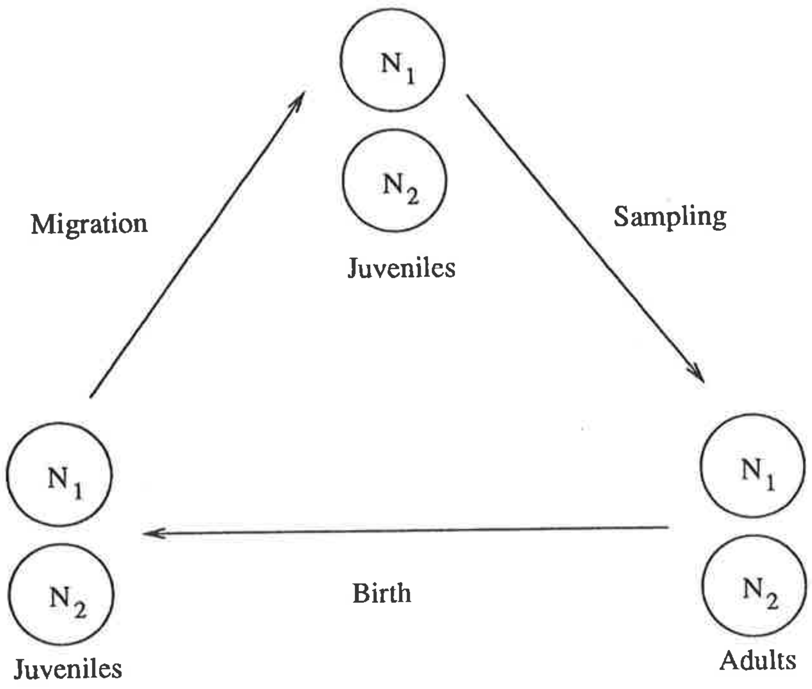


Figure 2.1: The life cycle assumed in this model.

is standard in Wright-Fisher models. Clearly in practice this will not be exact but in most situations it will be a close approximation unless infant mortality is very low and almost all juveniles recruit to the breeding stock. The approximation is very close for species with high adult fecundity and high juvenile mortality such as plants, many marine species and arthropods.

The term "juveniles" here can have different meanings depending on the context but it is always a non-breeding (pre-adult) stage which has the ability to migrate directly or indirectly from one subpopulation to another. Individuals themselves do not have to move but genes must move between the populations to have migration in the sense being considered here. For example, in the case of animals this will be juveniles in the normal sense of the word, in the case of some marine organisms this may be larvae being moved by currents, and in the case of plants the two main ways in which genes can move are as seeds and as pollen. In the case of plants, a pollenator - an insect or bird or the wind - can move genes between the populations without an individual (or even a seed) of the plant species of interest ever physically moving.

Where migration occurs it is assumed to be small relative to the population size. As a result of this, it is assumed that when juveniles migrate from a population this has no effect on the number of juveniles in the population which they are leaving.

The number of juveniles which actually migrate is not directly specified in the model, the probability of a migrant used in the models is actually the probability of a successful migrant. So juveniles are migrating in proportion to the migration probabilities and then being sampled with all of the other juveniles in the destination population. The important thing is that the relative proportion of juveniles from each population is known as this determines the probabilities which are used to calculate the distribution of alleles in the next generation. There is a restriction to a maximum of one successful migrant from subpopulation i to subpopulation j in a given generation for all pairs of subpopulations (i, j). This is to simplify the algebra and does not conceptually change the model. It is a reasonable assumption when populations are small and migration is low.

Although the model here uses only two alleles, the results can easily be extended to the case of more than two alleles by considering each allele separately. For example, if there are three alleles, A_1, A_2 and A_3, then the fixation probability for A_1 alleles can be found by thinking of the population as containing A_1 alleles and combining A_2 and A_3 alleles. This can then be repeated to find the fixation probability for alleles A_2 and A_3. This method will not work if there is selection as alleles with different selection coefficients can not be combined.

2.1 Description of models in this thesis

In this section I summarise the models which I use, how these models relate to models of other authors, and how these models are modified when considering each of the situations of interest.

The models which I develop differ from the models discussed in chapter 1 in that they specifically consider small populations and find exact solutions. In some instances it is even possible to find analytic solutions.

Using an exact stochastic model for a subdivided population I show that the probability of fixation of a neutral allele is equal to a weighted average of the fixation probabilities in each subpopulation considered separately. I then find a numeric formula for the fixation probability of fixation in a subdivided population when there is selection acting and show that the probability of fixation is not the same for a subdivided population as it is for a single homogeneous population. Interpretations of the results and examples are given. Finally I consider a population which varies cyclically over time and show that an effective population size approximation, N_e, Wright (1939) goves a reasonably accurate prediction of fixation probabilities when compared to the exact stocahstic model.

In chapter 3 I consider a two subpopulation model with no selection and find an analytic formula for the fixation probability. These results are compared with the results obtained by other authors. The effect of asymmetric migration and some examples are considered.

In chapter 4 I consider the same population structure as in chapter 3 and find a general matrix formula for the mean absorption and mean fixation times. This formula can be solved analytically for small populations. Numerical examples are given for larger populations. This is compared with the results of other authors for the case when migration is symmetric. The effect of asymmetric migration is also considered.

In chapter 5 I extend the results of chapters 3 and 4 to the case when there are three and four subpopulations. The formulas for fixation probabilities and mean absorption times rapidly become more complicated as the number of populations increases but there are still some informative special cases and examples which can be examined. Symmetric and asymmetric migration are compared and the relative importance of edge patches and centre patches is considered.

In chapter 6 I consider a general population structure with n subpopulations and use graph theory to find an analytic solution for the fixation probability in this case. The numbers of terms in the weight coefficients can be very large ($= n^{n-2}$) so it is not possible to interpret the general case. However, it is possible to give an interpretation of the terms that make up the coefficients using graph theory. Specific cases can also be considered to give some useful insights into fixation probabilities in some general cases.

In chapter 7 I introduce selection to the model. Initially I consider a single population with selection acting. This can be thought of as a haploid population or a diploid population with additive selection. Once again, explicit formulas can be found for small population sizes and a matrix formula is found for larger population sizes which can be solved numerically. A matrix formula is also found for the mean absorption time. This matrix formula is solved numerically for some special cases and its behaviour in general is discussed. The selection model is then extended to a two population model and a matrix formula for the fixation probability is found. Some special cases are solved numerically and discussed. These results are compared with the results of other authors who have considered this question using other models.

In chapter 8 I consider a single population model with cyclically varying size. Matrix formula are found and these are compared with the results that would be obtained using the effective population size concept discussed in section 1.8. I also consider populations with randomly varying size and derive matrix formulae for the fixation probability and mean absorption times for a general random process of population size fluctuation.

Chapter 9 is a summary of the thesis. I describe the major results of the thesis and discuss what further work could be done using the models which are used in this thesis.

Chapter 3

Fixation probability with two subpopulations

3.1 Introduction

In this chapter I consider the problem of finding the fixation probability for a simple population structure – a population which is divided into two subpopulations with migration between them. The model used in this chapter has a very simple population structure. It is presented primarily as a step towards the models which will be considered in later chapters. It is also useful to consider this model as it has a simpler state space than the models considered in later chapters so the results obtained here can be more easily displayed and interpreted than the results in later chapters where the state space becomes far more complicated. I find the fixation probability for an allele when migration is stochastic and asymmetric and there is no selection. The expression that I find for the fixation probability is then compared with expressions obtained by Maruyama (1970a; 1970b) using a diffusion model, Maruyama (1972) using a Moran model, and Slatkin (1981) using a diffusion model. For each of these comparisons I only consider the specific case of symmetric migration as each of the models by other authors assumes symmetric migration. Tachida and Iizuka (1991) also consider a problem similar to this with asymmetric migration using the approach of Slatkin (1981). They assume weak migration with modest to strong selection. The results I present here confirm and extend the results of the previous authors using a different method.

Finally the implications of this theorem are discussed and some examples of the application of the theorem are considered, including the case of asymmetric migration which extends the work of the other authors mentioned previously.

3.2 Description of model

The mathematical model used is a two population version of the model described in Chapter 2. Further to the assumptions outlined in Chapter 2, there is no selection and the two subpopulations are of constant size N_1 and N_2. Migration is allowed between the two populations and the mean number of migrants from subpopulation 1 to subpopulation 2 in a given generation is μ_{12}, the mean number of migrants from subpopulation 2 to subpopulation 1 is μ_{21}. The state variable in the model, (X_t, Y_t), is the number of alleles of type A_1 present in each subpopulation at a given time, t, where X_t is the number present in subpopulation 1 and Y_t is the number present in subpopulation 2. I also assume that $X_{t+1}|X_t, Y_t$ and $Y_{t+1}|X_t, Y_t$ are independent – this follows from the fact that sampling of juveniles within one subpopulation is independent of sampling in the other subpopulation.

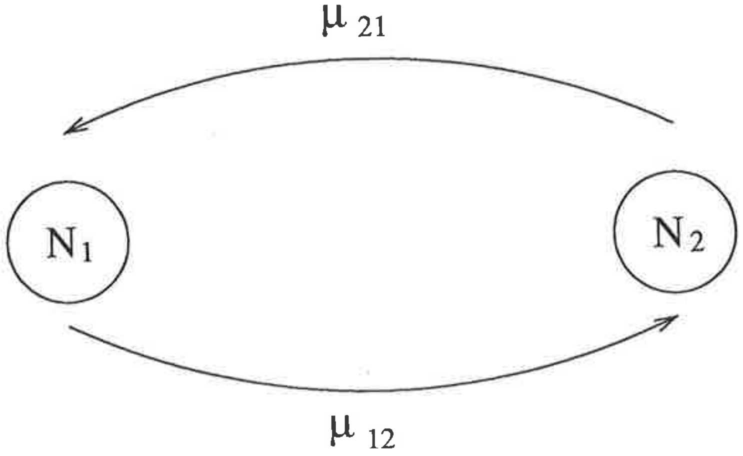


Figure 3.1: Diagramatic representation of the model

3.3 Transition probabilities

In this section I will derive the transition probabilities for subpopulation 1 and state the transition probabilities for subpopulation 2. The number of successful migrants that move between the subpopulations in a given generation is a random variable. The distribution of this random variable has a Bernoulli distribution as there is either 0 or 1 migrants in each direction in each generation. The mean number of successful-migrants from subpopulation 2 to subpopulation 1 in a generation is μ_{21} which is also the probability that there will be a migrant from subpopulation 2 to subpopulation 1 as the random variable has a Bernoulli distribution. The number of migrants from subpopulation 2 to subpopulation 1 in generation t is denoted by the random variable $M_{21}(t)$, so $E(M_{21}(t)) = \mu_{21}$. The number of these migrants which have A_1 alleles is denoted by the random variable $M'_{21}(t)$. The transition probabilities can then be derived by noticing that for subpopulation 1 to have k A_1 alleles at time $t+1$ when the subpopulations at time t contain X_t and Y_t individuals respectively. If there are $m'_{21}(t)$ successful migrants from subpopulation 2 with the A_1 allele where $m'_{21}(t)$ is an observation of the random variable $M'_{21}(t)$ there must be $k - m'_{21}(t)$ individuals born in subpopulation 1 with the A_1 allele which then become adults in subpopulation 1.

So, using the abbreviations "migs" for migrants and "indivs" for individuals,

$$\begin{aligned}P(X_{t+1}=k|X_t=i,Y_t=j) &= \sum_{m'=0,1} P(m'_{21}(t+1) \text{ migs from pop 2 with } A_1 \text{ allele},\\ &\qquad k-m'_{21}(t+1) \text{ indivs from pop 1 with } A_1 \text{ allele}).\end{aligned} \tag{3.1}$$

An analagous argument holds for the distribution of Y_{t+1} and

$$\begin{aligned}P(Y_{t+1}=l|X_t=i,Y_t=j) &= \sum_{m'=0,1} P(m'_{12}(t+1) \text{ migs from pop 1 with } A_1 \text{ allele},\\ &\qquad k-m'_{12}(t+1) \text{ indivs from pop 2 with } A_1 \text{ allele}).\end{aligned} \tag{3.2}$$

3.4 Fixation probability

Theorem 1 *The probability of fixation of an allele, when the initial number of alleles present is x_0 in subpopulation 1 and y_0 in subpopulation 2 is*

$$\alpha(x_0,y_0)=\frac{\mu_{12}x_0+\mu_{21}y_0}{\mu_{12}N_1+\mu_{21}N_2}, \tag{3.3}$$

Proof To simplify notation, let

$$\begin{aligned}P(i,j;k,l) &= P(X_{t+1}=k,Y_{t+1}=l|X_t=i,Y_t=j). &(3.4)\\ &= P(X_{t+1}=k|X_t=i,Y_t=j)P(Y_{t+1}=l|X_t=i,Y_t=j). &(3.5)\end{aligned}$$

using the independence of $X_{t+1}|X_t,Y_t$ and $Y_{t+1}|X_t,Y_t$.

The probability that the A_1 allele will be fixed in the whole population from a given starting state, (i,j), is equal to summing the probability that the first transition is to state (k,l), over all states (k,l), and that fixation then occurs from state (k,l). Written mathematically this is

$$P(\text{fixation from } (i,j)) = \sum_{(k,l)} P((i,j) \to (k,l)) P(\text{fixation from } (k,l))$$

or

$$\alpha(i,j) = \sum_{k=0}^{N_1} \sum_{l=0}^{N_2} P(i,j;k,l)\alpha(k,l) \tag{3.6}$$

Using equation 3.5, equation 3.6 can be rewritten as

$$\alpha(i,j) = \sum_{k=0}^{N_1} \sum_{l=0}^{N_2} P(X_{t+1} = k|X_t = i, Y_t = j) P(Y_{t+1} = l|X_t = i, Y_t = j)\alpha(k,l) \tag{3.7}$$

Substituting the solution from equation 3.3 into the left hand side of equation 3.7 gives

$$l.h.s = \frac{\mu_{12} i + \mu_{21} j}{\mu_{12} N_1 + \mu_{21} N_2}.$$

And substituting equation 3.3 into the right hand side of equation 3.7 gives

$$\begin{aligned}
r.h.s &= \sum_{k=0}^{N_1} \sum_{l=0}^{N_2} P(X_{t+1} = k|X_t = i, Y_t = j) P(Y_{t+1} = l|X_t = i, Y_t = j) \frac{\mu_{12} k}{\mu_{12} N_1 + \mu_{21} N_2} \\
&\quad + \sum_{k=0}^{N} \sum_{l=0}^{N} P(X_{t+1} = k|X_t = i, Y_t = j) P(Y_{t+1} = l|X_t = i, Y_t = j) \frac{\mu_{21} l}{\mu_{12} N_1 + \mu_{21} N_2} \\
&= \sum_{k=0}^{N_1} P(X_{t+1} = k|X_t = i, Y_t = j) \frac{\mu_{12} k}{\mu_{12} N_1 + \mu_{21} N_2} \sum_{l=0}^{N_2} P(Y_{t+1} = l|X_t = i, Y_t = j) \\
&\quad + \sum_{l=0}^{N_2} P(Y_{t+1} = l|X_t = i, Y_t = j) \frac{\mu_{21} l}{\mu_{12} N_1 + \mu_{21} N_2} \sum_{k=0}^{N_1} P(X_{t+1} = k|X_t = i, Y_t = j) \\
&= \sum_{k=0}^{N_1} P(X_{t+1} = k|X_t = i, Y_t = j) \frac{\mu_{12} k}{\mu_{12} N_1 + \mu_{21} N_2} \\
&\quad + \sum_{l=0}^{N_2} P(Y_{t+1} = l|X_t = i, Y_t = j) \frac{\mu_{21} l}{\mu_{12} N_1 + \mu_{21} N_2} \\
&= E\Big(\frac{\mu_{12} X_{t+1}}{\mu_{12} N_1 + \mu_{21} N_2} |X_t = i, Y_t = j\Big) + E\Big(\frac{\mu_{21} Y_{t+1}}{\mu_{12} N_1 + \mu_{21} N_2} |X_t = i, Y_t = j\Big) \\
&= \frac{\mu_{12}}{\mu_{12} N_1 + \mu_{21} N_2} E(X_{t+1}|X_t = i, Y_t = j) + \frac{\mu_{21}}{\mu_{12} N_1 + \mu_{21} N_2} E(Y_{t+1}|X_t = i, Y_t = j).
\end{aligned} \tag{3.8}$$

Now it remains to calculate the expected values in the above expression and substitute them into equation 3.8. The expected value of X_{t+1} can be calculated using its conditional distribution and the definition of the transition probability given in equations 3.1. The assumption that not all juveniles recruit to the breeding stock means that the juveniles which are recruited to the breeding stock (the only ones which we are interested in) are essentially randomly drawn from a large pool. So the binomial distribution is a fair approximation to the distribution of juveniles which recruit to the breeding stock. (Note: it is not necessary to use the binomial distribution for the proof to work, all that is required is for the distribution to have expected number of A_1 individuals equal to subpopulation size *times* proportion of A_1 in relevant subpopulation.)

In generation $t+1$ the observed number of successful migrants from subpopulation 2 to subpopulation 1 is $m_{21}(t+1)$. In this case X_{t+1} has conditional distribution

$$\begin{aligned} &X_{t+1}|X_t = i, Y_t = j, m_{21}(t+1) \text{ migs from pop 2} \\ &\quad\sim Bin(N_1 - m_{21}(t+1), i/N_1) + Bin(m_{21}(t+1), j/N_2) \end{aligned}$$

and when there are $m_{21}(t+1)$ migrants from subpopulation 2 to subpopulation 1 in generation t, X_{t+1} has expected value conditional on the number of migrants of

$$E(X_{t+1}|M_{21}(t+1) = m_{21}(t+1)) = \frac{i(N_1 - m_{21}(t+1))}{N_1} + \frac{jm_{21}(t+1)}{N_2}.$$

So the unconditional expected value of X_{t+1} is

$$\begin{aligned} E(X_{t+1}) &= E(E(X_{t+1}|\text{number of migrants})) \\ &= E\Big(\frac{i(N_1 - M_{21}(t+1))}{N_1} + \frac{jM_{21}(t+1)}{N_2}\Big) \\ &= \Big(1 - \frac{\mu_{21}}{N_1}\Big)i + \frac{\mu_{21}}{N_2}j \end{aligned}$$

and analogously using equation 3.2,

$$E(Y_{t+1}) = \left(1 - \frac{\mu_{12}}{N_2}\right)j + \frac{\mu_{12}}{N_1}i.$$

Substituting the expected values back into equation 3.8 gives

$$\begin{aligned} r.h.s &= \left[\frac{\mu_{12}}{\mu_{12}N_1+\mu_{21}N_2}\right]\left[\left(1-\frac{\mu_{21}}{N_1}\right)i+\frac{\mu_{21}}{N_2}j\right]+\left[\frac{\mu_{21}}{\mu_{12}N_1+\mu_{21}N_2}\right]\left[\left(1-\frac{\mu_{12}}{N_2}\right)j+\frac{\mu_{12}}{N_1}i\right] \\ &= \frac{\mu_{12}i+\mu_{21}j}{\mu_{12}N_1+\mu_{21}N_2} \end{aligned}$$

as required.

□

3.5 Discussion

The first important conclusion from theorem 1 is that the probability of fixation of a neutral allele is not equal to the initial frequency in a subdivided population unless migration is symmetric, that is, unless $\mu_{12} = \mu_{21}$. This result holds irrespective of whether the population sizes are equal. This is consistent with Maruyama (1970a; 1972) who showed that in the case where migration is symmetric, fixation probabilities are not affected by subdividing the population. Slatkin (1980) also showed that fixation probabilities are independent of μ_{ij} (migration parameters) when migration is symmetric. The result is a confirmation of the work of Maruyama and Slatkin using another method.

The result in Section 3.4 is a generalisation of the results of these other authors to the case where migration is asymmetric and is a solution to a more general problem than that tackled by previous authors. When there are two subpopulations with asymmetric migration, there is a *source* subpopulation and a *sink* subpopulation. The terms "source" and "sink" have been used by different authors to mean different things (Pulliam, 1988, Tuck and Possingham, 1994). In this context a subpopulation which is a net exporter of individuals for any reason is called a source. A subpopulation which is a net importer of individuals for any reason is called a sink. Some authors have an implicit or explicit assumption that a source subpopulation is an exporter because it produces excess juveniles due to high fecundity whereas the definition used here includes the case where a subpopulation with average fecundity exports more individuals for other reasons, such as prevailing winds. I also introduce the idea of a *migration ratio* in subsection 3.5.1. This is a measure of the relative flow of individuals between subpopulations and is important in determining the fixation probability when migration is not symmetric.

3.5.1 The "migration ratio"

Here I introduce the idea of the *migration ratio*. First, the result from theorem 1 is rewritten as

$$\alpha(x_0, y_0) = \frac{x_0 + \gamma y_0}{N_1 + \gamma N_2}, \tag{3.9}$$

where $\gamma = \mu_{21}/\mu_{12}$. The physical interpretation for γ is a *migration ratio*. The absolute values of μ_{12} and μ_{21} are not important in determining the fixation probability, the important quantity is the ratio between them. The migration ratio, γ, summarizes the way in which μ_{12} and μ_{21} determine the fixation probability. The qualititave effects of varying γ can be seen by rearranging equation 3.9 to give

$$\begin{aligned} \alpha(x_0, y_0) &= \left(\frac{N_1}{N_1 + \gamma N_2}\right)\frac{x_0}{N_1} + \left(\frac{\gamma N_2}{N_1 + \gamma N_2}\right)\frac{y_0}{N_2} \\ &= \left(\frac{N_1}{N_1 + \gamma N_2}\right)\frac{x_0}{N_1} + \left(1 - \frac{N_1}{N_1 + \gamma N_2}\right)\frac{y_0}{N_2}. \end{aligned}$$

So the fixation probability is a weighted average of the initial frequency in the two populations. The weights add to 1 and the function $N_1/(N_1 + \gamma N_2)$ is a monotonic function of γ when $\gamma > 0$,

$$\frac{N_1}{N_1 + \gamma N_2} \rightarrow 1 \text{ as } \gamma \rightarrow 0$$

$$\frac{N_1}{N_1+\gamma N_2} \rightarrow 0 \text{ as } \gamma \rightarrow \infty.$$

In the case where $\gamma = 1$ (symmetric migration), the initial frequencies x_0/N_1 and y_0/N_2 are weighted in proportion to the subpopulation sizes and the fixation probability is equal to the initial frequency in the whole population. As γ increases the weight of subpopulation 1 will decrease and the weight of subpopulation 2 will increase. So if $\gamma > 1$ ($\mu_{21} > \mu_{12}$) and genes tend to flow from subpopulation 2 to subpopulation 1, then subpopulation 2 is relatively more important in determining the fixation probability than subpopulation 1. This means that subpopulation 2 receives a greater weight than population 1.

Similarly, if $\gamma < 1$ ($\mu_{12} > \mu_{21}$) then individuals tend to flow from subpopulation 1 to subpopulation 2, and hence subpopulation 1 is relatively more important in determining the fixation probability than subpopulation 2. This means that subpopulation 1 receives a greater weight than subpopulation 2.

This shows that the value of γ can have a substantial impact on the fixation probabilities even though the number of migrants between subpopulations is at most one per generation.

It also follows that because the fixation probability is a weighted average of the initial frequency in each population and the weights sum to 1, the value of the fixation probability for the subdivided population is between the initial frequencies in the two populations.

3.5.2 Examples

Example 1: If $\mu_{12} = \mu_{21}$ ($\gamma = 1$) and $N_1 = N_2 = N$

then the fixation probability is

$$\alpha(x_0, y_0) = \frac{x_0 + y_0}{2N},$$

which is equal to the initial frequency of the allele in the whole population. This is the case considered by Maruyama (1970a; 1972) and Slatkin (1981) and the result is the same as the result they obtained using their own methods.

Example 2: If $\mu_{12} = \mu_{21}$ ($\gamma = 1$) and $N_1 \neq N_2$ then the fixation probability is

$$\alpha(x_0, y_0) = \frac{x_0 + y_0}{N_1 + N_2}$$

which once again is the initial frequency of the allele in the whole population. So when population sizes are different the result is qualitatively the same as the result in example 1 although this result is an extension to the earlier results of Maruyama and Slatkin who assumed that all of the subpopulations were the same size.

Example 3: When $\mu_{12} \neq \mu_{21}$, the fixation probability will no longer be equal to the initial frequency in the whole population. The probability depends on γ and the relative sizes of the two populations. This situation is qualitatively different from example 1 and example 2 as migration is no longer symmetric. The result obtained is also qualitatively different from the symmetric migration case as the subpopulation which is acting as a source, that is, a net exporter of migrants, has a relatively greater influence on the fixation probability than the subpopulation that is acting as a sink, that is, a net importer of migrants. There are many examples in biological systems where a population is subdivided and there is a tendency for migrants to move more in one direction than the other direction. Prevailing winds or currents can provide unequal migration between two otherwise similar habitats – for example Possingham and Roughgarden (1990), Richards *et al* (1995), Burton *et al* (1979), Leary *et al* (1993). Subpopulations in habitats of different value to the species can lead to different numbers of potential migrants; and subpopulations of different sizes will, in general, also produce different numbers of potential migrants. In each of these situations the subpopulation which produces more migrants is relatively more important in determining which allele will eventually become fixed in the whole population.

Example 4: A more extreme example is when migration is only in one direction. Suppose that subpopulation 1 is a source and subpopulation 2 is a sink, then $\mu_{21} = 0$ and $\mu_{12} > 0$, so $\gamma = 0$. This means that the probability of fixation of the A_1 allele is

$$\alpha(x_0, y_0) = \frac{x_0 + (0 \times y_0)}{N_1 + (0 \times N_2)} = \frac{x_0}{N_1}$$

which is equal to the initial frequency of the A_1 allele in subpopulation 1, the source subpopulation. Note that this is also equal to fixation probability in subpopulation 1 considered separately.

Example 5: Paradoxically, if subpopulation 1 is much larger than subpopulation 2, then subpopulation 1 may actually have the greatest effect on the fixation probability even if there is net migration from subpopulation 2 to subpopulation 1. That is, if $N_1 \gg N_2$ and $\mu_{21} > \mu_{12}$ then the initial frequency in subpopulation 1, x_0/N_1, may have a larger weight than the initial frequency in subpopulation 2, y_0/N_2, even though genes tend to flow from subpopulation 2 to subpopulation 1. So even though subpopulation 2 is a source and subpopulation 1 is a sink, subpopulation 1 will have a greater effect on the fixation probability for the whole population. From the biological point of view, subpopulation 1 will be less affected by immigrants as they will form a small proportion of the large subpopulation. Migrants from subpopulation 1 to subpopulation 2 will have a relatively larger effect on that subpopulation even though they occur less often.

In figure 3.2 the initial frequencies are fixed at 0.8 in subpopulation 1 and 0.2 in

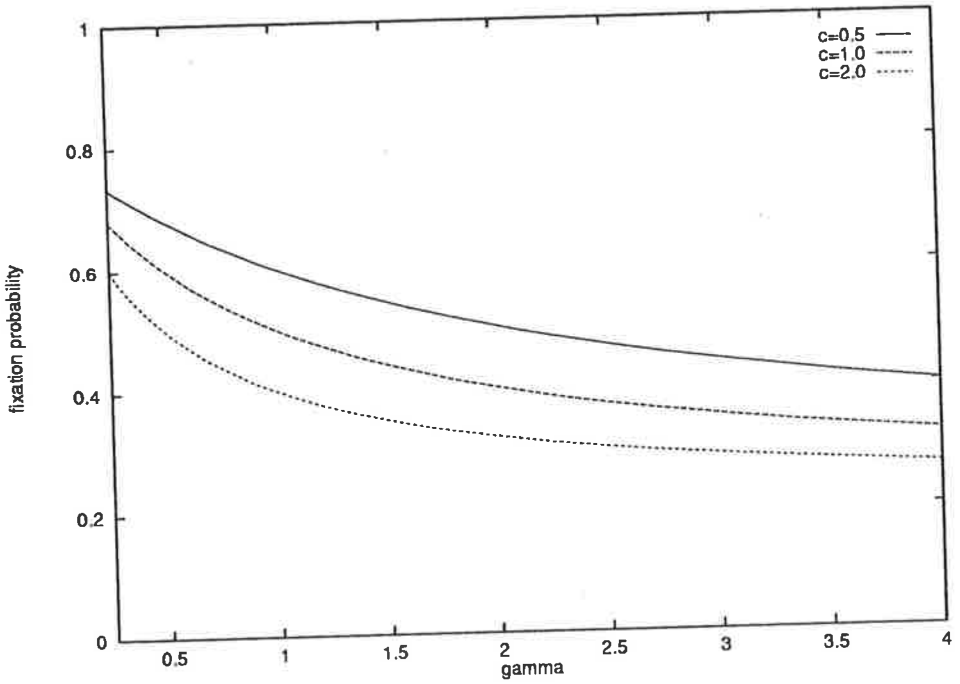


Figure 3.2: Model with $N_1 = cN_2$ for $c = 0.5, 1, 2$. Graph shows the probability of fixation of allele A_1 for varying values of γ, when initial frequency of A_1 is 0.8 in subpopulation 1 and 0.2 in subpopulation 2.

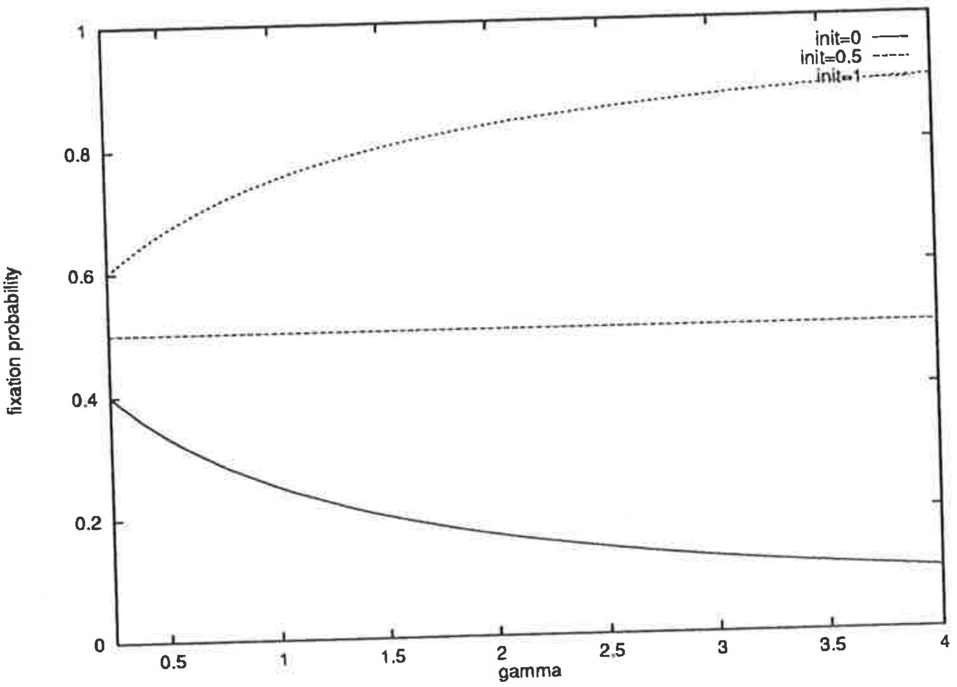


Figure 3.3: Model with $N_1 = N_2$ for $y_0 = 0, 0.5, 1$. Graph shows the probability of fixation of Allele when initial frequency is 0.5 in subpopulation 1 and varying migration ratio, γ.

subpopulation 2. The fixation probability is given as a function of the migration ratio, γ for three different ratios of the subpopulation sizes, c. Note that for each value of c, the fixation probability is closer to the initial frequency in subpopulation 1 when γ is low and closer to the initial frequency in subpopulation 2 when γ is high. This graph also shows that if subpopulation 2 is larger than subpopulation 1 then, for the same value of γ, the fixation probability is closer to the initial frequency in subpopulation 1 than the initial frequency in subpopulation 2.

In figure 3.3 the populations are equal in size. The initial frequency in population 1 is fixed at 0.5, and the fixation probability is shown as a function of γ for three different values of the initial frequency in population 2. This graph shows that when there is more migration from population 1 to population 2 the fixation probability is closer to the initial frequency in population 1 and when there is more migration from population 2 to population 1 the fixation probability is closer to the initial frequency in population 2. That is, the population which produces the most migrants has the greatest effect on deciding which gene will eventually be fixed. Note also that when the initial frequency is the same (0.5) in each population then the overall fixation probability is also equal to 0.5, regardless of migration rates and subpopulation sizes.

3.6 Conclusions

The theorem presented here proves that in a subdivided population, the fixation probability of an allele does not depend only on its initial frequency in the entire population but on its initial frequency in each subpopulation, the size of each subpopulation and the rates of migration between the two subpopulations. The overall conclusions are that in general,

i) the subpopulation which produces the most migrants has the greatest effect on the fixation probability,

ii) larger subpopulations have a greater effect on the fixation probability than smaller subpopulations,

iii) the absolute magnitude of the migration parameters, μ_{12} and μ_{21}, are not important in determining the fixation probability. The ratio, $\gamma = \mu_{21}/\mu_{12}$, of the migration parameters is the important quantity.

Note that when $\mu_{12} = \mu_{21} = 0$ the problem is not mathematically defined as there is an absorbing state which is not fixation or loss if the system goes to the state where one subpopulation is fixed for one allele and the other subpopulation is fixed for the other allele. Biologically this is the case when two subpopulations exist in isolation and become fixed for different alleles.

This model will be discussed and developed further in chapter 4 where mean absorption and fixation times of alleles will be considered. The model will be extended to the case of more than two subpopulations in chapter 5 and selection will be added to the two subpopulation model in chapter 7.

Chapter 4

Mean absorption and fixation times with two subpopulations

4.1 Introduction

In this chapter I investigate the mean time to absorption and the mean time to fixation of an allele in a population which is divided into two subpopulations when there is no selection. The model assumptions are the same as the model assumptions used to calculate fixation probabilities in chapter 3.

There are three time quantities that are of interest when considering the survival of an allele in a population. These are the mean times to fixation, loss and absorption. Each of these quantities can be used to compare the rate of evolution under different population structures with different rates of migration between the two populations (Kimura, 1983; Hastings and Harrison, 1994; Lande and Barrowclough, 1987).

The quantity which I consider first is the mean time to absorption of an allele. This is of interest when the main question under consideration is the rate of loss of heterogeneity in the population. A question of relevance to conservation genetics is the rate of loss of variability which is already present in the population (Kimura, 1983; Hastings and Harrison, 1994; Sampson *et al.*, 1988; Templeton *et al.*, 1990). Of particular interest here is the effect that different population structures have on the mean absorption time and hence the mean rate of loss of genetic variation under different population structures. This will hopefully provide some insight into the effect that different management options would have on the loss of genetic variability. One of the major questions in conservation management is how to spread resources – is it better to have a single large reserve or several small reserves? (Margules and Austin, 1990; Walters, 1991; McKelvey *et al.*, 1992). The effect that habitat fragmentation and loss has on the level of genetic variability in a population is a related question of concern. These issues have been considered by many authors, (for example Lande, 1988; Templeton *et al.*, 1990) but generally using population dynamics rather than population genetics.

In studies of evolutionary theory the mean time to fixation is the quantity of interest. Specifically the mean time to fixation of a mutant allele with frequency $1/2N$. This is because an evolutionist is interested in the rate of genetic change from the present state. The rate of evolutionary change is determined by the rate of production of mutant alleles, the probability of fixation of a mutant allele, and the mean time to fixation of those mutant alleles which become fixed. This quantity has been considered when analysing the neutral theory of evolution although the effect that population structure has on the mean fixation time is not usually considered. In contrast, conservation geneticists are interested in the mean absorption time, that is the the mean time to fixation or loss of an allele (Leary *et al.*, 1993; Lande, 1988; Meffe and Vrijenhoek, 1988). This is because both fixation and loss of an allele are a loss of heterogeneity. The analysis for the mean fixation time is exactly analagous to the analysis for the mean absorption time. The mean absorption time is considered here in more detail as it is of most relevance to loss of heterogenity, the question being considered.

4.2 Description of model

The model assumptions used here are exactly those used in the calculation of the fixation probability and are described in section 3.2.

4.2.1 Models for mean absorption and mean fixation times

To calculate the mean time to absorption or fixation of the A_1 allele it is necessary to specify the state of the system at each generation. The state of the system is described as follows: if at time t there are i A_1 alleles in subpopulation 1 and j A_1 alleles in subpopulation 2 then the state of the system is $(X_t, Y_t) = (i, j)$. This means that there are $(N_1 + 1)(N_2 + 1)$ possible states in the system.

In the calculation of mean absorption times, both (0,0) and (N_1, N_2) are absorbing states so there are $(N_1 + 1)(N_2 + 1) - 2$ transient states and 2 absorbing states in the system. The modified process used to calculate mean absorption times will be discussed in section 4.3.

In the calculation of mean fixation times the state (N_1, N_2) is an absorbing state but the state (0,0) is removed from the system as we are only interested in those cases that result in fixation of the allele of interest. In this case there are $(N_1+1)(N_2+1)-1$ states in the system of which $(N_1 + 1)(N_2 + 1) - 2$ are transient and one is absorbing. The modified process used to calculate mean fixation times will be discussed in section 4.5.

4.3 Calculation of mean absorption time

In this section I will derive equations for the mean absorption time. I define $T_{(i,j)}$ to be the expected time to absorption from state (i,j). The transition probability, $P(i,j;k,l)$, is defined in equation 3.4. Let $\mathcal{A}$ be the set of absorbing states for the case under consideration (in this case $\mathcal{A} = \{(0,0),(N_1,N_2)\}$) then the following relationship can be established between the $T_{(i,j)}$'s:

$$\begin{aligned} T_{(i,j)} &= 1 + \sum_{(k,l)\notin\mathcal{A}} P(i,j;k,l)T_{(k,l)} \qquad \text{for}(i,j) \notin \mathcal{A} \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad \text{for}(i,j) \in \mathcal{A} \end{aligned} \tag{4.1}$$

which is a set of $(N_1+1)(N_2+1)-2$ linear equations in $(N_1+1)(N_2+1)-2$ unknowns.

The equations can be rewritten in matrix form with $\mathbf{T}$ = the vector of expected times to absorption in the following format.

$$\mathbf{T} = \begin{pmatrix} T_{(0,1)} \\ T_{(0,2)} \\ \vdots \\ T_{(0,N_2)} \\ T_{(1,0)} \\ T_{(1,1)} \\ \vdots \\ T_{(N_1,N_2-1)} \end{pmatrix} \tag{4.2}$$

And $\mathbf{P}$ is the transition matrix,

$$\mathbf{P} = \begin{pmatrix} P_{(0,1)\to(0,1)} & P_{(0,1)\to(0,2)} & \cdots & P_{(0,1)\to(N_1,N_2-1)} \\ P_{(0,2)\to(0,1)} & P_{(0,2)\to(0,2)} & \cdots & P_{(0,2)\to(N_1,N_2-1)} \\ \vdots & \vdots & & \vdots \\ P_{(N_1,N_2-1)\to(0,1)} & P_{(N_1,N_2-1)\to(0,2)} & \cdots & P_{(N_1,N_2-1)\to(N_1,N_2-1)} \end{pmatrix} \tag{4.3}$$

The order that the states are listed in is not important provided the same order is always used.

Then,

$$\mathbf{T} = \mathbf{1} + \mathbf{PT} \tag{4.4}$$

which can be rearranged to give the following formula for calculating the absorption times for a given transition matrix, $\mathbf{P}$,

$$\mathbf{T} = (\mathbf{I} - \mathbf{P})^{-1}\mathbf{1} \tag{4.5}$$

In general it will be extremely difficult to solve equation 4.5 analytically as it involves inverting a $(N_1+1)(N_2+1)-2$ by $(N_1+1)(N_2+1)-2$ square matrix. The formula can usually be evaluated numerically for fixed values of each of the parameters, even with N_1 and N_2 reasonably large. Gale (1990) gives an analytic formula for upper and lower bounds of the mean fixation time in a single population but a general analytic solution is impossible. Solutions have also been found for other types of model (see Watterson, 1961; Ewens, 1963; Kimura and Ohta, 1969b; Gale, 1990).

4.4 Some examples

In this section I will give an explicit solution for equation 4.5 for the case $N_1 = N_2 = 2$. I will then give numerical solutions for a more complicated case and compare the results obtained with the absorption times for the equivalent single population. This result is also compared with an approximate result by Slatkin (1981).

4.4.1 An analytic example

In this example I evaluate equation 4.5 for the case where $N_1 = N_2 = 2$ and $\mu_{12} = \mu_{21} = \mu$ to give an example of how complicated the mathematics gets even for the simplest of population structures. The absorption times for the three classes of initial state are as follows:

$$\begin{aligned}
T_{(0,1)} = T_{(1,0)} = T_{(2,1)} = T_{(1,2)} &= \frac{4\mu^4 - 20\mu^3 - \mu^2 + 40\mu + 6}{4\mu^4 - 12\mu^3 + 2\mu^2 + 12\mu} \\
T_{(0,2)} = T_{(2,0)} &= \frac{-8\mu^3 + 22\mu + 6}{2\mu^4 - 6\mu^3 + \mu^2 + 6\mu} \\
T_{(1,1)} &= \frac{4\mu^4 - 16\mu^3 + \mu^2 + 25\mu + 3}{2\mu^4 - 6\mu^3 + \mu^2 + 6\mu} \qquad (4.6)
\end{aligned}$$

Each of these is plotted in figure 4.1. The general shape of each graph is fairly flat over most of the range of μ between 0.1 and 1. As μ drops below 0.1, the mean time to absorption rises rapidly, with limit ∞ as μ approaches 0. The reason for this rapid climb for low values of μ is variation between subpopulations, and the amount of time taken for a migrant to move from one subpopulation to the other to allow absorption, which causes delays in the time to absorption. This is because when migration is very low it is possible that the different subpopulations will be fixed for different alleles for a long time before a migrant transfers between the subpopulations. The average time for a migrant to transfer between the subpopulations is proportional to $1/\mu$ generations. When the subpopulations are fixed for different alleles absorption in the population as a whole can not occur until there is a migrant between the two subpopulations. So the absorption time for the whole population increases as the level of migration between the subpopulations decreases.

This can also be thought of in terms of Slatkin's low migration limit (Slatkin, 1981) discussed in section 1.7.4. Slatkin showed as a general rule for two different population structures that the absorption time is proportional to $1/\mu$ where μ is the level of

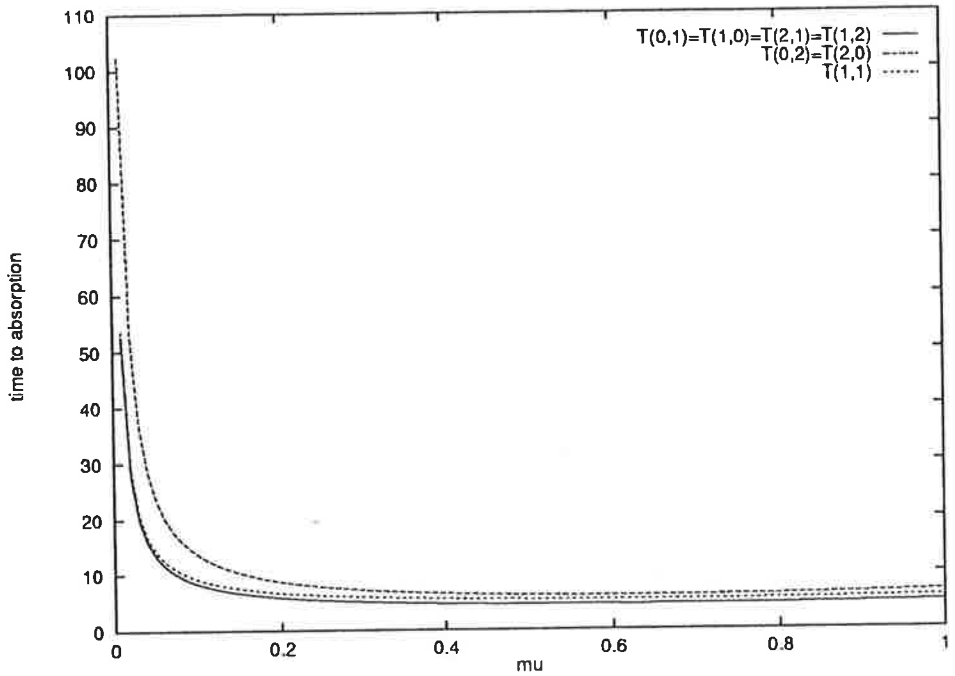


Figure 4.1: Plot of absorption time as a function of migration probability for the three different classes of initial state.

migration between the two subpopulations. As the level of migration becomes lower, equation 4.6 shows that the absorption times are approximately

$$\begin{aligned} T_{(0,1)} = T_{(1,0)} = T_{(2,1)} = T_{(1,2)} &\approx \frac{1}{2\mu} \\ T_{(0,2)} = T_{(2,0)} &\approx \frac{1}{\mu} \\ T_{(1,1)} &\approx \frac{1}{2\mu} \end{aligned} \tag{4.7}$$

As an example of how accurate this approximation is, figure 4.2 shows the absorption time for the states (0,1), (1,0), (2,1) and (1,2), and the approximation analagous to Slatkin's low migration limit. The approximation displays the same qualitative behaviour as the exact results. It is clear from this that for larger values of μ the approximation is poor, but when μ is less than 0.05, the approximation is more accurate.

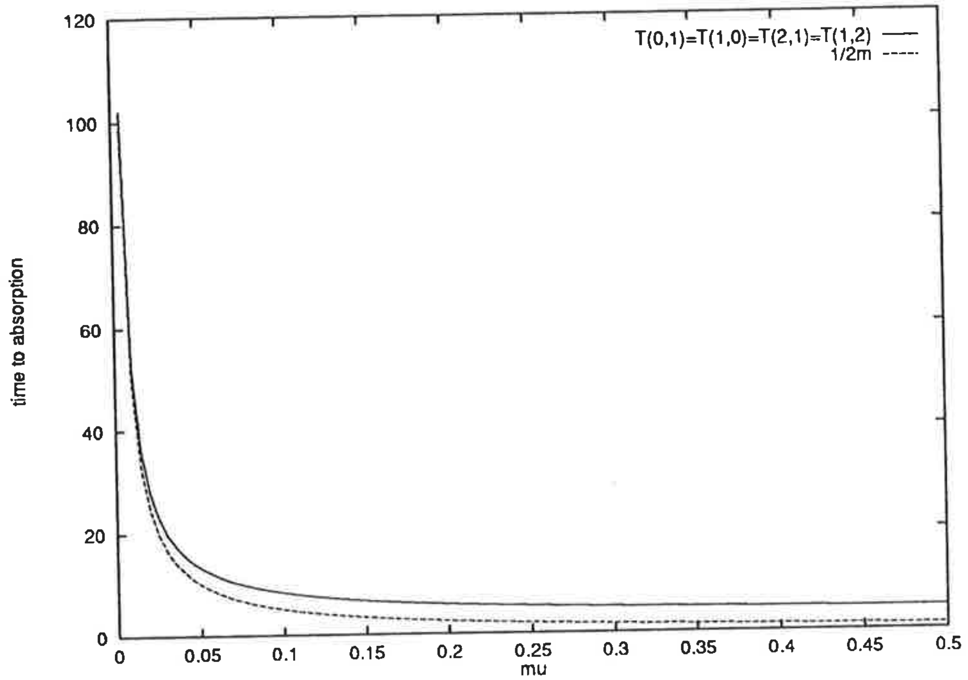


Figure 4.2: Plot of absorption time and approximate absorption time for one set of initial states.

4.4.2 Some numerical results - equal migration

Numerical results were obtained for the case when $N_1 = N_2 = 4$, $\mu_{12} = \mu_{21} = \mu$. They are displayed in figure 4.3. The form of the results generally agree with the results obtained analytically above for the case $N_1 = N_2 = 2$. When μ is larger than about 0.1, the mean time to absorption remains fairly constant. As μ gets closer to 0, the expected time to absorption rapidly increases. This also shows that the mean absorption time is dependent on the magnitude of the migration between the two subpopulations and not just the ratio of the two migration coefficients.

Another consequence of μ decreasing is the much more dramatic difference between absorption times for states with the same frequency in the whole population but different frequencies in the individual subpopulations. For example, the states (0,4) and (2,2) have the same initial frequency in the whole population but different frequencies in the two individual subpopulations. When μ is 0.99 the expected absorption times for these states are 11.8 and 10.5 respectively. When μ is 0.05 the expected absorption times are 48.3 and 28.8 respectively. The absorption times are not only larger, but are also much further apart than when $\mu = 0.99$. This is because when the population starts in the initial state (0,4) must wait for a migrant between the two subpopulations before absorption in the population as a whole can occur. When the population starts in the initial state (2,2) the population can go to absorption fairly quickly if both subpopulations drift to fixation of the same allele. If the two subpopulations drift to fixation of different alleles then they must wait for a migrant before fixation in the population as a whole can occur, but this will only happen half of the time rather than all of the time.

This is another example of how Slatkin's low migration limit is approximately correct. For a subpopulation of size 4 the mean absorption time for an allele when a new migrant arrives is about 3 generations. So if the average time between migration events is much larger than 3 generations (that is $\mu \ll 0.3$) then the low migration limit is a fair approximation to the absorption time.

4.4.3 Some numerical results - unequal migration

In the case of unequal migration between the two subpopulations, the situation is more complicated. No analytic results were obtained, but the following numerical results were obtained for the same population structure as the previous example with different migration parameters. Figures 4.4, 4.5 and 4.6 show three different values of μ_{21} (0.1, 0.5 and 0.9 respectively) and allow μ_{12} to vary between 0.01 and 0.99. The subpopulation sizes are $N_1 = N_2 = 4$. There are three main conclusions from these figures.

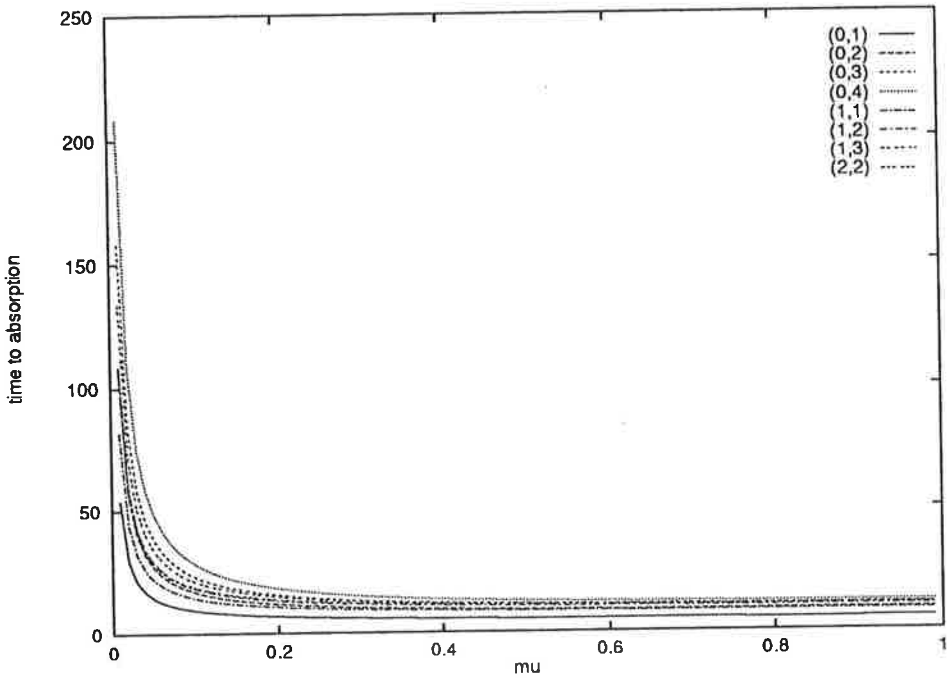


Figure 4.3: Plot of absorption time as a function of migration probability for the eight different classes of initial state when migration is symmetric.

- When there is a source subpopulation and a sink subpopulation the mean absorption time is largely determined by the mean absorption time in the source population. This is because once the source is fixed for a particular allele, the sink will soon be fixed for the same allele. Figure 4.6 shows that the mean absorption times for a source/sink population structure are much lower than those for strong symmetric migration between the two subpopulations. For the source/sink population structure the mean absorption time is often less than the mean absorption time for the corresponding single population ($N = 8$) except when the sink population is initially fixed for one of the alleles. This shows that the mean absorption time of the source subpopulation primarily determines the mean absorption time for the whole population.

- When there is reasonably strong and approximately equal migration between the two populations the mean absorption time is small and relatively constant as the level of migration varies slightly. The mean absorption time is generally between 4 and 11 generations (figure 4.5). The mean absorption times are approximately equal to, but higher than, the mean absorption time for the corresponding single population case ($N = 8$). The bump in figure 4.5 is because when $\mu_{21} = 0.5$ and μ_{12} is close to 0 or 1, there is a slight source/sink effect which tends to drive the population more quickly towards absorption. When $\mu_{12} \approx 0.5$ there is no source and sink effect so random fluctuations will tend to slightly increase the mean time until absorption.

- Figure 4.4 shows that when migration is low in both directions the mean absorption time is still much higher than when there is strong migration in one or both directions. This is essentially the same effect which was observed in section 4.4.1 when migration between the subpopulations is equal and at a low level, and is another example of Slatkin's (1981) low migration limit being a good approximation to the exact solution.

These points demonstrate well that there are two main factors which influence the mean absorption time. These are migration rates and genetic stochasticity. Absorption will often not occur until a migrant passes between the two subpopulations (half the time they will fix for different alleles). When migration is low in both directions this can take a considerable time and so the waiting time for a migrant will determine the mean time to absorption or fixation. When migration is higher, migrants will transfer much more often and the major factor determining the time to fixation is stochasticity, that is genetic drift. When migration is moderate to high, the mean absorption time is generally similar to a single population of size $N_1 + N_2$.

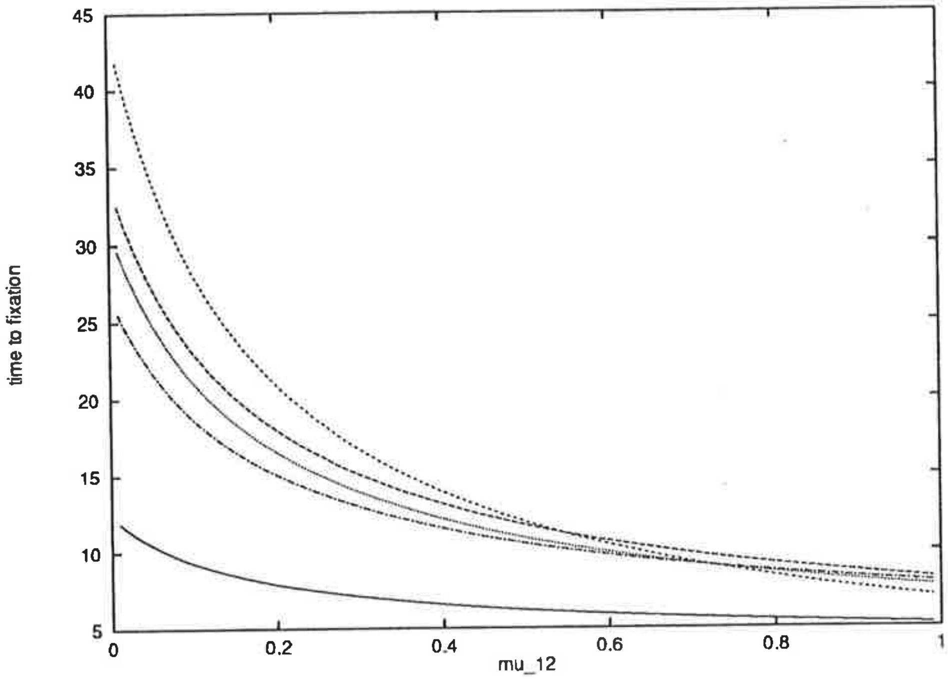


Figure 4.4: Plot of mean absorption time as a function of migration probability for some of the different classes of initial state when $\mu_{21} = 0.1$.

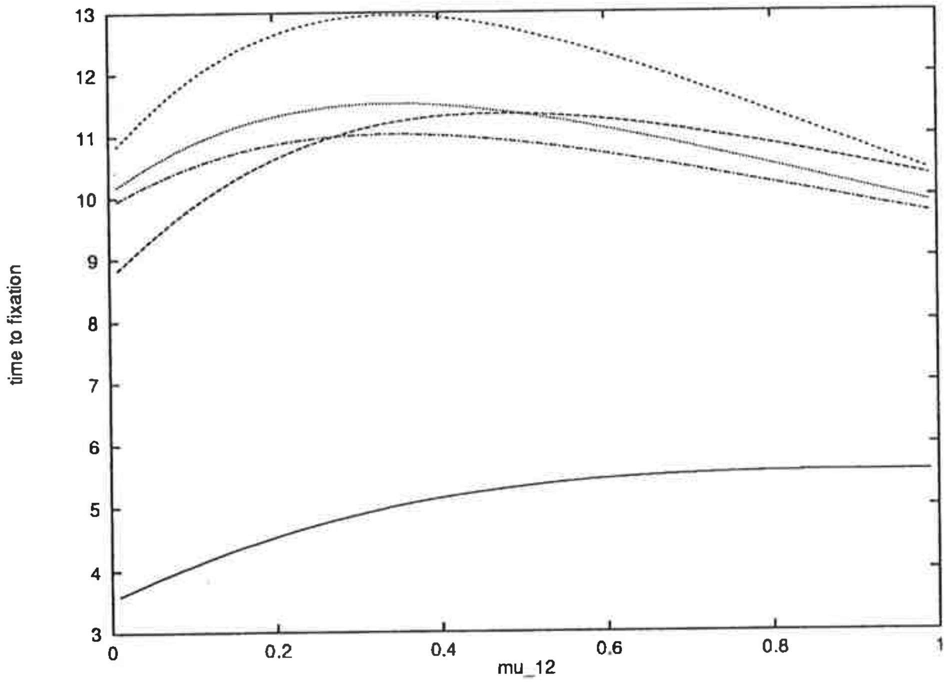


Figure 4.5: Plot of mean absorption time as a function of migration probability for some of the different classes of initial state when $\mu_{21} = 0.5$.

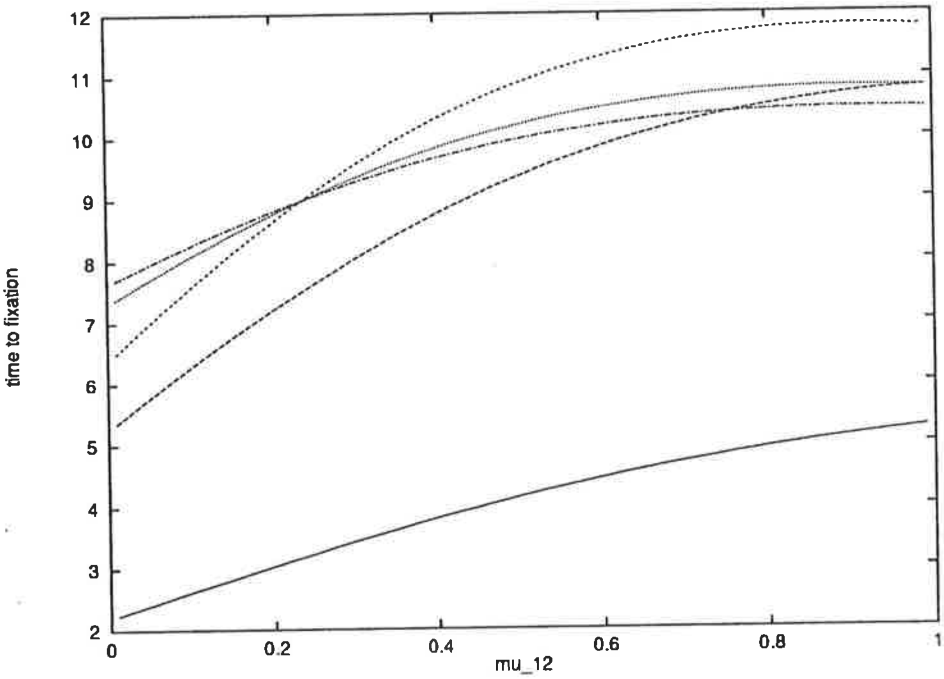


Figure 4.6: Plot of absorption time as a function of migration probability for some of the different classes of initial state when $\mu_{21} = 0.9$.

4.5 Calculation of mean fixation time

In order to calculate the mean fixation time for the A_1 allele, the stochastic process used in the first half of this chapter is modified and the same method is applied as for the calculation of the mean absorption time. As we are now only interested in those cases when fixation of the A_1 allele occurs we only consider those events which result in fixation of the A_1 allele. In other words, the process is modified by removing the "loss" absorbing state (0,0). Absorption in this modified process is equivalent to fixation in the original process so the mean absorption time for the modified process is equal to the mean fixation time for the original process. The transition probablities for the modified process are

$$P^{\star}(i,j;k,l) = \frac{P(i,j;k,l)}{1-P(i,j;0,0)}.$$

and define $T^{\star}_{(i,j)}$ as the expected time to fixation from state (i,j).

Now arguing analogously with the derivation of the mean absorption time in section 4.3, let $\mathcal{A}^{\star}$ be the set of absorbing states for the case under consideration – in this case $\mathcal{A}^{\star} = \{(N_1, N_2)\}$. The following relationship can be established between the $T^{\star}_{(i,j)}$'s:

$$\begin{aligned} T^{\star}_{(i,j)} &= 1 + \sum_{(k,l)\notin\mathcal{A}^{\star}} P^{\star}(i,j;k,l)T^{\star}_{(k,l)} && \text{for}(i,j) \notin \mathcal{A}^{\star} \\ &= 0 && \text{for}(i,j) \in \mathcal{A}^{\star} \qquad (4.8) \end{aligned}$$

which is a set of $(N_1+1)(N_2+1)-2$ linear equations in $(N_1+1)(N_2+1)-2$ unknowns. The difference between this case and the absorption times in equation 4.1 is that here there is only one absorbing state, (N_1, N_2).

The equations can be rewritten in matrix form with $\mathbf{T}^{\star}$ = the vector of mean times to fixation from each initial state in the following format.

$$\mathbf{T}^\star = \begin{pmatrix} T^\star_{(0,1)} \\ T^\star_{(0,2)} \\ \vdots \\ T^\star_{(0,N_2)} \\ T^\star_{(1,0)} \\ T^\star_{(1,1)} \\ \vdots \\ T^\star_{(N_1,N_2-1)} \end{pmatrix} \tag{4.9}$$

And $\mathbf{P}^\star$ is the transition matrix,

$$\mathbf{P}^\star = \begin{pmatrix} P^\star_{(0,1)\to(0,1)} & P^\star_{(0,1)\to(0,2)} & \cdots & P^\star_{(0,1)\to(N_1,N_2-1)} \\ P^\star_{(0,2)\to(0,1)} & P^\star_{(0,2)\to(0,2)} & \cdots & P^\star_{(0,2)\to(N_1,N_2-1)} \\ \vdots & \vdots & & \vdots \\ P^\star_{(N_1,N_2-1)\to(0,1)} & P^\star_{(N_1,N_2-1)\to(0,2)} & \cdots & P^\star_{(N_1,N_2-1)\to(N_1,N_2-1)} \end{pmatrix} \tag{4.10}$$

Once again, the order that the states are listed in is not important provided the same order is always used.

These matrices satisfy the equation

$$\mathbf{T}^\star = (\mathbf{I} - \mathbf{P}^\star)^{-1}\mathbf{1}. \tag{4.11}$$

As with equation 4.5 for mean absorption time, it will generally be extremely difficult to solve equation 4.11 analytically to obtain a formula for the mean fixation time as it involves inverting a $(N_1+1)(N_2+1)-2$ by $(N_1+1)(N_2+1)-2$ square matrix. The formula can be evaluated numerically for fixed values of each of the parameters, even with N_1 and N_2 reasonably large.

4.5.1 An example with an analytic solution

In this example I have evaluated the formula for the mean fixation time given in equation 4.11 for the case where $N_1 = N_2 = 2$ and $\mu_{12} = \mu_{21} = \mu$. As in the case of the mean absorption time, the formulae are quite complicated even in this simple population structure. The formulae for the mean fixation times from each initial state are as follows:

$$\begin{aligned}
T_{(0,1)} = T_{(1,0)} &= \frac{50\mu^5 - 155\mu^4 - 134\mu^3 + 291\mu^2 + 272\mu + 44}{48\mu^5 - 87\mu^4 - 54\mu^3 + 95\mu^2 + 22\mu} \\
T_{(0,2)} = T_{(2,0)} &= \frac{-\mu^6 - 170\mu^5 + 202\mu^4 + 560\mu^3 - 321\mu^2 - 502\mu - 88}{48\mu^6 - 183\mu^5 + 120\mu^4 + 203\mu^3 - 168\mu^2 - 44\mu} \\
T_{(1,2)} = T_{(2,1)} &= \frac{102\mu^6 - 479\mu^5 + 340\mu^4 + 799\mu^3 + 546\mu^2 - 448\mu - 64}{48\mu^6 - 183\mu^5 + 120\mu^4 + 203\mu^3 - 168\mu^2 - 44\mu} \\
T_{(1,1)} &= \frac{51\mu^6 - 291\mu^5 + 208\mu^4 + 597\mu^3 - 401\mu^2 - 316\mu - 44}{48\mu^6 - 183\mu^5 + 120\mu^4 + 203\mu^3 - 168\mu^2 - 44\mu}
\end{aligned} \tag{4.12}$$

Each of these is plotted in figure 4.7. As in the case of the mean absorption time, the general shape of each graph is fairly flat over most of the range of μ between 0.1 and 1. As μ drops below 0.1, the expected time to fixation rises rapidly, with limit ∞ as μ approaches 0. The reason for this rapid climb for low values of μ is that when migration is very low it is possible that the different subpopulations will be fixed for different alleles for a long time before a migrant transfers between the populations. As fixation in the population as a whole can not occur until this happens, the mean fixation time for the whole population increases as the level of migration between the subpopulations decreases (approximately with $1/\mu$). For all values of the migration parameter, μ, the mean fixation time is noticeably lower when one subpopulation is already fixed for the A_1 allele and the A_1 allele is also present in the other subpopulation. In this case, if the heterozygous subpopulation drifts to fixation of the A_1 allele then fixation in the whole population will be very fast. If the heterozygous subpopulation drifts to fixation

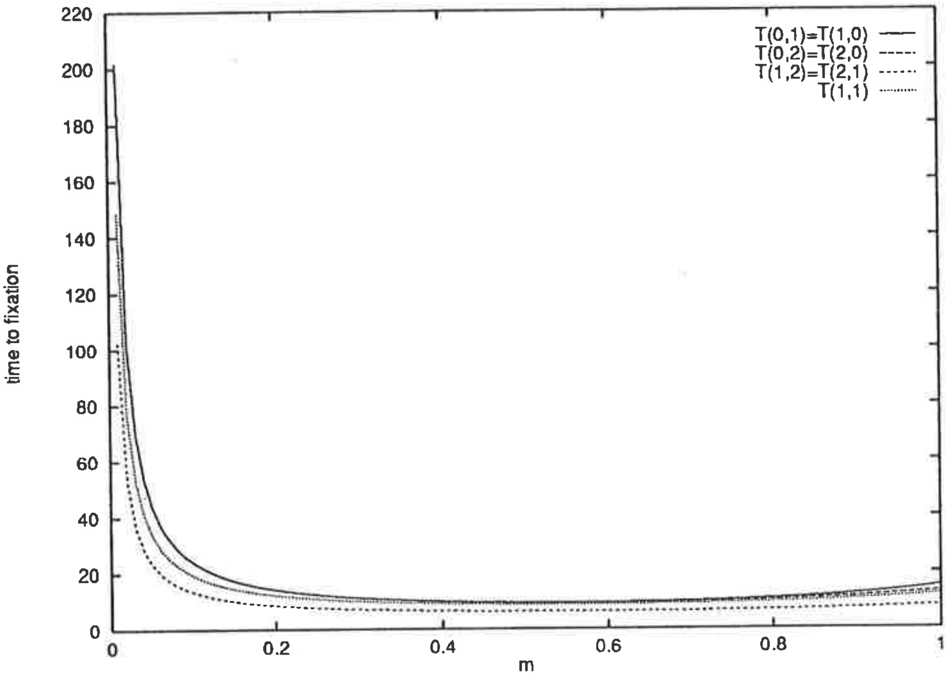


Figure 4.7: Plot of mean fixation time as a function of migration probability for the four different classes of initial state.

of the A_2 allele then fixation will take longer, especially when the level of migration is low.

If the A_1 allele is initially absent from one subpopulation then its initial frequency in the other subpopulation has no effect on mean fixation time for small values of μ. However, when $\mu > 0.8$ then the mean fixation time is higher if there are less A_1 alleles present in the other subpopulation.

4.6 Conclusions

One important conclusion drawn from examining the mean absorption and mean fixation times is that qualitatively the mean absorption and the mean fixation time show very similar behaviour. There are some general rules that apply for all of the populations considered in this chapter for both mean absorption and mean fixation times.

There are three separate cases to be considered: a) Strong and relatively equal migration both ways between the two populations, b) source and sink population structure (asymmetric migration), and c) weak migration between the two populations.

- When migration is strong and relatively equal between the two subpopulations, that is $\mu_{12} \approx \mu_{21} \gg 1/$(mean fixation time of each subpopulation considered separately), the mean absorption time and mean fixation time is relatively independent of the rate of migration.

- When the population has a source/sink structure, the mean absorption and mean fixation times are lower than for strong, equal migration. This is because the source subpopulation tends to dominate the behaviour of the whole population and once it is fixed for a given allele, the sink subpopulation will soon also be fixed for the same allele.

- When migration is weak in both directions, the mean absorption and mean fixation times are higher than in both of the above situations. This is because each subpopulation can be fixed for different alleles for a long time before an allele passes between the two subpopulations and absorption or fixation in the population as whole can not occur until this happens. This idea is analogous to Slatkin's low migration limit (Slatkin, 1981) where mean fixation time is proportional to $1/\mu$ when $\mu \ll 1/$(mean fixation time of each subpopulation separately).

- When the level of migration is low, the absolute level of migration is important in determining both the mean absorption and the mean fixation time. This effect is an example of Slatkin's (1981) low migration limit. As the level of migration decreases, stochasticity becomes more important in determining the mean times to absorption and fixation.

In chapter 5 the results obtained in chapters 3 and 4 are extended to three and four subpopulations using the same methods as before. This is then further extended in chapter 6 where an analogous but mathematically more complicated approach using matrices is used to derive the fixation probability for a general number, n, of subpopulations.

Chapter 5

Models for three and four subpopulations

5.1 Introduction

In this chapter I extend the results of chapters 3 and 4 to the case of three and four subpopulations. It is generally algebraically more complicated to obtain analytic results when there are more than two subpopulations and the solutions become more complicated with the addition of each extra subpopulation. Where analytic results are obtained, an exact interpretation for all situations is often difficult as there are a large number of parameters and special cases must be considered to understand the effect that population structure has on the fixation probabilities. However there are some useful general results and interpretations that can be obtained from the three and four subpopulation models for some more general population structures than was possible with the two subpopulation model. Specific examples of the effect of different population structures on fixation probability are also considered using these models. The interpretation of the three and four subpopulation models given here does not cover every possible structure that a subdivided population can have, however, they show some useful qualitative results which could not be demonstrated using the two subpopulation model.

I do not specifically consider the mean fixation or mean absorption time in this chapter. In principle it is possible to calculate each of these but it is very difficult to find an analytic solution for the mean fixation or mean absorption times when there are three or four subpopulations. It is possible to numerically find a solution for any given set of parameters although I do not do that here. The question of mean absorption and mean fixation times will be considered further in chapters 7 and 8 for populations with selection.

The models that I consider here are generally comparable with the results of Maruyama (1970a; 1970b; 1972; 1974) and the results of Slatkin (1981). There is no direct comparison made here although the conclusions made regarding the comparisons for the two subpopulation model in chapters 3 and 4 still hold. That is, the probability of fixation is independent of population structure when migration is symmetric and there is no selection. Although mena absorption times are not examined here, it can be shown that the low migration limit proposed by Slatkin (1981) is still a useful approximation to the mean absorption time when the level of migration is low.

5.2 Description of model for three subpopulations

This model is an extension to the model used for two subpopulations. The mathematical model used is a three subpopulation version of the model described in chapter 2. There is no selection. The three subpopulations are of constant size N_1, N_2 and N_3. Migration is allowed between the three subpopulations and the mean number of successful migrants from subpopulation i to subpopulation j in a given generation is μ_{ij}.

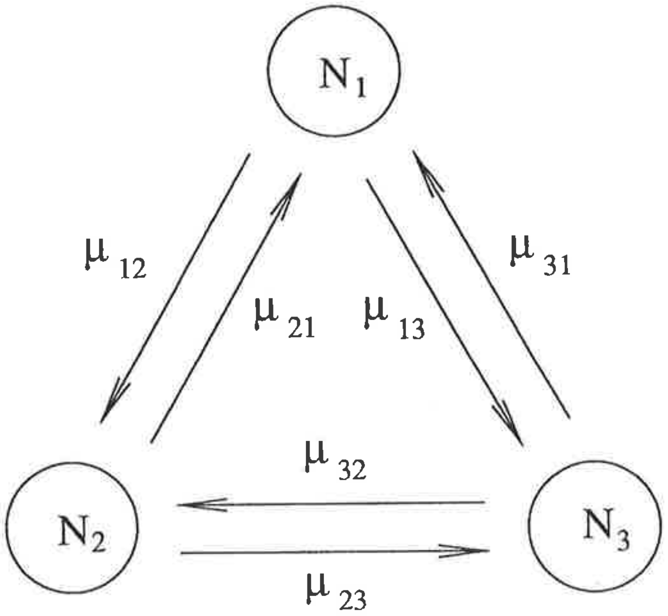


Figure 5.1: Diagramatic representation of the model

The variable of interest is the number of alleles of type A_1 present in each subpopulation at a given time, t, denoted (X_t, Y_t, Z_t) where X_t is the number present in subpopulation 1, Y_t is the number present in subpopulation 2 and Z_t is the number present in subpopulation 3. I also assume that $X_{t+1}|X_t, Y_t, Z_t$, $Y_{t+1}|X_t, Y_t, Z_t$ and $X_{t+1}|X_t, Y_t, Z_t$ are independent of each other. That is, sampling of the juveniles occurs independently in each of the subpopulations

5.2.1 Transition probabilities

The transition probability for a given subpopulation can be found by considering the ways in which that subpopulation can have k individuals with the A_1 allele at time $t+1$. The number of successful migrants from subpopulation i to subpopulation j in a given generation is a random variable, $M_{ij}(t)$. The chance of a migrant from subpopulation i to subpopulation j is μ_{ij} and as the number of migrants is either 0 or 1, it has a Bernoulli distribution with mean μ_{ij}. The number of these migrants which have the A_1 allele is denoted by the variable $M'_{ij}(t)$. There is a maximum of one successful migrant from a given population to any other given population in a generation. The observed number of successful migrants from population i to population j in a particular generation is denoted by $m_{ij}(t)$ and the observed number of successful migrants with the A_1 allele from population i to population j in a particular generation is denoted by $m'_{ij}(t)$. The transition probability for subpopulation 1 can then be written as follows:

$$\begin{aligned} P(X_{t+1} = k | X_t = i, Y_t = j, Z_t = k) \;=\; & \sum_{m'_{21}, m'_{31}} P(m'_{21}(t+1) \; A_1 \text{ alleles from pop 2}, \\ & m'_{31}(t+1) \; A_1 \text{ alleles from pop 3}, \\ & k - m'_{21}(t+1) - m'_{31}(t+1) \; A_1 \text{ alleles from pop 1}) \end{aligned}$$

and an analogous formula applies for Y_{t+1} and Z_{t+1}.

Because of the assumed independence,

$$\begin{aligned} P(X_{t+1}=i, Y_{t+1}=j, Z_{t+1}=k | X_t=x_t, Y_t=y_t, Z_t=z_t) = P(X_{t+1}=i | X_t=x_t, Y_t=y_t, Z_t=z_t) \\ \times P(Y_{t+1}=j | X_t=x_t, Y_t=y_t, Z_t=z_t) \times P(Z_{t+1}=k | X_t=x_t, Y_t=y_t, Z_t=z_t) \qquad (5.1) \end{aligned}$$

5.3 Fixation probability

Theorem 2 *The probability of fixation of an allele, when the initial number of individuals possessing the allele present is x_0 in subpopulation 1, y_0 in subpopulation 2 and z_0 in subpopulation 3 is*

$$\alpha(x_0, y_0, z_0) = \frac{\gamma_1 x_0 + \gamma_2 y_0 + \gamma_3 z_0}{\gamma_1 N_1 + \gamma_2 N_2 + \gamma_3 N_3} \qquad (5.2)$$

where

$$\gamma_1 = \mu_{13}\mu_{32} + \mu_{13}\mu_{12} + \mu_{12}\mu_{23}$$

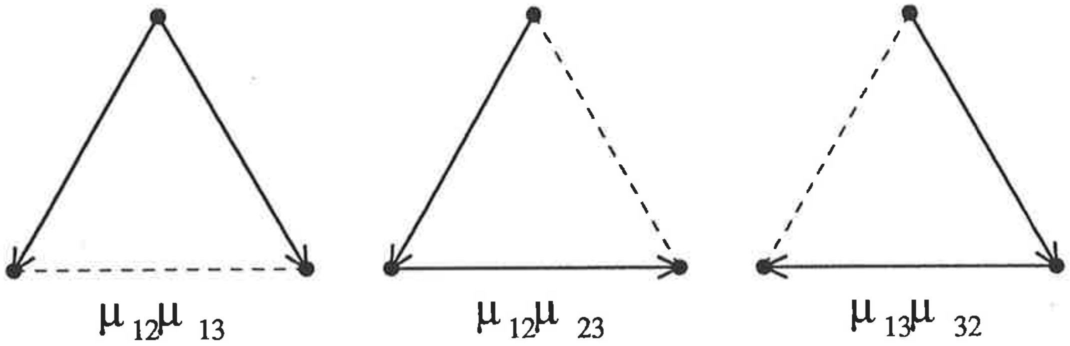


Figure 5.2: The terms that make γ_1 correspond to these three diagrams.

$$\begin{aligned} \gamma_2 &= \mu_{21}\mu_{13} + \mu_{23}\mu_{21} + \mu_{23}\mu_{31} \\ \gamma_3 &= \mu_{32}\mu_{21} + \mu_{32}\mu_{31} + \mu_{31}\mu_{12} \end{aligned} \tag{5.3}$$

The proof of this theorem is exactly analogous to the proof in the case of two subpopulations and is not given here.

The coefficient γ_1 can be considered as the sum of a set of three terms, each of which corresponds to one of the diagrams in figure 5.2. Each term is the product of the probabilities associated with the transitions in that diagram. Each of the diagrams can be thought of as representing a *path* consisting of two transitions which leads from the point of interest to each other point, either directly or indirectly, exactly once. A term for each of the possible paths of this type is included in the sum which makes γ_1. The reason for thinking of the coefficient in this way will become clearer when the more complicated case $n \geq 4$ is considered. It can also be useful in calculating the coefficients when the population structure is altered, such as when one of the links joining two subpopulations is removed.

5.3.1 Example of calculation of ceofficients

As an example of the application of formula 5.2 to calculate the coefficients, γ_1, γ_2 and γ_3 consider the case for three subpopulations arranged with the minimal number of

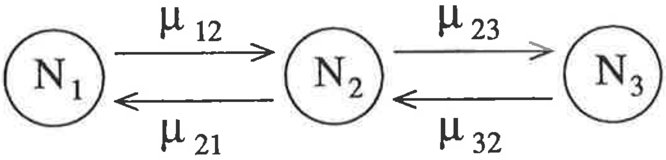


Figure 5.3: Three subpopulations arranged with the minimal number of connections.

connections, for example in a straight line, as in figure 5.3. Now to calculate γ_1 the only path that goes from subpopulation 1 to each other subpopulation under the modified structure is the path to subpopulation 2 then from subpopulation 2 to subpopulation 3 as illustrated below.

So the coefficient for subpopulation 1 is $\gamma_1 = \mu_{12}\mu_{23}$. By looking at the paths for the other two subpopulations it can be seen that the coefficients are $\gamma_2 = \mu_{21}\mu_{23}$ and $\gamma_3 = \mu_{32}\mu_{21}$.

5.3.2 Interpreting the fixation probability

Define the quantities $f_1 = x_0/N_1$, $f_2 = y_0/N_2$ and $f_3 = z_0/N_3$ – the initial frequencies of the A_1 allele in each subpopulation. Taking the formula in equation 5.2 and substituting in the f_i's, the probability of fixation is

$$\alpha(f_1, f_2, f_3) = \frac{\gamma_1 f_1 N_1 + \gamma_2 f_2 N_2 + \gamma_3 f_3 N_3}{\gamma_1 N_1 + \gamma_2 N_2 + \gamma_3 N_3} \tag{5.4}$$

Equation 5.4 shows that the fixation probability is in fact the weighted average of the initial frequencies of allele A_1 in each of the subpopulations, weighted by the quantity $\gamma_i N_i$. So, when subpopulation sizes are equal, a subpopulation with a larger γ has a greater influence on the fixation probability than a subpopulation with a smaller γ. In general, the greater the value of $\gamma_i N_i$ for a subpopulation, the greater the influence it will have on the fixation probability.

In the case of no selection being considered here, the f_i's are also equal to the fixation probability in each subpopulation considered separately. So, the fixation probability in the population as a whole can also be thought of as the weighted sum of the fixation probabilities in each of the subpoulations considered separately. This observation appears to have potential use in finding approximations in the case where there is selection. This will be investigated in Chapter 7.

5.4 Some examples

Example 1: Equal migration between subpopulations when all subpopulations are connected.

The population structure used here is the structure in figure 5.1 with equal migration, so $\mu_{ij} = \mu \ \forall i, j$. This means that each path contributes μ^2 to the coefficient. There are three paths for each coefficient so

$$\gamma_1 = \gamma_2 = \gamma_3 = 3\mu^2$$

and hence

$$\alpha(x_0, y_0, z_0) = \frac{x_0 + y_0 + z_0}{N_1 + N_2 + N_3}$$

which is the overall initial frequency in the whole population.

Example 2: Equal migration between subpopulations in the model with the minimal number of connections.

In the case of equal migration in the model with the minimal number of connections in figure 5.3, $\mu_{13} = \mu_{31} = 0, \mu_{ij} = \mu$ for all other migration parameters, μ_{ij}. This means that each path contributes μ^2 to the coefficient and there is one path for each coefficient so

$$\gamma_1 = \gamma_2 = \gamma_3 = \mu^2$$

and

$$\alpha(x_0, y_0, z_0) = \frac{x_0 + y_0 + z_0}{N_1 + N_2 + N_3}$$

which once again is the overall initial frequency in the whole population.

These two examples show that if there is equal symmetric migration, that is $\mu_{ij} = \mu \forall i, j$, between the connected subpopulations then the fixation probability is independent of the population structure and is equal to the initial frequency in the population considered as a whole.

Example 3: Symmetric migration between subpopulations with all subpopulations connected.

In this example I will show that even when the migration is symmetric but not equal between pairs of subpopulations, that is $\mu_{ij} = \mu_{ji} \ \forall \ i, j$ the fixation probability of the A_1 allele is still equal to the initial frequency of the allele in the whole population.

Let $\mu_{ij} = \mu_{ji} \ \forall i, j$ then γ_1 can be rearranged using $\mu_{ij} = \mu_{ji}$ as follows:

$$\begin{aligned}
\gamma_1 &= \mu_{12}\mu_{13} + \mu_{12}\mu_{23} + \mu_{13}\mu_{32} \\
&= \mu_{21}\mu_{13} + \mu_{21}\mu_{23} + \mu_{23}\mu_{31} = \gamma_2 \\
&= \mu_{31}\mu_{12} + \mu_{32}\mu_{21} + \mu_{32}\mu_{31} = \gamma_3
\end{aligned}$$

so,

$$\alpha(x_0, y_0, z_0) = \frac{x_0 + y_0 + z_0}{N_1 + N_2 + N_3}$$

Once again all of the coefficients are equal and the fixation probability is equal to the initial frequency in the whole population. This is also true for the model with the minimal number of connections with symmetric migration as $\mu_{13} = \mu_{31} = 0$ is just a special case of the formula considered above.

Example 4: Two subpopulations close together, third subpopulation further away.

In a situation where there are two subpopulations clustered close together with a third subpopulation further away, the level of migration between the two subpopulations which are close together will generally be higher than the level of migration between either of these two and the distant subpopulation. This situation occurs naturally where there is a single remnant subpopulation separated from the rest of the subpopulations – (see for example, Templeton *et al.*, 1990; Meffe and Vrijenhoek, 1988). Suppose that three subpopulations are arranged with the minimal number of connections as in figure 5.3 and migration between subpopulations 1 and 2 is twice the level of migration between subpopulations 2 and 3. That is, $\mu_{13} = \mu_{31} = 0$, $\mu_{23} = \mu_{32} = \mu$ and $\mu_{12} = \mu_{21} = 2\mu$. In this situation, the coefficents are $\gamma_1 = \gamma_2 = \gamma_3 = 2\mu^2$. So,

$$\alpha(x_0, y_0, z_0) = \frac{x_0 + y_0 + z_0}{N_1 + N_2 + N_3}$$

This means that each subpopulation is equally important in determining the fixation probabilities. Initially this seems counter intuitive but the model specifies that there are as many migrants going from the cluster to the distant subpopulation as from the distant subpopulation to the cluster. So although the distant subpopulation has a weak effect on the cluster, the cluster also has a weak effect on the distant subpopulation. If the migration parameters are changed so that there are more migrants going from the cluster to the distant subpopulation than from the distant subpopulation to the

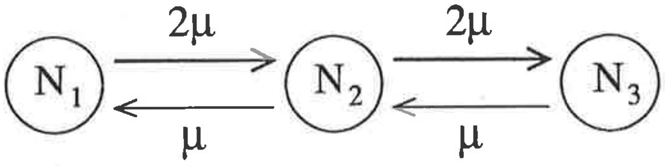


Figure 5.4: Directional flow in the model with the minimal number of connections.

cluster (possibly a more realistic assumption in some situations) then this would no longer be true.

This model is also appropriate for three subpopulations being equally spaced but two of the subpopulations being better connected than the third. This would be the case, for example, if there was a habitat corridor between two subpopulations but not between these two subpopulations and the third subpopulation.

Example 5: Directional flow in the model with the minimal number of connections.

Suppose now that the model with the minimal number of connections is considered with a constant directional flow. For example, this could be thought of as a current or a prevailing wind blowing more seeds in one direction than are blown in the other direction. In the example illustrated in figure 5.4, the transition probabilities are $\mu_{13} = \mu_{31} = 0$, $\mu_{32} = \mu_{21} = \mu$ and $\mu_{12} = \mu_{23} = 2\mu$. Upon substituting these transition probabilities into the equations for the coefficients, we get $\gamma_1 = 4\mu^2$, $\gamma_2 = 2\mu^2$ and $\gamma_3 = \mu^2$. So,

$$\alpha(x_0, y_0, z_0) = \frac{4x_0 + 2y_0 + z_0}{4N_1 + 2N_2 + N_3}$$

This is fairly intuitive as it suggests that if the flow is from left to right in figure 5.4, then the further left a subpopulation is, the further "upstream" it is and hence the more important it is in determining the fixation probability.

Note that in the similar case where $\mu_{12} = \mu_{23} = 2\mu$ and $\mu_{21} = \mu 32 = 0$, that the

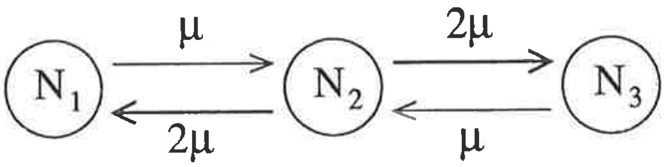


Figure 5.5: Central subpopulation is most productive.

fixation probability is simply x_0/N_1. This is because there is no flow into subpopulation 1, so once subpopulation 1 has become fixed (A_1 with probability x_0/N_1) the other subpopulations must eventually both become fixed for the same allele.

Example 6: Central subpopulation most productive.

In this example, the central subpopulation in the model with the minimal number of connections is the most productive. For example, this could be thought of as the central subpopulation being in the most productive habitat, the "core" habitat, and the outlying subpopulations being in less favourable habitat and thus producing less migrants. There are many instances of this in natural systems – see, for example (Downes *et al.*, 1997; Lamberson *et al.*, 1992). The example is illustrated in figure 5.5 and has transition probabilities $\mu_{13} = \mu_{31} = 0$, $\mu_{32} = \mu_{12} = \mu$ and $\mu_{21} = \mu_{23} = 2\mu$. Substitution of these transition probabilities into the formulae for the coefficients gives, $\gamma_1 = 2\mu^2$, $\gamma_2 = 4\mu^2$ and $\gamma_3 = 2\mu^2$. so,

$$\alpha(x_0, y_0, z_0) = \frac{x_0 + 2y_0 + z_0}{N_1 + 2N_2 + N_3}.$$

Once again this result is fairly intuitive in showing that when the central subpopulation is most productive, it will have the greatest influence on the fixation probability. It can be shown that no matter which subpopulation is the source subpopulation, it will have twice the effect that the other two subpopulations have if it produces twice the number of migrants that each of the other two subpopulations produce.

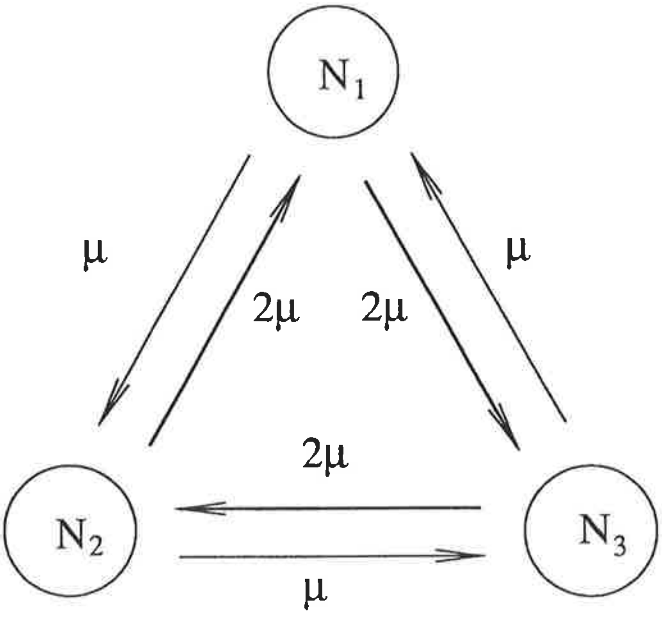


Figure 5.6: An example of asymmetric migration with no net migration.

The main conclusion which can be drawn from examples 5 and 6 is that the subpopulations which produce the most migrants have the greatest effect on the fixation probability. This is true regardless of the position of a subpopulation relative to other subpopulations, that is, edge or centre.

Example 7: Asymmetric migration with no net migration.

In this example there is no net migration, that is, each subpopulation sends out as many migrants as it receives alhtough there is no requirement that this is true between each pair of subpopulations. This means that

$$\begin{aligned}\mu_{12} + \mu_{13} &= \mu_{21} + \mu_{31} \\ \mu_{21} + \mu_{23} &= \mu_{12} + \mu_{32} \\ \mu_{31} + \mu_{32} &= \mu_{13} + \mu_{23}. \end{aligned} \tag{5.5}$$

This includes symmetric migration as a special case. An example of this is shown in figure 5.6 where there is asymmetric migration but more migrants move clockwise than anti-clockwise around a circle of subpopulations in such a way that there is no net flow of migrants. For example, $\mu_{12} = \mu_{23} = \mu_{31} = \mu$ and $\mu_{21} = \mu_{32} = \mu_{13} = 2\mu$.

As a result of this, each of the weights, γ_i, is equal, for example using equation 5.5 it can be shown that

$$\begin{aligned} \gamma_1 &= \mu_{13}\mu_{32} + \mu_{12}\mu_{13} + \mu_{12}\mu_{23} \\ &= \mu_{31}\mu_{23} + \mu_{13}\mu_{21} + \mu_{21}\mu_{23} \\ &= \gamma_2. \end{aligned}$$

and similarly, $\gamma_1 = \gamma_3$.

So when there is no net migration between the subpopulations, the fixation probability is independent of the population structure and is equal to the initial frequency in the population as a whole. This agrees with the results obtained by Maruyama (1970a; 1974) where he implicitly assumes no net migration.

5.5 Mean absorption and fixation times

The basic results which hold for two subpopulations also hold for three subpopulations. The state space is far more complicated and analytic results are possible in theory though extremely messy to calculate in practice.

5.6 Description of the model for four subpopulations

This model is a further extension of the two subpopulation model and the three subpopulation model. The mathematical model used is a four subpopulation version of the model described in Chapter 2. There is no selection. There are four subpopulations of constant size N_1, N_2, N_3 and N_4. Migration is allowed between the four subpopulations and the mean number of successful migrants per generation from subpopulation i to subpopulation j is a random variable. The probability of a successful migrant from subpopulation i to subpopulation j in a generation is μ_{ij}. The maximum number of migrants in a given generation is one so $M_{ij}(t)$, the random variable representing the number of successful migrants from subpopulation i to subpoulation j in generation t, has a Bernoulli distribution with mean μ_{ij}. The variables of interest in the model are the numbers of alleles of type A_1 present in each subpopulation at a given time, t. The variables are denoted W_t, X_t, Y_t and Z_t for the number of A_1 alleles in subpopulations 1, 2, 3 and 4 respectively. The mathematics of the model will not be written out in full for the case of four subpopulations as it is exactly analagous to the case of two and three subpopulations except that the algebra is far more complicated.

5.7 Fixation Probabilities

The fixation probabilities are given by the formula

$$\alpha(w_0, x_0, y_0, z_0) = \frac{\gamma_1 w_0 + \gamma_2 x_0 + \gamma_3 y_0 + \gamma_4 z_0}{\gamma_1 N_1 + \gamma_2 N_2 + \gamma_3 N_3 + \gamma_4 N_4} \tag{5.6}$$

where w_0, x_0, y_0 and z_0 are the initial frequencies in subpopulations 1, 2, 3 and 4 respectively. The migration coefficients γ_i are each made up of sixteen terms. Each

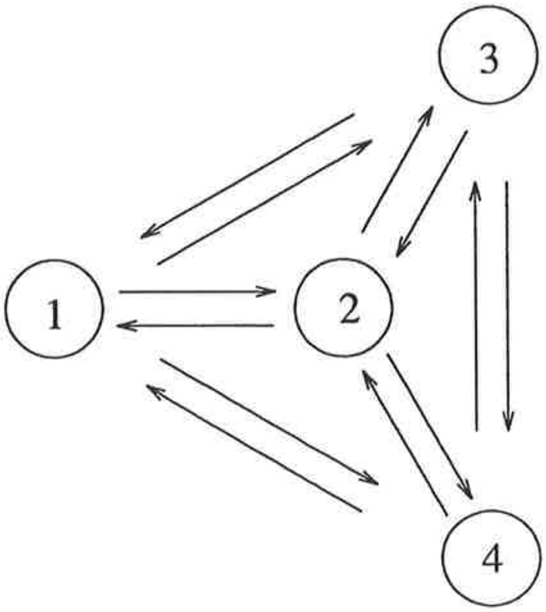


Figure 5.7: Diagramatic representation of the model

of these terms is a product of three migration magnitudes of the form μ_{ij}. As an example, the migration coefficient γ_1 is given along with a geometric interpretation of its component terms given in section 5.7.1.

$$\begin{aligned}\gamma_1 \;=\;& \mu_{12}\mu_{13}\mu_{14} + \mu_{13}\mu_{32}\mu_{24} + \mu_{13}\mu_{34}\mu_{42} + \mu_{12}\mu_{23}\mu_{34} + \mu_{12}\mu_{24}\mu_{43} + \mu_{14}\mu_{43}\mu_{32} \\ & + \mu_{14}\mu_{42}\mu_{23} + \mu_{13}\mu_{12}\mu_{24} + \mu_{13}\mu_{14}\mu_{42} + \mu_{12}\mu_{13}\mu_{34} + \mu_{12}\mu_{14}\mu_{43} + \mu_{14}\mu_{13}\mu_{32} \\ & + \mu_{14}\mu_{12}\mu_{23} + \mu_{13}\mu_{32}\mu_{34} + \mu_{12}\mu_{23}\mu_{24} + \mu_{14}\mu_{42}\mu_{43} \end{aligned} \tag{5.7}$$

The other coefficients, γ_2, γ_3 and γ_4 are calculated and interpreted in the same manner.

5.7.1 Interpreting the fixation probability

As in the case of three subpopulations, the fixation probability can be rewritten as

$$\alpha(f_1, f_2, f_3, f_4) = \frac{\gamma_1 f_1 N_1 + \gamma_2 f_2 N_2 + \gamma_3 f_3 N_3 + \gamma_4 f_4 N_4}{\gamma_1 N_1 + \gamma_2 N_2 + \gamma_3 N_3 + \gamma_4 N_4}$$

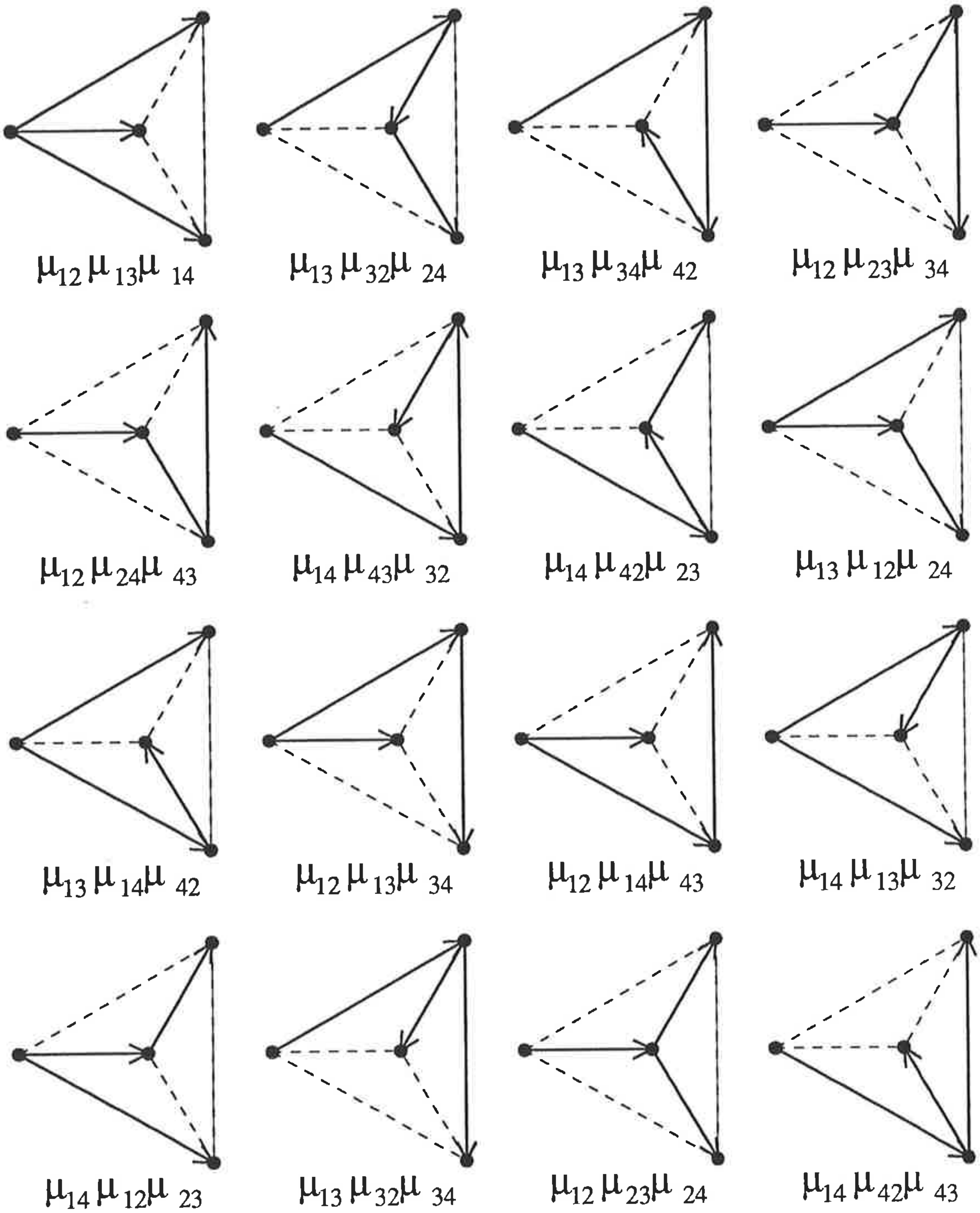
where $f_1 = w_0/N_1$, $f_2 = x_0/N_2$, $f_3 = y_0/N_3$ and $f_4 = z_0/N_4$. This means that, as in the case of three subpoulations, the fixation probability for the whole population can be thought of as the weighted sum of the fixation probabilities in each subpopulation considered separately.

Each migration coefficient has a physical interpretation. The migration coefficient γ_1 can be thought of as being the sum of the terms corresponding to each of the "paths" in figure 5.8. Each path consists of three transitions which starts at the point of interest and goes through each other point either directly or indirectly exactly once. All of the possible paths of this type are included in the sum so there is a term corresponding to each way of getting from the subpopulation of interest to each other subpopulation exactly once. The term corresponding to a given path is equal to the product of the migration strength associated with each of the transitions in that path. This result can be proven using graph theory and this is used to prove the analogous general result for n subpopulations in Chapter 6.

5.8 Different population structures

There are six different ways in which four subpopulations can be connected by links which allow migration (ignoring the strengths of migration). They are represented in figure 5.9. The lines represent links between two subpopulations. The actual physical distance between two adjacent subpopulations in the diagram does not necessarily represent the magnitude of migration between the two subpopulations.

When migration between the subpopulations is equal and symmetric (that is $\mu_{ij} = \mu \ \forall i, j$), the method used in section 5.4 for calculating the migration coefficients, γ_i, under different population structures of counting the number of paths from subpopulation i is a useful way of finding the coefficients quickly. This works here because each



$\mu_{12}\mu_{13}\mu_{14}$ $\mu_{13}\mu_{32}\mu_{24}$ $\mu_{13}\mu_{34}\mu_{42}$ $\mu_{12}\mu_{23}\mu_{34}$

$\mu_{12}\mu_{24}\mu_{43}$ $\mu_{14}\mu_{43}\mu_{32}$ $\mu_{14}\mu_{42}\mu_{23}$ $\mu_{13}\mu_{12}\mu_{24}$

$\mu_{13}\mu_{14}\mu_{42}$ $\mu_{12}\mu_{13}\mu_{34}$ $\mu_{12}\mu_{14}\mu_{43}$ $\mu_{14}\mu_{13}\mu_{32}$

$\mu_{14}\mu_{12}\mu_{23}$ $\mu_{13}\mu_{32}\mu_{34}$ $\mu_{12}\mu_{23}\mu_{24}$ $\mu_{14}\mu_{42}\mu_{43}$

Figure 5.8: These diagrams represent the 16 terms that make up γ_1.

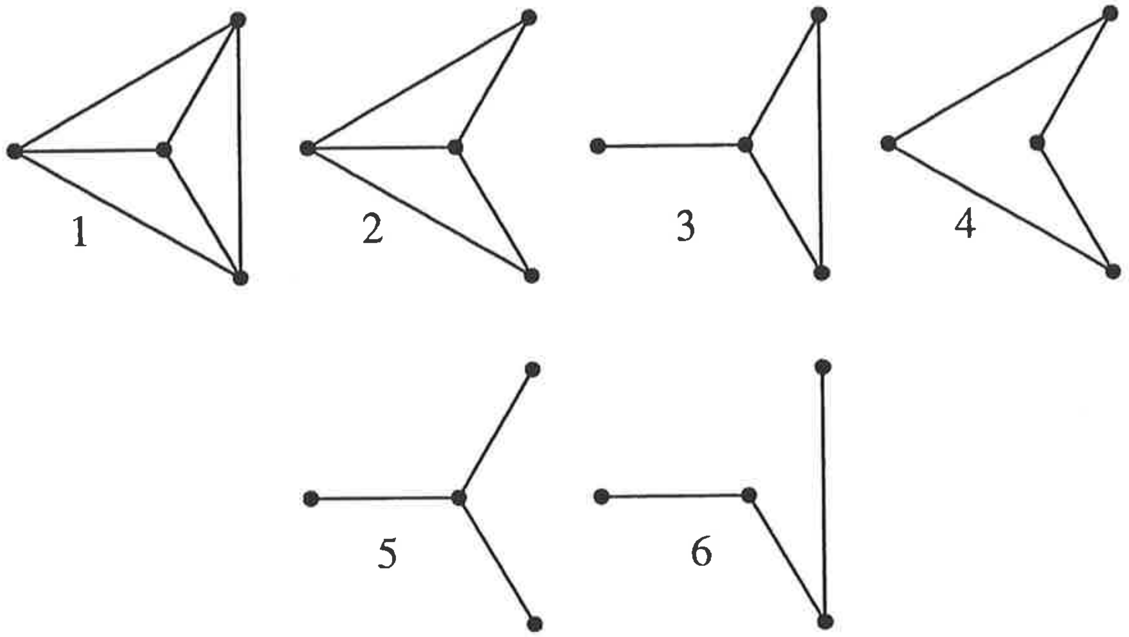


Figure 5.9: The different ways in which four subpopulations can be linked.

path contributes μ^3 to the migration coefficient under consideration.

5.8.1 Fixation probabilities under equal migration

In each of the following cases there is equal and symmetric migration between connected subpopulations, that is $\mu_{ij} = \mu \ \forall i, j$ where a link exists between subpopulations i and j. The numbers refer to the diagram numbers in figure 5.9.

Structure 1: This is the fully connected model where each coefficient consists of the full 16 terms. Each term contributes μ^3 to the coefficient so $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 16\mu^3$. So the fixation probability is equal to the initial frequency in the population as a whole.

Structure 2: In this model two of the subpopulations are connected to all three of the other subpopulations and two of the subpopulations are only connected to two of the other subpopulations. For each subpopulation there are 8 paths connecting it to each of the other subpopulations, each contributing μ^3, so $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 8\mu^3$.

Structure 3: In this model three of the subpopulations are connected to each other and the fourth subpopulation is connected with one of these subpopulations only. Intuitively it would seem that the subpopulation which is connected with only one other should have less influence on the fixation probability than the other subpopulations which are all connected with each other. However, there are three paths from each subpopulation so each coefficient has three terms giving $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 3\mu^3$. This is because each subpopulation has no net flow of migrants – they each recieve as many migrants from other subpopulations as they send out to other subpopulations. The distant subpopulation sends less and receives less migrants than each of the other subpopulations.

Structure 4: This model is equivalent to having four subpopulations equally spaced on the corners of a square. Intuitively it would seem that each subpopulation should have the same effect on the fixation probability as the system is symmetric, and this is true. There are four paths from each subpopulation to the other subpopulations, so $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 4\mu^3$. So each subpopulation has the same effect on the fixation probability as expected.

Structure 5: One subpopulation is central and is connected to each other subpopulation, but each other subpopulation is connected only to the central subpopulation. Once again it would seem that the central subpopulation would have the most influence on the fixation probabilities, but for each subpopulation there is only one path connecting it to all of the other subpopulations, so $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \mu^3$. This means each subpopulation is equally important which seems counterintuitive. As with structure 3, the reason for this is that there is no net flow of migrants to or from any subpopulation. Each subpopulation sends out as many migrants as it receives from other subpopulations

Structure 6: The minimally connected model. The subpopulations are only connected to there nearest neighbours (the minimal number of connections). Once again, for each subpopulation there is only one path connecting it to all of the other subpopulations, so $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \mu^3$. Again each subpopulation is equally important regardless of whether it is an edge or a centre subpopulation. This is because those subpopulations which produce the most migrants (the central subpoulations) also receive the most migrants, and those subpopulations which produce the least migrants (the end subpopulations) also receive the least migrants.

So for each of the population structures, $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4$ when migration is symmetric. This means that for each of the population structures, when migration is equal and symmetric the fixation probability is equal to the initial frequency in the population as a whole. This result at first seems counter intuitive as, for example in population structure 5, the centre subpopulation sends out three times as many migrants as any of the other subpopulations. However, being the centre subpopulation it also recieves three times as many immigrants as any other subpopulation. The important point to note here is that each subpopulation sends as many migrants as it receives and thus has no net flow of migrants. So each subpopulation has an equal effect on the fixation probability. That is, those subpopulations which have a large impact on other subpopulations by producing a lot of migrants are also more readily affected by other subpopulations because they receive a large number of immigrants.

5.8.2 Symmetric migration

In section 5.4 it was shown for the case of three subpopulations that when there is symmetric migration between adjacent subpopulations, the fixation probability is equal to the overall initial frequency in the whole population regardless of the population structure. Now consider the case $\mu_{ij} = \mu_{ji}\ \forall i, j$ with four subpopulations arranged as in structure 1 in figure 5.9. By substituting equivalent terms in the formula for γ_1 it can be seen that $\gamma_1 = \gamma_2$.

$$\begin{aligned}
\gamma_1 &= \mu_{12}\mu_{13}\mu_{14} + \mu_{13}\mu_{32}\mu_{24} + \mu_{13}\mu_{34}\mu_{42} + \mu_{12}\mu_{23}\mu_{34} + \mu_{12}\mu_{24}\mu_{43} + \mu_{14}\mu_{43}\mu_{32} \\
&\quad +\mu_{14}\mu_{42}\mu_{23} + \mu_{13}\mu_{12}\mu_{24} + \mu_{13}\mu_{14}\mu_{42} + \mu_{12}\mu_{13}\mu_{34} + \mu_{12}\mu_{14}\mu_{43} + \mu_{14}\mu_{13}\mu_{32} \\
&\quad +\mu_{14}\mu_{12}\mu_{23} + \mu_{13}\mu_{32}\mu_{34} + \mu_{12}\mu_{23}\mu_{24} + \mu_{14}\mu_{42}\mu_{43} \\
&= \mu_{21}\mu_{13}\mu_{14} + \mu_{31}\mu_{23}\mu_{24} + \mu_{31}\mu_{43}\mu_{24} + \mu_{21}\mu_{23}\mu_{34} + \mu_{21}\mu_{24}\mu_{43} + \mu_{41}\mu_{34}\mu_{23} \\
&\quad +\mu_{14}\mu_{24}\mu_{23} + \mu_{13}\mu_{21}\mu_{24} + \mu_{13}\mu_{41}\mu_{24} + \mu_{21}\mu_{13}\mu_{34} + \mu_{21}\mu_{14}\mu_{43} + \mu_{14}\mu_{31}\mu_{23} \\
&\quad +\mu_{14}\mu_{21}\mu_{23} + \mu_{31}\mu_{23}\mu_{34} + \mu_{21}\mu_{23}\mu_{24} + \mu_{41}\mu_{24}\mu_{43} \\
&= \mu_{21}\mu_{23}\mu_{24} + \mu_{31}\mu_{43}\mu_{24} + \mu_{41}\mu_{34}\mu_{23} + \mu_{13}\mu_{41}\mu_{24} + \mu_{21}\mu_{13}\mu_{34} + \mu_{21}\mu_{14}\mu_{43} \\
&\quad +\mu_{14}\mu_{31}\mu_{23} + \mu_{21}\mu_{23}\mu_{34} + \mu_{14}\mu_{21}\mu_{23} + \mu_{21}\mu_{24}\mu_{43} + \mu_{13}\mu_{21}\mu_{24} + \mu_{31}\mu_{23}\mu_{24} \\
&\quad +\mu_{41}\mu_{24}\mu_{23} + \mu_{41}\mu_{24}\mu_{43} + \mu_{31}\mu_{23}\mu_{34} + \mu_{21}\mu_{13}\mu_{14} \\
&= \gamma_2
\end{aligned}$$

It can be shown analogously that $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4$.

As all of the other population structures are a special case of this situation (for population structures 2 to 6, $\mu_{ij} = \mu_{ji} = 0$ for some of the pairs i, j) this is true for all population structures shown in figure 5.9.

This means that all of the results of subsection 5.8.1 apply here essentially for the same reason – when migration is symmetric there is no net flow of migrants between any two adjacent subpopulations. So all subpopulations which send out migrants receive an equal number of immigrants from other subpopulations and hence the fixation probability is just equal to the initial frequency in the whole population.

5.8.3 No net migration

As in the case of three subpopulations, when there is no net migration into or out of any of the four subpopulations, the probability of fixation of an allele in the whole population is equal to its initial frequency in the whole population. That is, when

$$\begin{aligned} \mu_{12} + \mu_{13} + \mu_{14} &= \mu_{21} + \mu_{31} + \mu_{41} \\ \mu_{21} + \mu_{23} + \mu_{24} &= \mu_{12} + \mu_{32} + \mu_{42} \\ \mu_{31} + \mu_{32} + \mu_{34} &= \mu_{13} + \mu_{32} + \mu_{43} \\ \mu_{41} + \mu_{42} + \mu_{43} &= \mu_{14} + \mu_{24} + \mu_{34} \end{aligned}$$

it can be shown by substitution that the fixation probability is equal to the initial frequency in the whole population.

5.9 Unequal migration between subpopulations

There are endless possible combinations of migration parameters that could be considered with the four subpopulation model. Some important cases will be considered here to compare the qualitative effects with the three subpopulation case.

Example 1: Constant flow of individuals in one direction with a minimally connected population structure.

It is assumed that gene flow is twice as likely in one direction as in the other. So

$$\mu_{12} = \mu_{23} = \mu_{34} = 2\mu$$

$$\mu_{21} = \mu_{32} = \mu_{43} = \mu$$

$$\mu_{13} = \mu_{31} = \mu_{14} = \mu_{41} = \mu_{24} = \mu_{42} = 0$$

Each point has just one path to the other three points, so there is only one term in each coefficient. The coefficients are :

$$\gamma_1 = 8\mu^3, \quad \gamma_2 = 4\mu^3, \quad \gamma_3 = 2\mu^3, \quad \gamma_4 = \mu^3$$

giving the fixation probability as

$$\alpha(w_0, x_0, y_0, z_0) = \frac{8w_0 + 4x_0 + 2y_0 + z_0}{8N_1 + 4N_2 + 2N_3 + N_4} \tag{5.8}$$

So as in the case of three dimensions, when the flow is from left to right, the subpopulations on the left are the most important in determining the fixation probability.

Example 2: The central subpopulation is most productive

Consider a population structured as in diagram 5 in figure 5.9. Once again there is only one path from each of the subpopulations, so each coefficient has only one term. If the migration parameters are:

$$\mu_{21} = \mu_{23} = \mu_{24} = 2\mu$$

$$\mu_{12} = \mu_{32} = \mu_{42} = \mu$$

$$\mu_{13} = \mu_{31} = \mu_{14} = \mu_{41} = \mu_{34} = \mu_{43} = 0$$

So,

$$\gamma_1 = 4\mu^3, \quad \gamma_2 = 8\mu^3, \quad \gamma_3 = 4\mu^3, \quad \gamma_4 = 4\mu^3.$$

giving the fixation probability as

$$\alpha(w_0, x_0, y_0, z_0) = \frac{w_0 + 2x_0 + y_0 + z_0}{N_1 + 2N_2 + N_3 + N_4} \tag{5.9}$$

Showing once again that the subpopulation which produces the most migrants is the most important subpopulation in determining the fixation probability. Although the central subpopulation being the most important is intuitive, the relative magnitudes of the coefficients in this example at first seem counter intuitive. The central subpopulation is sending out twice as many migrants as the rest of the populations put together, yet it only has the weight of each other subpopulation conisdered separately. This is because the outer subpopulations are sending their migrants to the central subpopulation – the most important subpopulation in terms of fixation probability thus increasing the survival chance of these alleles in subsequent generations. The central subpopulation is sending its migrants to the relatively less important edge subpopulations.

Example 3: The central subpopulation is most productive, all subpopulations directly linked.

Now consider a population with structure 1 from figure 5.9. The central subpopulation is twice as productive as the other subpopulations, so

$$\mu_{21} = \mu_{23} = \mu_{24} = 2\mu$$

$$\mu_{12} = \mu_{13} = \mu_{14} = \mu_{31} = \mu_{32} = \mu_{34} = \mu_{41} = \mu_{42} = \mu_{43} = \mu.$$

Then by using formula 5.7 for γ_1 and analogous formulae for γ_2, γ_3 and γ_4 it can be seen that

$$\gamma_1 = \gamma_3 = \gamma_4 = 25\mu^3$$

$$\gamma_2 = 50\mu^3.$$

giving the fixation probability as

$$\alpha(w_0, x_0, y_0, z_0) = \frac{w_0 + 2x_0 + y_0 + z_0}{N_1 + 2N_2 + N_3 + N_4} \tag{5.10}$$

So again the central subpopulation which is twice as productive as the other subpopulations is twice as important in determining the fixation probability as the other subpopulations are.

5.10 Mean absorption and fixation times

As in the case of three subpopulations, an analytic solution for the mean absorption and fixation times of an allele is possible but is so complicated arithmetically for any real situation that it could not reasonably be calculated even with symmetric equal migration. However, it is possible to numerically find the mean absorption and fixation times for any given set of parameters. Mean fixation times are not found here.

5.11 Summary of results for three and four subpopulations

In summary, the following results have been found for the three and four subpopulation case when there is no selection.

- The fixation probability in the whole population is equal to a weighted sum of the initial frequency in each subpopulation. This is also equal to a weighted sum of the fixation probability in each subpopulation considered separately.
- In the case of symmetric migration, the fixation probability in the whole population is equal to the initial frequency in the whole population. This is also true in the case when migration is not symmetric but there is no net migration between the subpopulations.
- In the case of asymmetric migration between the subpopulations, subpopulations which have a net outflow of migrants in general have a greater weight than subpopulations which have a net inflow of migrants. This idea is expanded on further when the general population model is considered.

- Mean absorption and fixation times were not examined for three and four subpopulations.

Chapter 6

A general model for fixation probabilities

In this chapter I consider the general case of a population divided into n subpopulations. I find a general result for asymmetric migration in a subdivided population with no selection. This is then interpreted using a result from graph theory and some general principles are found for fixation probabilities in a subdivided population.

This is a very significant extension of previous chapters and of the work of other authors. Other authors have modelled a general number of subpopulations with symmetric and no net migration (for example Maruyama, 1970a; Maruyama, 1974; Slatkin, 1981) or have only two subpopulations with asymmetric migration (Tachida and Iizuka, 1991). Considering more than two populations allows the exploration of other more general structures - for example edge/centre subpopulation relationships, rings of populations and hub/spoke configurations.

In addition to the general results, this model could be used in specific cases to determine which subpopulations are the most significant in terms of loss of genetic variability, and to assess the consequences of management decisions such as removing a subpopulation or reducing the ability for individuals to move between subpopulations by removing a habitat corridor (Walters, 1991; Lamberson *et al.*, 1992; Winter *et al.*, 1997). As with the simpler models, the outcome depends on the relative, rather than the absolute, magnitudes of migration between the subpopulations. This means that if data is available but is rough, it is only necessary to estimate the relative magnitudes of migration rather than the average number of succuessful migrants between each pair of populations in each generation. For the purposes of estimating times to fixation it would be necessary to estimate absolute nagnitudes of migration, but that problem is not considered here for the general n subpopulations.

6.1 Description of the model

In this model the subpopulation is divided into n subpopulations and the population structure can be represented diagramatically as in figure 6.1. The model is essentially a generalization of the assumptions for the two, three and four subpopulation cases to the case of n subpopulations. It is an n subpopulation version of the model described in chapter 2 and the n subpopulations are of constant size N_1, N_2, $\ldots$,N_n. Migration is allowed between any pair of subpopulations and the mean number of migrants from subpopulation i to subpopulation j in a given generation is μ_{ij}, which is also the probability of a migrant from subpopulation i to subpopulation j in a given generation. There is a maximum of one migrant from any population, i, to any other population, j, in a single generation. The state variables in the model are the number of A_1 alleles in each subpopulation, with $X_{i,t}$ being the number of A_1 alleles in subpopulation i at time t. The state of the system at time t is denoted by the state vector $\mathbf{x}_t = [x_{1,t}, x_{2,t}, \ldots, x_{n,t}]'$. As with the previous models it is assumed that the adult breeding stock are recruited from a large pool of juveniles, so $X_{i,t+1}|\mathbf{x}_t$ and $X_{j,t+1}|\mathbf{x}_t$

are independent of each other for all pairs i, j.

Defining

$$\gamma_i = \frac{|A_{ii}|}{\sum_{i=1}^{n} N_i |A_{ii}|},$$

where $\mathbf{A}$ is the matrix with

$$\begin{aligned} a_{ij} &= -\mu_{ij} \quad i \neq j \\ a_{ii} &= \sum_{k \neq i} \mu_{ik}, \end{aligned}$$

that is

$$\mathbf{A} = \begin{bmatrix} \sum\limits_{k=2}^{n} \mu_{1k} & -\mu_{21} & \cdots & -\mu_{n1} \\ -\mu_{12} & \sum\limits_{\substack{k=1 \\ k\neq 2}}^{n} \mu_{2k} & \cdots & -\mu_{n2} \\ \vdots & \vdots & & \vdots \\ -\mu_{1n} & -\mu_{2n} & \cdots & \sum\limits_{k=1}^{n-1} \mu_{nk} \end{bmatrix} \tag{6.1}$$

and A_{ii} is the matrix obtained by removing row i and column i from matrix $\mathbf{A}$, the following result holds.

Theorem 3 *For the model described above the fixation probability is*

$$\alpha(x_{1,0}, x_{2,0}, \ldots, x_{n,0}) = \sum_{i=1}^{n} \gamma_i x_{i,0} \tag{6.2}$$

where

Formula 6.2 has a geometrical interpretation which will be discussed after the proof of this theorem.

Proof:

Step 1: Show $\sum \gamma_i x_{i,t} = \sum \gamma_i E[x_{i,t+1}|\mathbf{x}_t]$.

The probability of fixation from state $\mathbf{x}_t$ is equal to the probability that the system changes to state $\mathbf{x}_{t+1}$ in the next generation and is then fixed from this new state,

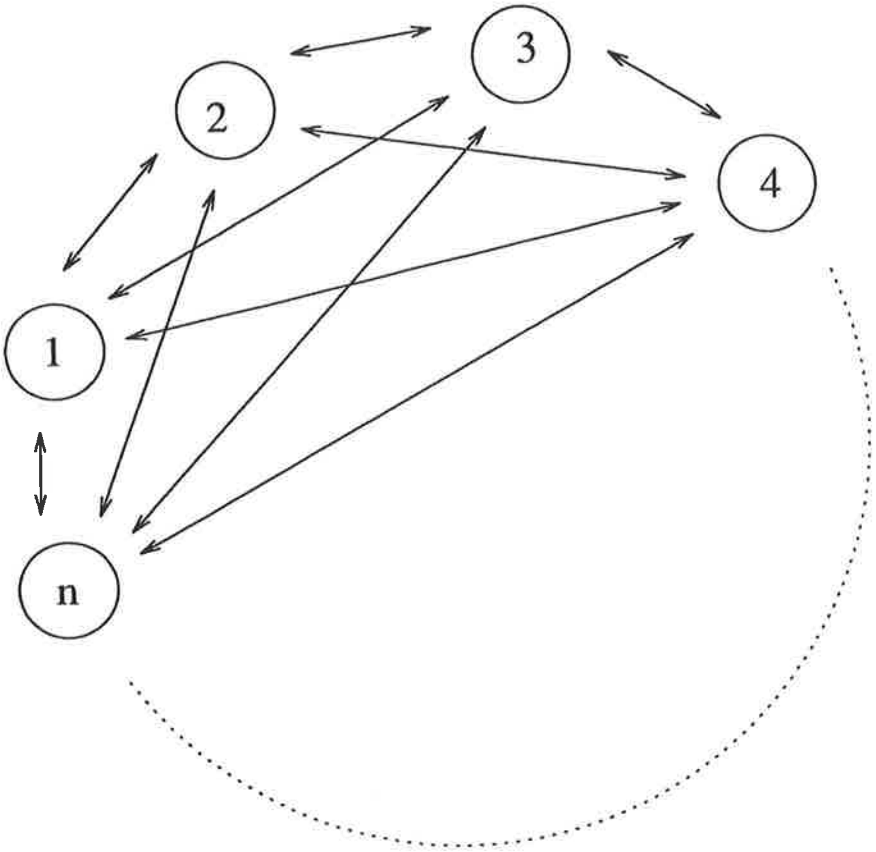


Figure 6.1: Diagramatic representation of the model when there are n subpoulations. Each double-ended arrow between two subpoulations represents a two way link between the two subpopulations.

summed over all possible states that it could transfer to in one generation. In symbols,

$$P(\text{fixation from } \mathbf{x}_t) = \sum_{\mathbf{X}_{t+1}} P(\mathbf{x}_t \to \mathbf{x}_{t+1}) P(\text{fixation from } \mathbf{x}_{t+1}). \tag{6.3}$$

This is the backward Kolmogorov equation and is analagous to equation 3.6 in the derivation of the two subpopulation formula.

Also, as $X_{i,t+1}|\mathbf{x}_t$ and $X_{j,t+1}|\mathbf{x}_t$ are assumed independent of each other

$$P(\mathbf{x}_t \to \mathbf{x}_{t+1}) = P(\mathbf{x}_{t+1}|\mathbf{x}_t) = \prod_{l=1}^{n} P(x_{l,t+1}|\mathbf{x}_t).$$

The left hand side of equation 6.3 is the formula given in equation 6.2. That is, $\sum_{i=1}^{n} \gamma_i x_{i,t}$.

The right hand side of equation 6.3 is

$$\begin{aligned}
&\sum_{\mathbf{X}_{t+1}} P(\mathbf{x}_t \to \mathbf{x}_{t+1}) P(\text{fixation from } \mathbf{x}_{t+1}) \\
=\ & \sum_{\mathbf{X}_{t+1}} \prod_{l=1}^{n} P(x_{l,t+1}|\mathbf{x}_t) \sum_{i=1}^{n} \gamma_i x_{i,t+1} \\
=\ & \sum_{x_{1,t+1},\ldots,x_{n,t+1}} \sum_{i=1}^{n} \gamma_i x_{i,t+1} \prod_{l=1}^{n} P(x_{l,t+1}|\mathbf{x}_t) \\
=\ & \sum_{i=1}^{n} \sum_{x_{1,t+1}} \cdots \sum_{x_{i-1,t+1}} \sum_{x_{i+1,t+1}} \cdots \sum_{x_{n,t+1}} P(x_{1,t+1},\ldots,x_{i-1,t+1},x_{i+1,t+1},\ldots,x_{n,t+1}|\mathbf{x}_t) \\
& \times \gamma_i \sum_{x_{i,t+1}} P(x_{i,t+1}|\mathbf{x}_t) x_{i,t+1} \\
=\ & \sum_{i=1}^{n} 1.\gamma_i \sum_{x_{i,t+1}} P(x_{i,t+1}|\mathbf{x}_t) x_{i,t+1} \\
=\ & \sum_{i=1}^{n} \gamma_i E[X_{i,t+1}|\mathbf{x}_t]
\end{aligned} \tag{6.4}$$

Step 2: show $E[X_{l,t+1}] = \left(1 - \sum_{\substack{k=1\\k\neq l}}^{n} \frac{\mu_{kl}}{N_l}\right) x_{l,t} + \sum_{\substack{k=1\\k\neq l}}^{n} \frac{\mu_{kl}}{N_k} x_{k,t}$.

The expected number of A_1 alleles in subpopulation l at time $t+1$, conditional on $\mathbf{x}_t$, the number of A_1 alleles present in each subpopulation at time t is equal to the sum of the expected numbers of A_1 alleles in subpopulation l which originated in each of the subpopulations.

The mean number of migrants from subpopulation k ($\neq l$) is μ_{kl}. The probability that each of these migrants is an A_1 allele is $x_{k,t}/N_k$ from subpopulation k. So the expected number of A_1 alleles in subpopulation l which migrated from subpopulation k in generation $t+1$ is $\mu_{kl}x_{k,t}/N_k$. This means the expected number of A_1 alleles in subpopulation l which migrated from subpopulations other than subpopulation l in this generation is

$$\sum_{k \neq l} \frac{\mu_{kl} x_{k,t}}{N_k} \tag{6.5}$$

The expected total number of immigrants to subpopulation l at time $t+1$ is $\sum_{k\neq l} \mu_{kl}$ and so the expected number of individuals which are born in subpopulation l and stay there is $N_l - \sum_{k\neq l} \mu_{kl}$. Each of these individuals has probability $x_{l,t}/N_t$ of having an A_1 allele. So the expected number of A_1 individuals being born in subpopulation l and being recruited to the breeding population in subpopulation l is

$$\left(N_l - \sum_{k\neq l} \mu_{kl}\right) \frac{x_{l,t}}{N_l} \tag{6.6}$$

Adding equations 6.5 and 6.6 gives

$$\begin{aligned} E[X_{l,t+1}|\mathbf{x}_t] &= \left(N_l - \sum_{k\neq l} \mu_{kl}\right) \frac{x_{l,t}}{N_l} + \sum_{k\neq l} \frac{\mu_{kl} x_{k,t}}{N_k} \\ &= \left(1 - \sum_{k\neq l} \frac{\mu_{kl}}{N_l}\right) x_{l,t} + \sum_{k\neq l} \frac{\mu_{kl}}{N_k} x_{k,t} \end{aligned} \tag{6.7}$$

as required.

Step 3: Derive $\mathbf{A\Gamma} = 0$.

Using the equations 6.4 and 6.7 it is possible to find a matrix equation which the γ_i must satisfy. Combining these equations gives

$$\sum_{i=1}^{n} \gamma_i x_{i,t} = \sum_{l=1}^{n} \gamma_l \left[\left(1 - \sum_{k\neq l} \frac{\mu_{kl}}{N_l}\right) x_{l,t} + \sum_{k\neq l} \frac{\mu_{kl}}{N_k} x_{k,t}\right] \tag{6.8}$$

and rearranging this gives

$$-\sum_{l=1}^{n}\gamma_l\left(\sum_{k\neq l}\frac{\mu_{kl}}{N_l}\right)x_{l,t}+\sum_{l=1}^{n}\sum_{k\neq l}\gamma_l\frac{\mu_{kl}}{N_k}x_{k,t}=0 \tag{6.9}$$

Equating coefficients of $x_{i,t}$ gives the following n equations

$$\begin{aligned} -\gamma_i\sum_{k\neq i}\frac{\mu_{ki}}{N_i}+\sum_{l\neq i}\gamma_l\frac{\mu_{il}}{N_i} &= 0 \qquad \text{for } i=1,\ldots,n \\ \Rightarrow -\gamma_i\sum_{k\neq i}\mu_{ki}+\sum_{l\neq i}\gamma_l\mu_{il} &= 0 \end{aligned}$$

defining $\Gamma=[\gamma_1,\ldots,\gamma_n]'$ and using the matrix $\mathbf{A}$ from equation 6.1, this can be rewritten in matrix form as

$$\begin{bmatrix} \sum\limits_{k=2}^{n}\mu_{1k} & -\mu_{21} & \cdots & -\mu_{n1} \\ -\mu_{12} & \sum\limits_{\substack{k=1\\k\neq 2}}^{n}\mu_{2k} & \cdots & -\mu_{n2} \\ \vdots & \vdots & & \vdots \\ -\mu_{1n} & -\mu_{2n} & \cdots & \sum\limits_{k=1}^{n-1}\mu_{nk} \end{bmatrix}\begin{bmatrix}\gamma_1\\ \gamma_2\\ \vdots\\ \gamma_n\end{bmatrix}=0 \tag{6.10}$$

Step 4: Solve $\mathbf{A}\Gamma=0$.

Note that the matrix $\mathbf{A}$ has the property that all of its column sums are zero, so $|\mathbf{A}|=0$, that is, it is non-invertible. This means that equation 6.10 has a non-trivial solution where $\Gamma\neq 0$.

Scale the variables γ_i to create the new variables $\beta_i=\gamma_i/c_{ii}$, where $c_{ii}=|A_{ii}|=(-1)^{i+i}|A_{ii}|$. Multiplying column i of $\mathbf{A}$ by c_{ii} gives the matrix

$$\mathbf{B}=\begin{bmatrix} c_{11}\sum\limits_{k=2}^{n}\mu_{1k} & -c_{22}\mu_{21} & \cdots & -c_{nn}\mu_{n1} \\ -c_{11}\mu_{12} & c_{22}\sum\limits_{\substack{k=1\\k\neq 2}}^{n}\mu_{2k} & \cdots & -c_{nn}\mu_{n2} \\ \vdots & \vdots & & \vdots \\ -c_{11}\mu_{1n} & -c_{22}\mu_{2n} & \cdots & c_{nn}\sum\limits_{k=1}^{n-1}\mu_{nk} \end{bmatrix}$$

and $\mathbf{B}\beta=0$. Also each column sum of $\mathbf{B}$ is zero, so $|\mathbf{B}|=0$.

Lemma 1 *If* **C** *is a matrix with all column sums equal to zero then* $(-1)^{x+i}|C_{xi}| = (-1)^{y+i}|C_{yi}|$ *for all* $i, x, y = 1, 2, \ldots, n$.

Proof of lemma 1: This is seen using a lemma from Chen (Lemma 4.11971). The lemma says that if a matrix A has all row sums equal to zero then $(-1)^{i+x}|A_{ix}| = (-1)^{i+y}|A_{iy}|$. The transpose matrix C' has row sums equal to zero so $(-1)^{i+x}|C'_{ix}| = (-1)^{i+y}|C'_{iy}|$, that is $(-1)^{i+x}|C_{xi}| = (-1)^{i+y}|C_{yi}|$ as required.

□

Now, denoting the (i, j)th element of **A** and **B** by a_{ij} and b_{ij} respectively, note that the matrix **B** has the property that the sum of row j is

$$\begin{aligned}\sum_j b_{ij} &= \sum_j c_{jj} a_{ij} \\ &= \sum_j a_{ij}|A_{jj}| \\ &= \sum_j (-1)^{i+j} a_{ij}|A_{ij}| \\ &= |\mathbf{A}| \\ &= 0\end{aligned}$$

So **B** has each row sum equal to zero. This means that **B** is an equicofactor matrix, that is $(-1)^{i+j}|B_{ij}| = (-1)^{k+l}|B_{kl}| \ \forall i, j, k, l$.

Clearly, as the row sums of **B** are all zero, $\beta = \theta[1, 1, \ldots, 1]'$ is a solution of $\mathbf{B}\beta = 0$. That is $\beta_i = \theta$, or, $\gamma_i = \beta_i c_{ii} = \theta|A_{ii}|$ is a solution of $A\Gamma = 0$.

Using the boundary condition $\alpha(N_1, N_2, \ldots, N_n) = 1$ gives the constant of proportionality,

$$\sum_i \gamma_i = \sum_i \theta|A_{ii}|N_i = 1$$

so $\theta = 1/\sum_i N_i|A_{ii}|$ and hence

$$\gamma_i = \frac{|A_{ii}|}{\sum_{i=1}^n N_i|A_{ii}|}$$

as required.

□

6.2 Interpretation of the general result using graph theory

The interpretation for the general result is an extension of the interpretation for the result for the cases of 2, 3 and four subpopulations. The weight given to a subpopulation in determining the fixation probability is the sum of a weight term for each path originating from that subpopulation. In short, the result is analogous to the previous cases – the subpopulations with the largest excesses of migrants will have the largest impact on the fixation probability. This enables us to make some useful generaliastions about fixation probabilities under various different population structures and migration rates.

Some background is required for the interpretation. One useful reason for constructing the equicofactor matrix, **B**, in the proof of theorem 3 rather than use an alternate proof is that there is a result from graph theory which leads to a geometrical interpretation of the terms of the cofactors of equicofactor matrices. First, some terms that need to be defined:

graph A graph $G(V,E)$, or simply G, consists of a set, V, of *nodes* together with a set, E, of unordered pairs of the form (i,j), with $i,j \in V$, called the *edges* of G. The nodes i and j are called the *endpoints* of (i,j). In this case each subpopulation is a node and a link between two subpopulations is an edge.

n = the number of nodes in the graph.

degree The degree of the node i is the number of edges incident with node i. Here this is the number of subpopulations directly linked to subpopulation i.

labelled graph A graph in which each node has been properly labelled.

directed-graph A directed graph is a set, V, of nodes with a set, E, of ordered pairs (i,j), with $i,j \in V$, called the *edges* of G. The node i is called the initial node of (i,j) and the node j is called the terminal node of (i,j). The only difference between a graph and a directed graph is that the edges of a directed graph are ordered pairs of nodes while the edges of a graph are unordered pairs. Here, the edges are a link between subpopulations with a direction in which the migrants travel – the migrants travel from subpopulation i to subpopulation j.

outgoing degree The number of edges which have i as there initial node. Here, the number of subpopulations which subpopulation i sends migrants too.

weighted directed-graph A directed graph in which each edge has been assigned a weight. Here this a link between two subpopulations which has a direction and a magnitude – in this application this is the mean number of migrants per generation. I denote by $f(i,j)$ the weight associated with edge (i,j). If G_S is a subgraph of G, then I denote by $f(G_S)$ the weight of the subgraph, G_S, and it is defined as

$$\begin{aligned} f(G_S) &= \prod_{(i,j)\in G_S} f(i,j) \quad \text{if } G_S \neq \emptyset \\ f(\emptyset) &= 0 \end{aligned}$$

associated directed graph For an equicofactor matrix, $\mathbf{Y}$, the *associated directed graph*, denoted by the symbol $G(\mathbf{Y})$ is an n-node, weighted, labelled, directed graph. The nodes are labelled by the integers from 1 to n such that if $i \neq j$ and

$y_{ij} \neq 0$, then there is an edge directed from node i to node j with associated weight y_{ij} for $i, j = 1, 2, \ldots, n$. So for example, the matrix $\mathbf{A}$ in equation 6.1 has the associated directed graph in figure 6.1.

tree A spanning subgraph of a directed graph is said to be a *tree* if and only if it is connected and contains no circuits.

directed tree with reference node i A subgraph of a graph G, denoted by the symbol t_i, is said to be a *directed tree* of G with reference node i if and only if (1) it is a tree of G, and (2) the outgoing degree of each node of t_i is 1 except the node i which has outgoing degree 0.

The result which is important here is from Chen (1971, Theorem 4.3).

Theorem 4 *If $\mathbf{Y}$ is an equicofactor matrix then the cofactor $(-1)^{i+j}|Y_{ij}|$ is given by*

$$(-1)^{i+j}|Y_{ij}| = \sum_{t_k} f(t_k) \tag{6.11}$$

for $i, j, k = 1, 2, \ldots, n$ where t_k is a directed tree in $G(\mathbf{Y})$.

or more importantly here

$$|Y_{ii}| = \sum_{t_i} f(t_i) \tag{6.12}$$

As $\mathbf{B}$ is an equicofactor matrix, each of the cofactors of $\mathbf{B}$ is equal. The value in terms of the matrix $\mathbf{A}$ can be calculated by using the following standard result from matrix theory.

Result: If $\mathbf{A}^*$ is the matrix obtained by multiplying each element of column j of matrix $\mathbf{A}$ by the scalar constant k then $|\mathbf{A}^*| = k|\mathbf{A}|$, so

$$|B_{ii}| = \prod_{l \neq i} c_{ll} |A_{ii}|$$

$$\begin{aligned} &= \prod_{l \neq i} |A_{ll}||A_{ii}| \\ &= \prod_{l=1}^{n} |A_{ll}|. \end{aligned} \tag{6.13}$$

Combining equation 6.13 with the result of theorem 4 gives

$$|B_{ii}| = \sum_{t_k} f(t_k) = \prod_{l=1}^{n} |A_{ll}|$$

where t_k is a directed tree in $G(\mathbf{B})$. The directed graph $G(\mathbf{B})$ is the same as the directed graph $G(\mathbf{A})$ with the weights μ_{ij} replaced by $c_{ii}\mu_{ij}$.

From the definition of t_k, a directed tree in $G(\mathbf{B})$, the outgoing degree of each node of t_k is 1 except for node k which has degree 0. So each term in the sum will be of the form $\prod_{l \neq k} \mu_{lr} c_{ll}$ for some $r \neq k, r \neq l$ which is equal to $f(t_k) = \prod_{l \neq k} c_{ll} \prod_{l \neq k} \mu_{lr}$. And denoting the corresponding directed tree in $G(\mathbf{A})$ by s_k gives $f(s_k) = \prod_{l \neq k} \mu_{lr} = f(t_k) / \prod_{l \neq k} c_{ll}$, so

$$\begin{aligned} \sum_{s_k} f(s_k) &= \sum_{t_k} f(t_k) / \prod_{l \neq k} c_{ll} \\ &= |B_{kk}| / \prod_{l \neq k} c_{ll} \\ &= \prod_{l=1}^{n} |A_{ll}| / \prod_{l \neq k} c_{ll} \\ &= |A_{kk}|. \end{aligned} \tag{6.14}$$

This means that for the matrix $\mathbf{A}$ the cofactor $|A_{ii}|$ can also be considered as a sum of terms representing the directed trees in $G(\mathbf{A})$ with reference node i. That is, the coefficients γ_i can be represented in terms of a set of subgraphs of the graph $G(\mathbf{A})$.

Consider the coefficient γ_i. From theorem 6.2 we know that $\gamma_i = |A_{ii}|/(\sum_l |A_{ll}| N_i)$. From theorem 4 and the above corollary we know that $|A_{ii}| = \sum_{s_i} f(s_i)$ where s_i is a directed tree in $G(\mathbf{A})$ with reference node i, so

$$\gamma_i = \frac{\sum_{s_l} f(s_l)}{\sum_l N_l \sum_{s_l} f(s_l)}.$$

This is the normalised sum of the terms for trees with reference node i. this formula will now be used to determine the coefficients, γ_i, in some specific cases.

6.2.1 Some examples

In general there are n^{n-2} terms in the expansion of $|A_{ii}|$. In the cases considered in previous chapters – 2, 3 and 4 subpopulations – there were 1, 3 and 16 terms in the expansion which agrees with this formula. The 16 graphs shown in figure 5.9 for the terms in the expansion of γ_1 are in fact the directed trees in $G(\mathbf{A})$ with reference node 1.

Example 1: Equal, symmetric migration, $\mu_{ij} = \mu \ \forall i, j$. There are n^{n-2} terms each of which is μ^{n-1} so $|A_{ii}| = n^{n-2}\mu^{n-1}$ for all i. This gives subpopulation i the weight

$$\gamma_i = \frac{\mu^{n-1} n^{n-2}}{\sum_l N_l \mu^{n-1} n^{n-2}} \quad \forall i.$$

All of the subpopulation weights are the same, so all of the subpopulations have the same impact on the fixation probability. This means that the overall fixation probability for the whole population is equal to the initial frequency in the whole population.

Example 2: Symmetric migration, $\mu_{ij} = \mu_{ji}$. The row sums of matrix $\mathbf{A}$ are all zero so $\mathbf{A}$ is equicofactor. This means $\gamma_i = |A_{ii}|/(\sum_l |A_{ll}|N_l) = |A_{jj}|/(\sum_l |A_{ll}|N_l) = \gamma_j$ for all i, j. So once again the fixation probability is equal to the initial frequency in the whole population.

Examples 1 and 2 confirm the results that were given by Maruyama (1970a; 1970b; 1972; 1974).

Example 3: Population structure with minimal number of connections. With a this population structure there is only one directed tree with reference node i for each i. This means that each term γ_i has only one term in its expansion. So γ_i can be shown to be

$$\gamma_i \propto \prod_{l=1}^{i-1} \mu_{l,l+1} \prod_{l=i+1}^{n} \mu_{l,l-1}.$$

In the case where there is a prevailing wind along a line of subpopulations such that $\mu_{i,i+1} = 2\mu,\ \mu_{i,i-1} = \mu$ this can be evaluated to give $\gamma_i = 2^{i-1}\mu^{n-1}$. This means that each subpopulation is twice as important as the subpopulation immediately down wind from it as the wind is causing twice as many migrants to go with the wind as against it. For this reason, Subpopulations which are a long way down wind will have little influence on the overall population fixation probabilities.

Example 4: Central subpopulation connected to $n-1$ subpopulations with central subpopulation most productive. In this situation there is only one path in each of the weights. The weight of the central subpopulation is $\prod_{i=2}^{n} \mu_{1i}$ and the weight of each of other subpopulation j is $\mu_{j1} \prod_{i=2, i\neq j}^{n} \mu_{1i}$. If the central subpopulation sends μ_a migrants to each of the outer subpopulations and each of the outer subpopulations sends μ_b migrants where $\mu_a > \mu_b$, then the weight of each subpopulation is $\gamma_1 = \mu_a^{n-1}$ and $\gamma_j = \mu_b \mu_a^{n-2}$. So the central, most productive subpopulation here has a weight μ_a/μ_b times greater than each of the outer subpopulations. If the central subpopulation was no more productive than each of the outer subpopulations then each subpopulation would have an equal weight as migration would be symmetric.

Example 5: Now consider the above population structure with one of the outer subpopulations producing the greatest excess of migrants. Call the outer subpopulation which is most productive subpopulation 2 and $\mu_{21} = \mu_a$, $\mu_{j1} = \mu_b$ when $j = 3, \ldots, n$ and $\mu_{1j} = \mu_b$. Then the weights for each of the subpopulations is $\gamma_2 = \mu_b \mu_a^{n-2}$ and $\gamma_j = \mu_a^{n-1}$ for all other values of j. This shows that position of a subpopulation is not important. The important thing in determining the weight of a subpopualtion is whether or not a subpopulation produces an excess of migrants.

Example 6: No net migration – circular migration. An example of no net migration without having symmetric migration is circular migration. In this case, the subpopulations are arranged in a circle with each subpopulation sending migrants only to the subpopulations to the immediately adjacent subpopulations on either side. In this situation, each of the weight terms is made up of n terms. As an example, $\gamma_1 = \sum_{k=1}^{n} \prod_{i=1}^{k-1} \mu_{i,i+1} \prod_{i=k+1}^{n} \mu_{i+1,i}$. If $\mu_{i,i+1} = \mu_{n1} = \mu_a$ and $\mu_{i+1,i} = \mu_{1n} = \mu_b$, then $gamma_1 = \sum_{k=1}^{n} \mu_a^k \mu_b^{n-k}$. Similarly, each of the weights can be shown to be equal to γ_1 and so each subpopulation has equal weight in determining the fixation probability.

6.3 Conclusions

In this chapter a formula has been found for the fixation probability of a neutral allele A_1 for a population divided into n subpopulations. An interpretation of this formula has been given using graph theory which can be used to interpret some special cases such as the population structure with the minimal number of connections. The fixation time formula can not be solved in a general form and would need to be solved numerically for specific cases. The conclusions relating to the fixation probability support the conclusions of other authors who have considered the problem using other models. This result also extends the results of other authors to the case where there is asymmetric migration.

The size of a subpopulation is also important in determining how much it influences fixation probability. The influence of a subpopulation is directly proportional to its size (ignoring the effect of differential migration).

I found that in general when migration is asymmetric, and ignoring the impact of size on the weight of a subpopulation, the subpopulations which produce the greatest surplus of migrants are the most important subpopulations in determining the fixation probability. That is, those subpopulations which send out more migrants than they receive are the most important subpopulations in determining the fixation probability of the whole population. I also found that the position of a subpopulation within the population is not important in determining its impact on the fixation probability of the whole population. That is, an edge and a centre subpopulation have an equal impact on the fixation probability of the whole population if they are each producing the same net number of migrants provided the migrants are going to subpopulations of equal importance. The position relative to source/sink populations does have an effect on the weight of a subpopulation. If a subpopulation tends to send migrants to a subpopulation which in turn has a large surplus of outgoing migrants, whereas another subpopulation tends to send migrants to a subpopulation which has a net deficit of outgoing migrants then the first of these subpopulations will have a greater impact on the fixation probability of the population as a whole. This is because the first of these subpopulations is indirectly having its alleles spread more widely than the second subpopulation. In terms of the graph theory used here, the trees referenced to the first subpopulation have a greater weight. The trees referenced to the second of these subpopulations have a lesser weight.

6.4 Summary of chapters 3, 4, 5 and 6

In chapters 3, 4, 5 and 6 I have considered the problem of fixation probability and mean fixation and absorption times for subpopulations with two, three and four subpopulations. I have also considered the fixation probability for a population subdivided into a general n subpopulations. The models used have several assumptions in common. Firstly they are all based on the Wright-Fisher model. They have no selection. Subpopulations are all of constant size.

There are several general conclusions that have been found to be true for each of these models.

- The fixation probability in the whole population is equal to a weighted average of the intial frequency in each subpopulation. This is also equal to a weighted average of the fixation probability in each subpopulation considered separately.
- When migration is symmetric between connected pairs of subpopulations, the fixation probability of an allele in the whole population is equal to the overall initial frequency of that allele in the whole population.
- When there is no net migration between the subpopulations, for example when there is circular migration, as in example 6 in section 6.2.1, the fixation probability is also equal to the initial frequency in the whole population. This is equivalent to the result obtained by Maruyama (1970a).
- When migration is not symmetric then the subpopulations which produce the greatest excess of migrants tend to have the largest impact on the fixation probability. The fixation probability will not generally be equal to the overall initial frequency but will be closer to the initial frequency of those subpopulations which produce the most migrants.
- The position of a subpopulation relative to other subpopulations is not important in detemining how much influence that subpopulation will have on the fixation probability. That is, whether a subpopulation is near the edge or the centre is not in itself important in determining influence it has on the fixation probability.
- In terms of the graph theory used here, the subpopulations which have the greatest effect on the fixation probability are those subpopulations whose referenced trees have the greatest weight. As mentioned in the previous point, this will generally be those subpopulations which produce the most migrants.

However, as mentioned in section 6.3, it is also important how many migrants are produced by the subpopulations which receive the migrants from the subpopulation in question. Essentially, this is saying that the sum of the weights of the trees referenced to a subpoulation is the important quantity, which is equivalent in words to the main theorem(3) of this chapter.

Chapter 7

Fixation Probabilities with selection

7.1 Introduction

In this chapter I introduce selection into a single population Wright-Fisher model. In the preceding chapters it has always been assumed that the alternative alleles, A_1 and A_2, at the locus under consideration are selectively neutral. This means that in the models used up until now, A_1 and A_2 juveniles have been produced by a given subpopulation according to the proportion of A_1 and A_2 adults in the previous generation in that subpopulation. It has been shown when there is no selection that if each population produces the same number of migrants on average (equal symmetric migration) then the fixation probability of an allele is equal to the initial frequency of that allele in the whole population. If this is not true but the weaker condition of symmetric migration holds then it is still true that the probability of fixation of an allele is equal to the inital frequency of that allele in the whole population. In general, when there is no selection, the fixation probability is a weighted sum of the initial frequency in each of the subpopulations, or a weighted sum of the fixation probability in each subpopulation considered separately.

With the introduction of selection into the model these results no longer hold. Each time a new generation of individuals is born one allele is selectively favoured over the other, leading to production of more of the selectively favoured allele than when there is no selection acting. This increases its chance of fixation to a level greater than its initial frequency in the whole population. The inclusion of selection in the model used in chapters 3, 4 and 5 makes the model far more complicated as the model is no longer linear. Few authors have attempted to find exact analytic formulas for fixation probabilities when there is selection acting. Those authors that have made progress with this problem have either assumed the population is very large and used deterministic models to approximate the system (Maruyama, 1970a; Maruyama, 1972; Slatkin, 1981), or have used simulation studies (Slatkin, 1981). More recently Tachida and Iizuka (1991) have used an idea introduced by Slatkin (1981) where it is assumed that the time between migration events is much greater than the mean time to absorption in each subpopulation to gain some useful results. This is known as the low migration limit.

Solving the Kolmogorov equations ($\alpha_i = P_{ik}\alpha_k$) when selection is allowed is difficult because the solution is no longer a linear combination of the initial frequencies in each of the subpopulations with the introduction of selection. It is not just a matter of solving a system of linear equations to find the weights of the initial allele frequencies in each subpopulation.

In this chapter I will find general formulae which the fixation probability and the mean fixation and absorption times must satisfy. I will solve these for some small populations to get exact analytic results in some special cases.

I then combine this model with models from chapters 3, 4 and 5 to get a subdivided population model with selection. I describe a model which combines these two assumptions and find a matrix formula for the fixation probability and for the mean absorption time. Numerical solutions will be given for a simple case and these solutions will be discussed. The model used here is essentially the same as the model used in chapter 3 with selection added in. In general it is very difficult to find analytic solutions for this model. However, exact numerical solutions can be found for many useful cases.

This work uses an exact Wright-Fisher model to address some of the same questions sddressed by Slatkin (1981), Barton (1993), Tachida and Iizuka (1991) and Maruyama (1970a) using various approximations. I will compare some of the relevant results of Slatkin (1981), Maruyama (1970a), and Tachida and Iizuka (1991) with the symmetric migration case for this model. All of these models show the same qualitative effects for the fixation probability. Some authors consider mean fixation time rather than mean absorption time which I consider so it is more difficult to compare this aspect of each of the models. There is less agreement between the models when mean fixation time is considered. Barton's (1993) results appear qualitatively different from this model and the other models considered here for both fixation probability and mean fixation time even when only weak selection is considered.

7.2 Single Population model

Consider a single (unsubdivided) population with a constant size of N individuals. The assumptions are as described in Chapter 2 with the addition of selection. The state variable is $X(t)$ which is the number of A_1 individuals in the population at time t. There is selection acting in the model which is parameterised by the variable $s \in (-1, \infty)$. I assume the selection variable only affects the system through its effect on the function $\sigma(i, s)$ which is used to determine the transition probablilities. The transition probabilities under selection will now be described in more detail.

7.3 Transition probabilities

In order to calculate the transition probabilities it is first necessary to specify how selection takes place. As specified in Chapter 2, it is assumed that there are many juveniles born and that the next generation of adults is then sampled from these juveniles. The selection in the population is assumed to be on these juveniles before they recruit to the breeding stock. The function $\sigma(i, s)$ is defined to be the probability that an individual which is in the pool of juveniles and available to be recruited to the breeding stock at generation $t+1$ has the A_1 allele, given there were i individuals with the A_1 allele and $N-i$ individuals with the A_2 allele at generation t.

A selection function is required which is equivalent to A_1 and A_2 alleles being born in the ratio $(1+s)i : N-i$ and these juveniles then being sampled binomially to make the next generation of breeding individuals. Because of the way s is defined, selection is not symmetric, that is a selection coefficient of, say -0.5, is not opposite in its selective effect to a selection coefficient of 0.5. If a new variable, s', is defined which is equal to s for negative values of s and equal to $s' = s/(1+s)$ for positive values of s, then the action of s' is symmetric. This transformation is used on the graphs in this section as it makes the action of selection easier to interpret because negative and positive values have the same effect. All calculations are performed using untransformed s.

I define the function $\sigma(i, s) = \frac{(1+s)i}{N+si}$. With this formula for $\sigma(i, s)$, s must lie in the range $-1 < s < \infty$ and the neutral case is when $s = 0$. It has the properties $\sigma(0, s) = 0$, $\sigma(N, s) = 1$, and $0 < \sigma(i, s) < 1$ for $0 < i < N$. A negative value for s signifies selection against allele A_1 and a positive value for s signifies selection in favour of the A_1 allele.

The transition probabilities for moving from state i to state k are given by the formula

$$\begin{aligned} p_{i,k} &= \binom{N}{k}\sigma(i,s)^k(1-\sigma(i,s))^{N-k} \\ &= \binom{N}{k}\left(\frac{(1+s)i}{N+si}\right)^k\left(1-\frac{(1+s)i}{N+si}\right)^{N-k}. \end{aligned}$$

The fixation probability when there are initially i A_1 alleles in the population is denoted α_i and can be found by solving the Kolmogorov equations.

$$\alpha_i = \sum_{k=0}^{N} p_{i,k}\alpha_k, \quad i = 0, 1, 2, \ldots, N. \tag{7.1}$$

In words this equation means that the probability of fixation from initial state i is equal to the probability that the first transition is to state k and then fixation occurs from state k, summed over all possible values for k.

As the fixation probabilities $\alpha_0 = 0$ and $\alpha_N = 1$ are already known these equations can be simplified to give

$$\alpha_i = \sum_{k=1}^{N-1} p_{i,k}\alpha_k + p_{i,N} \quad i = 1, 2, \ldots, N-1.$$

This can be formulated in matrix form as follows.

$$\alpha = \mathbf{P}\alpha + P_N \tag{7.2}$$

or

$$\alpha = (\mathbf{I} - \mathbf{P})^{-1}P_N \tag{7.3}$$

where

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{N-1} \end{bmatrix}, \tag{7.4}$$

$$\mathbf{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N-1} \\ \vdots & \vdots & & \vdots \\ p_{N-1,1} & p_{N-1,2} & \cdots & p_{N-1,N-1} \end{bmatrix} \tag{7.5}$$

and

$$P_N = \begin{bmatrix} p_{1,N} \\ \vdots \\ p_{N-1,N} \end{bmatrix}. \tag{7.6}$$

Equation 7.3 is a set of $N - 1$ linear equations in $N - 1$ unknowns and can then be solved to find the fixation probabilities of the system in question.

This analysis applies whether or not there is selection as the transition probabilities, p_{ij}, have not been explicitly stated yet.

7.4 Finding the fixation probabilities

With the stated assumptions, it is possible to find an analytic solution for the fixation probability of the A_1 allele. In practice this will be arithmetically difficult for large populations, but this method can still be used to find numeric solutions for medium sized populations. For large populations a diffusion model (see for example, Maruyama, 1970a; Slatkin, 1981; Moran, 1960) would be better from a computational point of view and the assumptions of the diffusion model are generally satisfied.

In this section I will use p_{ij} as a short hand notation for $p_{ij}(s)$ as the transition probabilities are affected by the level of selection.

7.4.1 Solution for $N = 2$

In the case $N = 2$, α and $\mathbf{P}$ become

$$\alpha = [\alpha_1], \qquad \mathbf{P} = [p_{11}]$$

and equation 7.3 becomes

$$\alpha_1 = (1 - p_{11})^{-1} p_{12} = \frac{p_{12}}{1 - p_{11}}. \tag{7.7}$$

This is graphed in figure 7.1 for the case when there is initially one A_1 allele. The graph shows that when $s = 0$ (neutral alleles) the fixation probability is equal to the initial frequency, 0.5. When there is selection in favour of the A_1 allele ($s > 0$) it is more likely to be fixed than the neutral case, and when there is selection against the A_1 allele ($s < 0$) then it is less likely to be fixed than the neutral case as would be expected.

Now substituting $\sigma(i, s) = \frac{(1+s)i}{N+si}$ into equation 7.7 we get the fixation probability for the A_1 allele from the initial state where there is one A_1 allele out of the two alleles in the population is

$$\alpha_1 = \frac{\left(\frac{1+s}{2+s}\right)^2}{1 - \binom{2}{1}\frac{1+s}{2+s}\left[1 - \frac{1+s}{2+s}\right]}$$

7.4.2 Solution for $N = 3$

In the case $N = 3$, α and $\mathbf{P}$ become

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \qquad \mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

and

$$\alpha = \begin{bmatrix} 1 - p_{11} & -p_{12} \\ -p_{21} & 1 - p_{22} \end{bmatrix}^{-1} \begin{bmatrix} p_{13} \\ p_{23} \end{bmatrix} \tag{7.8}$$

so

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} [p_{13}(1 - p_{22}) + p_{12}p_{23}]/[(1 - p_{11})(1 - p_{22}) - p_{12}p_{21}] \\ [p_{23}(1 - p_{11}) + p_{21}p_{13}]/[(1 - p_{11})(1 - p_{22}) - p_{12}p_{21}] \end{bmatrix}$$

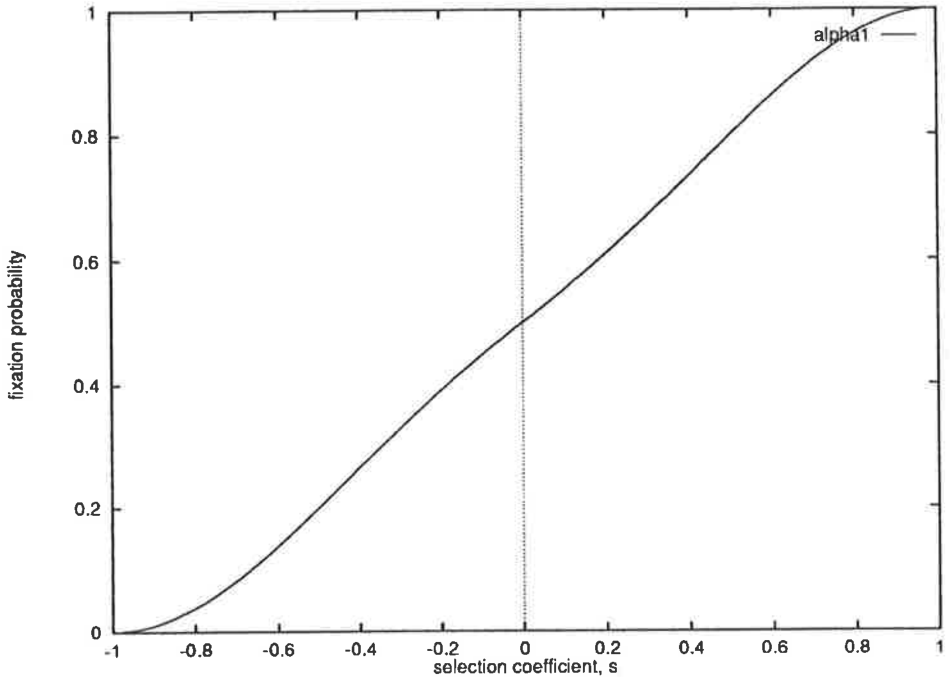


Figure 7.1: Plot of fixation probability as a function of the selection coefficient, s, when there is initially one A_1 allele in a single population of size two.

Substituting $\sigma(i,s)=\frac{(1+s)i}{N+si}$ we get

$$\alpha=\begin{bmatrix}\dfrac{\binom{3}{3}\left(\frac{1+s}{3+s}\right)^3\left(1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)\right)+\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)\binom{3}{3}\left(\frac{2(1+s)}{3+2s}\right)^3}{\left(1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2\right)\left(1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)\right)-\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2}\\ \dfrac{\binom{3}{3}\left(\frac{2(1+s)}{3+2s}\right)^3\left(1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2\right)+\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2\binom{3}{3}\left(\frac{1+s}{3+s}\right)^3}{\left(1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2\right)\left(1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)\right)-\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2}\end{bmatrix}.$$

This is graphed in figure 7.2. The results here are qualitatively the same as for the case of two populations. As would be expected, the probability of fixation of the A_1 allele is higher than the neutral case when there is selcetion in favour of the A_1 allele and lower than the neutral case when there is selection against the A_1 allele. It can also be seen that for any level of selection, the probability of fixation is higher for the case where the initial frequency is higher.

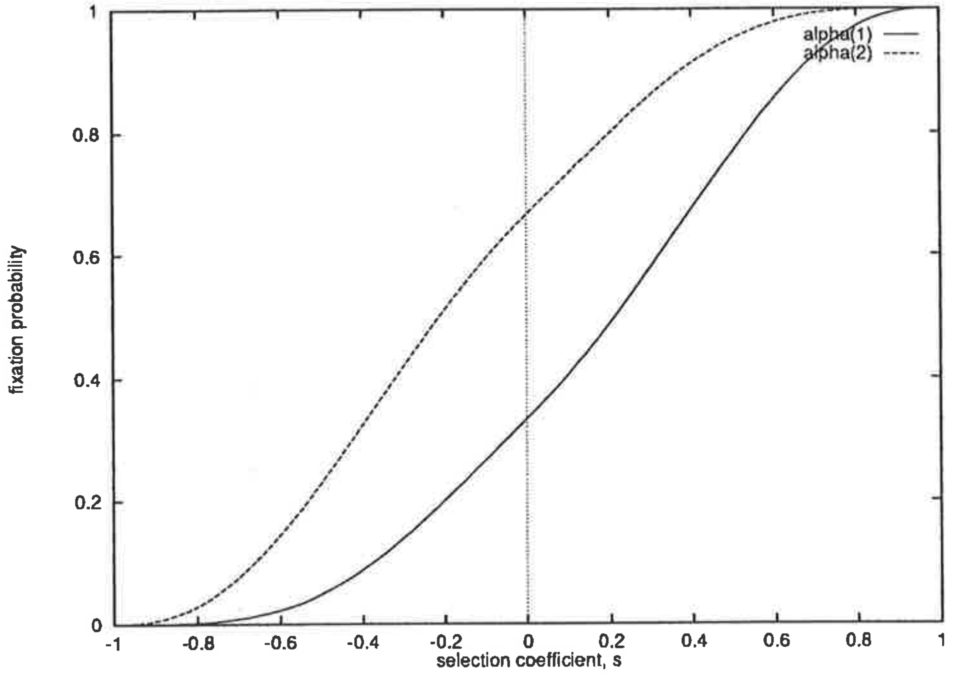


Figure 7.2: Plot of fixation probability as a function of the selection coefficient, s, when there are initially one (alpha(1)) and two (alpha(2)) A_1 alleles in a single population of size three. Increasing the selection for an allele increases its chance of fixation regardless of starting state.

7.4.3 Solution for $N \geq 4$

Equation 7.3 is a general formula which will not in general have a simple analytic solution for large N, however a matrix solution can be found. Equation 7.3 can be rewritten as

$$\mathbf{Q}\alpha = P_N$$

where $\mathbf{Q} = \mathbf{I} - \mathbf{P}$. Using the fact that $\mathbf{Q}$ is non-singular the solution can be written as

$$\alpha_i = \frac{\sum_{j=1}^{N-1} p_{j,N} |Q_{ji}|}{|\mathbf{Q}|}$$

where Q_{ji} is the matrix $\mathbf{Q}$ with row j and column i removed.

7.5 Finding the mean time to absorption

In chapter 4 a general formula was derived for the mean times to absorption and fixation for the case of no selection in a subdivided population and is given in equation 4.5. An analogous formula applies here for the mean absorption time with the entries of the probability matrix $\mathbf{P}$ changed to account for the selection acting in the population. The formula derived is

$$\mathbf{T} = (\mathbf{I} - \mathbf{P})^{-1}\mathbf{1} \tag{7.9}$$

or defining $\mathbf{Q} = (\mathbf{I} - \mathbf{P})$, we get

$$\mathbf{T} = \mathbf{Q}^{-1}\mathbf{1} \tag{7.10}$$

and once again there is not generally a straightforward analytic solution for this equation except when N is small.

In this chapter I will only consider mean absorption time as the calculation of mean fixation time is exactly analagous using a scaled probability $p^*_{i,k}$ instead of $p_{i,k}$.

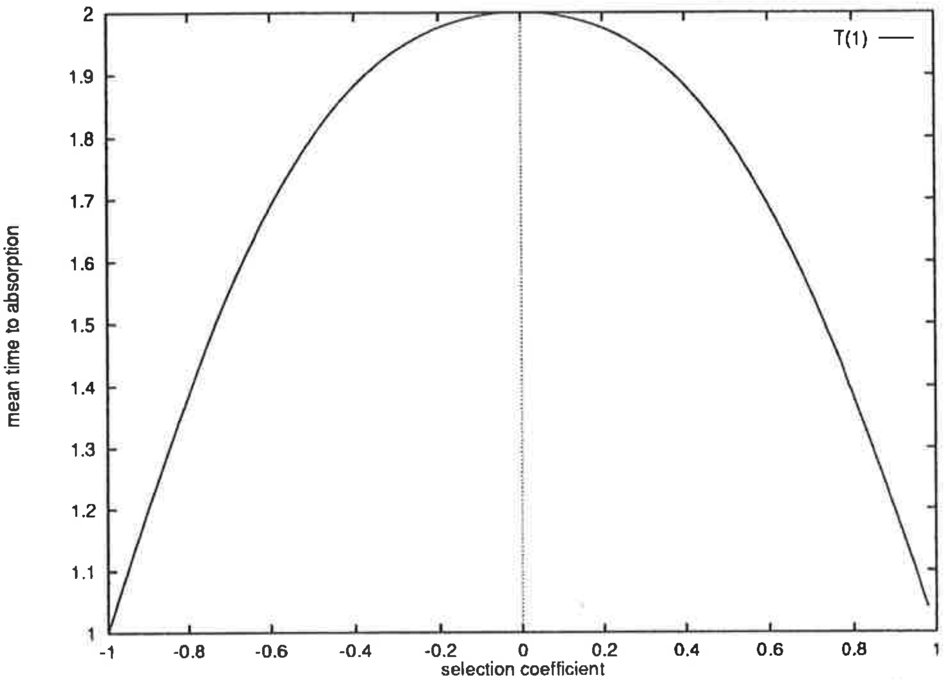


Figure 7.3: Plot of mean absorption time as a function of the selection coefficient, s, when there is initially one A_1 allele in a single population of size two. Increasing selection rapidly reduces time to absorption.

7.5.1 Mean absorption time in the case $N = 2$

In the case $N = 2$, $T_1 = (1 - p_{11})^{-1}$ and substituting $\sigma(i, s) = \frac{(1+s)i}{N+si}$, we get

$$T_1 = \frac{1}{1 - \binom{2}{1}\frac{1+s}{2+s}\left[1 - \frac{1+s}{2+s}\right]}.$$

This is graphed in figure 7.3. The main observation here is that the maximum of the mean absorption time occurs when selection is neutral ($s = 0$). At this point the mean absorption time is 2 generations, which is equal to the number of individuals in the population.

7.5.2 Mean absorption time in the case $N = 3$

The mean absorption times for the case $N = 3$ can be found using the equation

$$\mathbf{T} = \begin{bmatrix} 1 - p_{11} & -p_{12} \\ -p_{21} & 1 - p_{22} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{7.11}$$

Solving and substituting $\sigma(i, s) = \frac{(1+s)i}{N+si}$, we get

$$\begin{aligned} \mathbf{T} &= \begin{bmatrix} \frac{1-p_{22}+p_{12}}{\det \mathbf{Q}} \\ \frac{1-p_{11}+p_{21}}{\det \mathbf{Q}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)+\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)}{\left(1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2\right)\left(1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)\right)-\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2} \\ \frac{1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2+\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2}{\left(1-\binom{3}{1}\left(\frac{1+s}{3+s}\right)\left(1-\frac{1+s}{3+s}\right)^2\right)\left(1-\binom{3}{2}\left(\frac{2(1+s)}{3+2s}\right)^2\left(1-\frac{2(1+s)}{3+2s}\right)\right)-\binom{3}{2}\left(\frac{1+s}{3+s}\right)^2\left(1-\frac{1+s}{3+s}\right)\binom{3}{1}\left(\frac{2(1+s)}{3+2s}\right)\left(1-\frac{2(1+s)}{3+2s}\right)^2} \end{bmatrix}, \end{aligned}$$

which is graphed in figure 7.4. Notice that as in the case $N = 2$, the mean number of generations until absorption is equal to the population size when there is no selection acting. When $N = 3$ the maximum of the mean absorption times occur when the A_1 allele is not neutral. The maximum of the mean absorption time when there is initially one A_1 allele occurs when the selection decreases the probability that the system will quickly go to the nearest absorbing state, zero A_1 alleles, by weakly selecting in favour of the A_1 allele. Similarly when there are initially two A_1 alleles, the maximum mean absorption time occurs when weak selection is forcing the system away from the nearest absorbing state, that is, when there is weak selection against the A_1 allele.

7.5.3 Mean absorption time in the case $N \geq 4$

Equation 7.9 is a general formula which does not have a simple analytic solution for large N. However, a general matrix solution for the mean absorption time can be found. Equation 7.9 can be rewritten as

$$\mathbf{QT} = \mathbf{1},$$

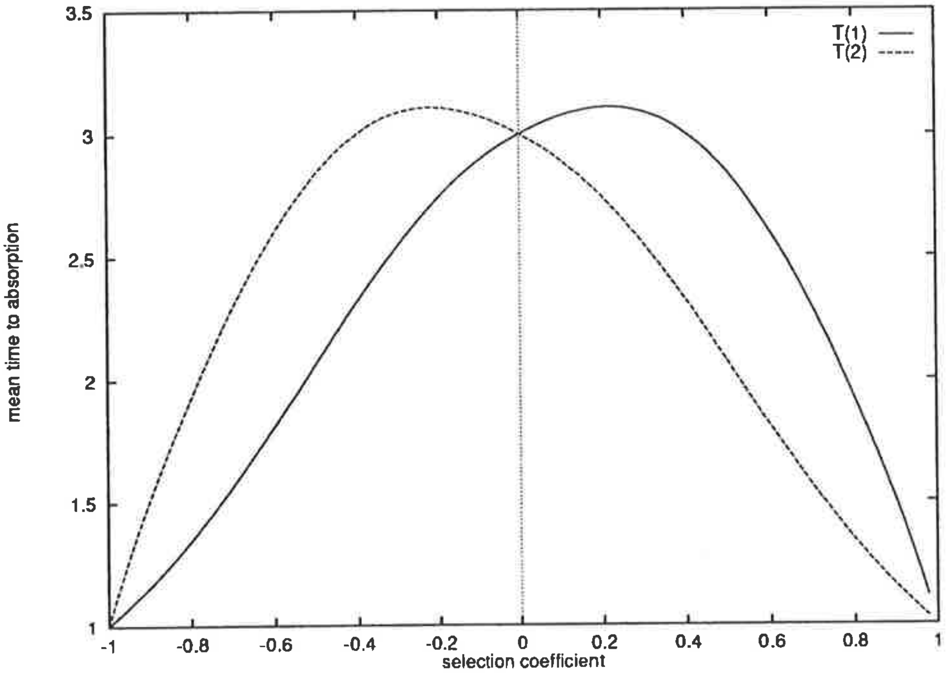


Figure 7.4: Plot of mean absorption time as a function of the selection coefficient, s, when there are initially one (T(1)) and two (T(2)) A_1 alleles in a single population of size three. Time to absorption rapidly decreases as selection strengthens.

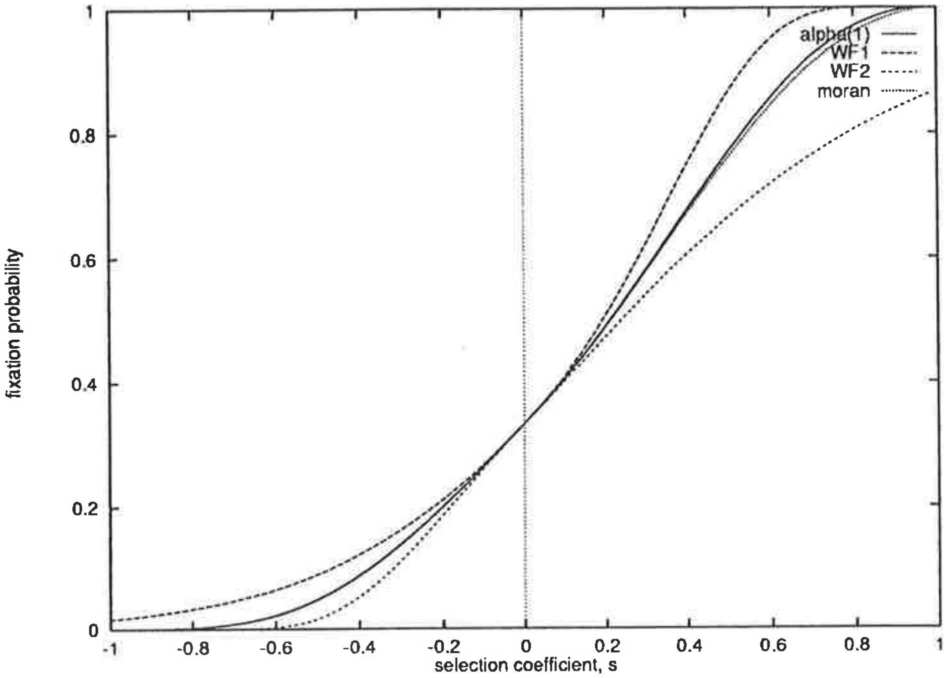


Figure 7.5: Comparison of fixation probability obtained using the exact Wright-Fisher model with two other methods from previous authors.

then because **Q** is non-singular the solution can be written as

$$T_i = \frac{\sum_{j=1}^{N-1} |Q_{ji}|}{|\mathbf{Q}|}$$

where Q_{ji} is the matrix **Q** with row j and column i removed.

7.6 Comparisons with other models which include selection

As the results are qualitatively the same for the cases $N = 2$ and $N = 3$ I will only consider here the case $N = 3$. The results obtained here are compared to the results obtained using the diffusion model (Maruyama, 1970a) and using the model used by Moran (1960).

The two Wright-Fisher graphs represent 1) the approximate Wright-Fisher model using $\theta = s$ and 2) the approximate Wright-Fisher model using $\theta = s/(1+s)$. The formula used here is

$$\frac{1 - e^{-2N\theta x_0}}{1 - e^{-2N\theta}}$$

as I use a population size of N, not $2N$ as was assumed by Moran (1960). Note that the diffusion method approach gives the same results as "WF1".

Figure 7.5 shows that the Moran approach gives a very close result to the result obtained here using what is essentially an exact Wright-Fisher method. The approximate Wright-Fisher result on the other hand gives a poor approximation to the results obtained here unless selection is very weak. It also gives a very large range of possible values between the upper and lower bounds. The diffusion approach also gives a poor approximation to the results obtained here as it is equal to the approximate Wright-Fisher model with $\theta = s$. It gives an estimate which is consistently closer to the initial frequency than the exact fixation probability is.

7.7 Subdivided Population model

The model developed in this section combines the subdivided population structure for two subpopulations as used in chapters 3 and 4 with the selection introduced for a single population in section 7.2. The model is a two subpopulation version of the model described in Chapter 2 with selection. The two subpopulations are of constant size N_1 and N_2. Migration is allowed between the two subpopulations but is limited to a maximum of one migrant in each direction per generation. The probability of a migrant from population 1 to 2 is μ_{12} and from population 2 to population 1 is μ_{21}. One haploid locus is considered and at this locus there are two alternative alleles A_1 and A_2. Selection is assumed to be of equal intensity in the two populations and is parameterised by the variable $s \in (-1, \infty)$. The selection variable, s, affects the functions $\sigma_1(i,s)$ and $\sigma_2(j,s)$ which are then used to determine the transition probabilities in subpopulations 1 and 2 respectively. As with the single population case where selection was included, in subpopulation 1, $\sigma_1(0,s) = 0$, $\sigma_1(N_1,s) = 1$ and $0 < \sigma_1(i,s) < 1$ for $0 < i < N_1$, and in subpopulation 2, $\sigma_2(0,s) = 0$, $\sigma_2(N_2,s) = 1$ and $0 < \sigma_2(j,s) < 1$ for $0 < j < N_2$.

The function $\sigma_1(i,s)$ is the probability that an individual in the pool of juveniles from subpopulation 1 at time $t+1$ has the A_1 allele, given there were i individuals with the A_1 allele and $N_1 - i$ individuals with the A_2 allele in subpopulation 1 at time t. It is defined as

$$\sigma_1(i,s) = \frac{(1+s)i}{N_1 + si}. \tag{7.12}$$

The function can be thought of as equivalent to A_1 alleles and A_2 alleles being born in subpopulation 1 in the ratio $(1+s)i : N_1 - i$ before migration and these individuals being sampled binomially after migration.

Similarly for population 2, the function used to determine the probability of an allele A_1 being born is

$$\sigma_2(j,s) = \frac{(1+s)j}{N_2 + sj} \tag{7.13}$$

and can be thought of as equivalent to A_1 alleles and A_2 alleles being born in subpopulation 1 in the ratio $(1+s)j : N_2 - j$ before migration and these individuals being sampled binomially after migration.

The variable of interest in the model is the number of A_1 alleles present in each population at time t, where X_t is the number present in population 1 and Y_t is the number present in population 2 at time t.

With probability μ_{21} there will be a migrant from population 2 to population 1. In this case X_{t+1} has the conditional distribution

$$X_{t+1}|X_t = i, Y_t = j, \text{migrant from pop } 2 \sim Bin(N_1 - 1, \sigma_1(i,s)) + Bin(1, \sigma_2(j,s)) \tag{7.14}$$

and with probability $1 - \mu_{21}$ there will be no migrant from population 2 to population 1, giving conditional distribution

$$X_{t+1}|X_t = i, Y_t = j, \text{no migrant from pop } 2 \sim Bin(N_1, \sigma_1(i,s)) \tag{7.15}$$

so the conditional distribution of X_{t+1} can be represented as

$$X_{t+1}|X_t = i, Y_t = j \sim \mu_{21}\left[Bin(N_1 - 1, \sigma_1(i,s)) + Bin(1, \sigma_2(j,s))\right] + (1-\mu_{21})Bin(N_1, \sigma_1(i,s)) \tag{7.16}$$

7.8 Finding the fixation probabilities

As in the case where there was no selection, the probability of fixation from state (i,j) is

$$\alpha(i,j) = \sum_{(k,l)} P((i,j) \to (k,l))\alpha(k,l) \tag{7.17}$$

and we know that $\alpha_{(0,0)} = 0$ and $\alpha_{(N_1,N_2)} = 1$ so equation 7.17 can be rewritten in matrix form as

$$\alpha' = \alpha'\mathbf{P} + P'_{(N_1,N_2)} \tag{7.18}$$

where

$$\alpha' = (\alpha(0,1), \alpha(0,2), \ldots, \alpha(0,N_2), \alpha(1,0), \ldots, \alpha(N_1, N_2 - 1)) \tag{7.19}$$

and

$$P'_{(N_1,N_2)} = (P_{(0,1)\to(N_1,N_2)}, P_{(0,2)\to(N_1,N_2)}, \ldots, P_{(N_1,N_2-1)\to(N_1,N_2)}) \tag{7.20}$$

is the vector containing the fixation probabilities for each state and $\mathbf{P}$ is the transition probability matrix. Putting $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ the system can be rewritten as

$$\mathbf{Q}'\alpha = P_{(N_1,N_2)}. \tag{7.21}$$

It can be shown using a counter example that $\alpha(i,j)$ is not of the form $f(i) + g(j)$ when there is selection acting, so the sum in equation 7.17 can not be simplified as it was in chapter 3 in the case when there was no selection acting. Without this simplification it becomes extremely difficult to find analytic solutions except when the populations are very small. The difficulty in calculating fixation probabilities arises because the matrix $\mathbf{P}$ is a $[(N_1+1)(N_2+1)-2] \times [(N_1+1)(N_2+1)-2]$ square matrix and to find the fixation probabilities the matrix $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ must be inverted. As population size increases, this matrix will increaes in size very rapidly with each dimension of the matrix being roughly equal to the square of the population size.

The solution of equation 7.21 can be written as

$$\alpha_{i,j} = \frac{\sum_{k,l} p_{(k,l)\to(N_1,N_2)} |Q_{(k,l),(i,j)}|}{|\mathbf{Q}|} \tag{7.22}$$

where $Q_{(k,l),(i,j)}$ is the matrix $\mathbf{Q}$ with the row corresponding to state (k,l) and the column corresponding to state (i,j) removed.

7.8.1 Example with symmetric migration

The equation can be solved numerically in some interesting cases. As an initial example I consider two small subpopulations with symmetric migration. I assume $N_1 = N_2 = 3$, $\mu_{21} = \mu_{12} = 0.2$ and compare this with the single population $N = 6$. Figures 7.6 and 7.7 below show the fixation probabilities for these two cases.

To give a clearer indication of the effect of subdividing the population, figure 7.8 shows the absorption probabilities for a given set of initial conditions in each of the population structures which have the same overall initial frequency.

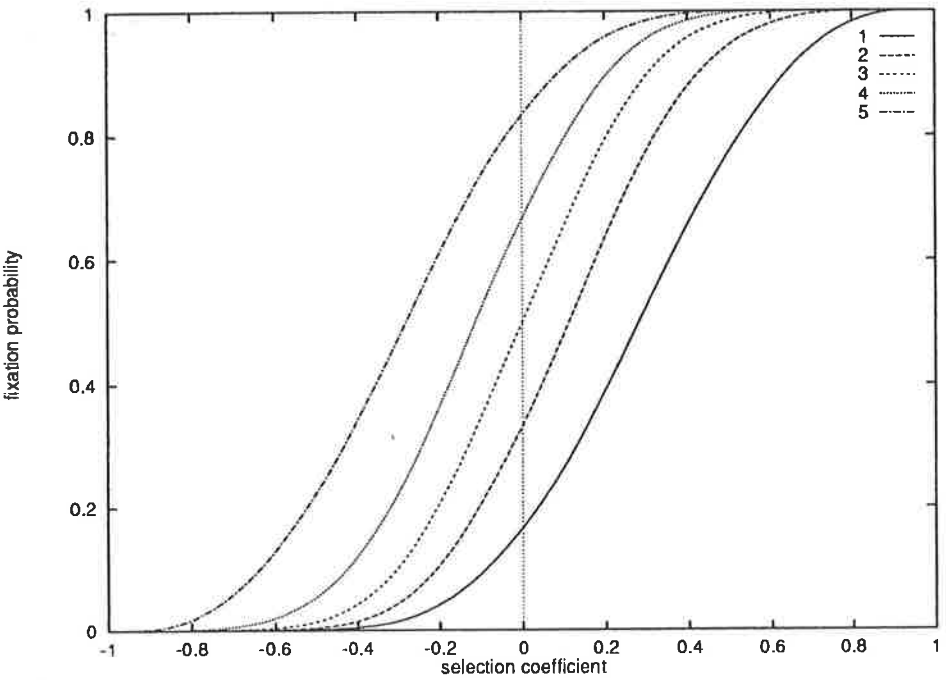


Figure 7.6: Plot of fixation probability as a function of the selection coefficient for $N = 6$.

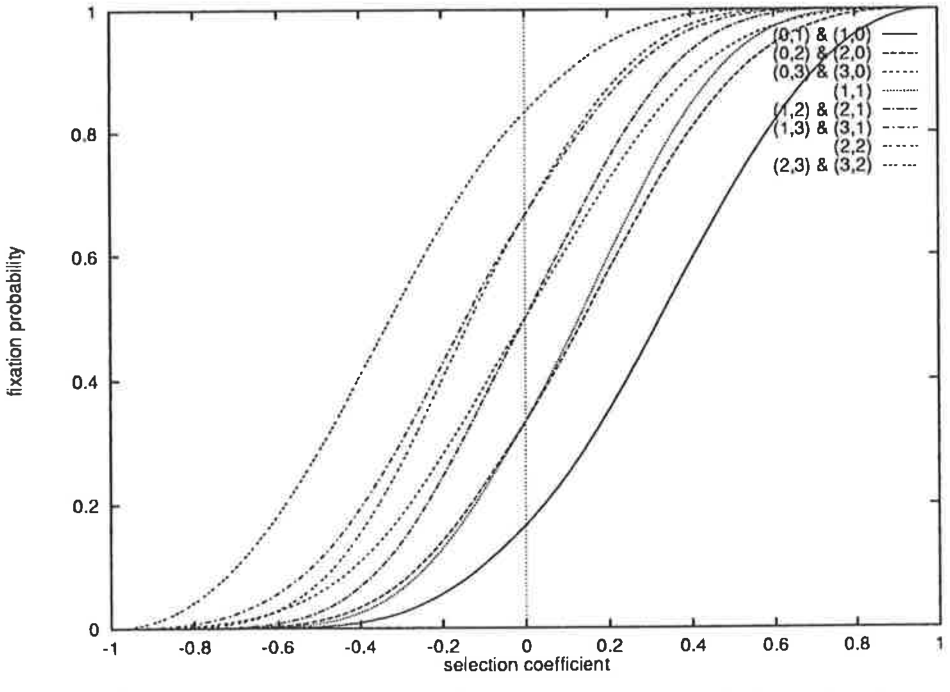


Figure 7.7: Plot of fixation probability as a function of the selection coefficient for $N_1 = N_2 = 3$.

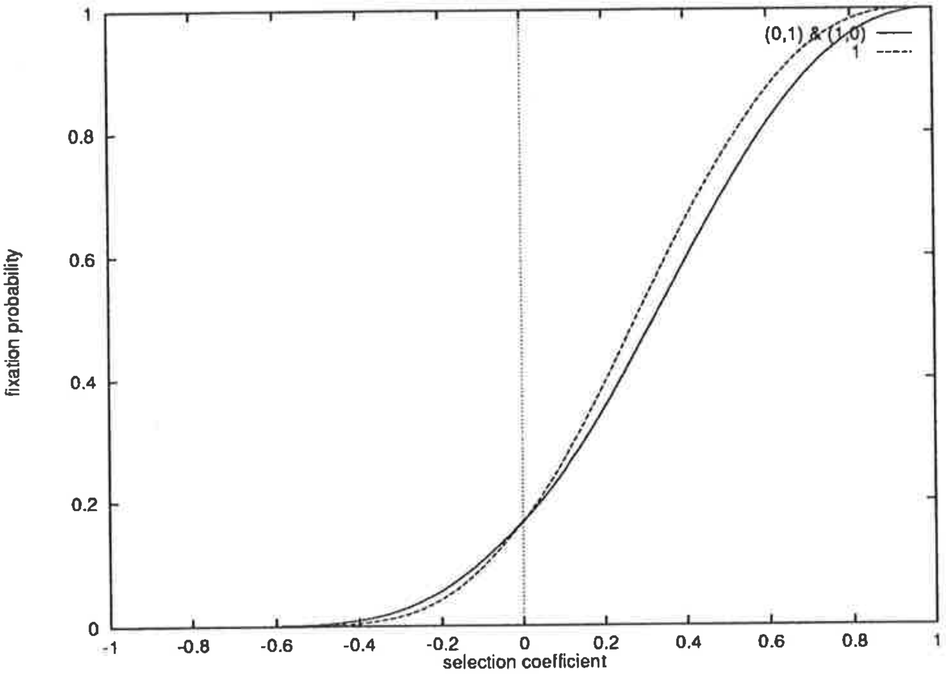


Figure 7.8: Plot of fixation probability as a function of the selection coefficient where there is initially one A_1 allele in a single population with six individuals and two subpopulations each with three individuals. In each case there are six individuals in total.

It is clear that when $s = 0$, that is there is no selection, the subdivision of the population has no effect on fixation probability as determined in chapter 3 and the fixation probability is just equal to the initial frequency in the whole population. The greatest difference between the subdivided and single population case occurs when there are initially few A_1 alleles in the population, in particular when one of the populations has no A_1 alleles.

If there is selection in favour of the A_1 alleles, the probability that one of the smaller subpopulations becomes fixed for the A_1 allele is higher than the probability of the single larger population becoming fixed for the A_1 allel. This is beacuse if both subpopulations initially have one A_1 allele, then the initial frequency in the subpopulation is twice that in the single population and the level of migration in this example is fairly low. Once one of the subpopulations becomes fixed for A_1 and there is selection in favour of this allele, the other subpopulation will also generally become fixed. This is because if one subpopulation is fixed for the A_1 allele and there is a migrant A_2 allele from the other subpopulation, the selection in favour of A_1 will generally ensure that A_1 again becomes fixed in that subpopulation. If there is a migrant from the subpopulation which is fixed for A_1 to the other subpopulation then because it is selectively advantaged it will generally also become fixed in that subpopulation.

A similar argument to the above holds when there are initially a large number of A_1 alleles and selection against A_1. The argument is the same with the roles of A_1 and A_2 reversed. However, when there are initially large numbers of A_1 alleles then the A_1 allele will generally become fixed anyway, so selection in favour of the A_1 allele will have less effect.

For the two population structures considered here even in the most extreme case, subdivision in itself causes relatively small changes in the fixation probability. In absolute terms it is always less than 0.1.

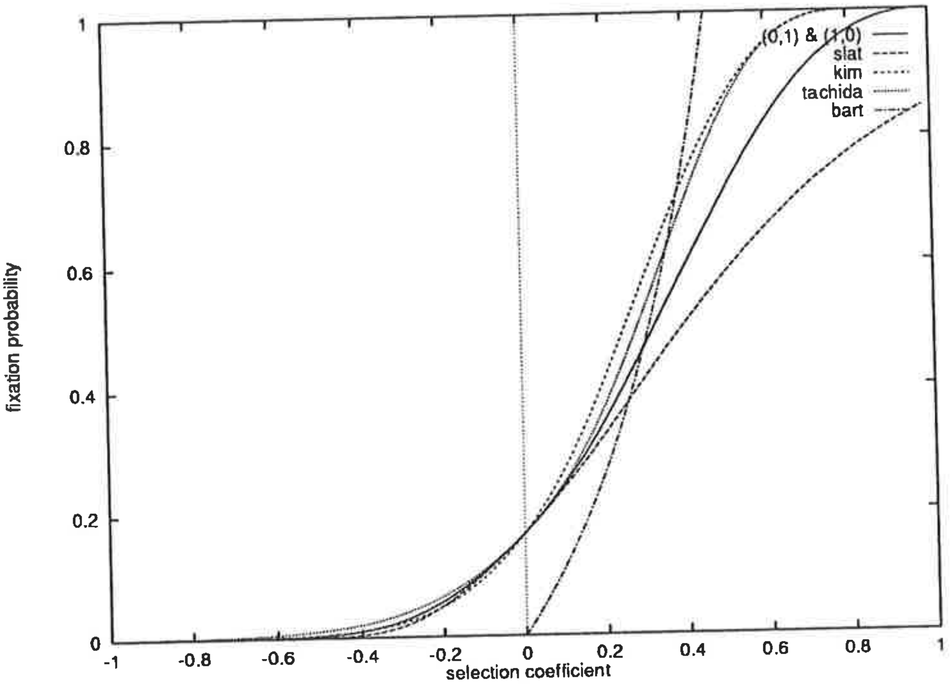


Figure 7.9: Comparison of results of Slatkin (1981), Maruyama (1970a), Tachida and Iizuka (1991) and Barton (1993) with the results of this chapter. The different methods achieve similar results, especially for small s, with the exception of Barton (1993) which appears incorrect.

7.8.2 Comparison with other results

The results from equation 11 of Slatkin (1981) were graphed using additive selection, $2N = 3$ and compared with the case $N = 3$, $\mu_{12} = \mu_{12} = 0.05$ for this model. (Note that Slatkin uses $2N$ as the population size whereas I use N as the population size so these examples are actually equivalent.) These are both compared with Maruyama's (1970a) result for subdivided populations in the high migration limit. With additive selection the probability density, $G(x, N)$, from Slatkin's model becomes $G(x, N) = \exp\{-4Ns_1x\}$. This is then integrated to find $u(N)$, the probability of fixation for an allele in a population of size N when there is initially one allele, and U, the probability that all subpopulations become fixed once one subpopulation has become fixed. The probability of fixation for a mutant gene is then given by $u(N)U$ in the low migration limit and $u(nN)$ in the high migration limit (this is the same result as Maruyama, 1970a). Tachida and Iizuka (1991) give a formula for fixation probability using a probabilistic model in the particular case where migration is asymmetric but low and selection is acting at different strengths in each subpopulation. I consider their model with $\mu_{12} = \mu_{21} = 0.05$ ($m_1 = m_2 = 0.05$ in their notation) and with selection uniform across both subpopulations. There are two subpopulations each of size $N = 3$ and selection is of equal intensity in each subpopulation.

The results are shown in figure 7.9. The results obtained using the model described in this chapter are intermediate between Slatkin's (1981) low and high migration limits for most values of s. Tachida and Iizuka's (1991) model produces results which are quite close to the high migration limit. All of the models show that when $s = 0$, that is there is no selection, subdivision of the population has no effect on the fixation probability. The models generally agree at least in the qualitative effects that subdivision and selection have on the fixation probability.

The models of Slatkin (1981), and Tachida and Iizuka (1991) are very close to the results obtained here when s is small. The results of Maruyama (1970a) (and the high migration limit of Slatkin, 1981) are not very close to the results I obtain, even for small s.

7.9 An approximation

In subsection 5.3.2 it was noted that when there is no selection the fixation probability in the whole population is equal to the weighted average of the fixation probabilities in each subpopulation considered separately. This approximation is tested here for the case where there is selection.

Consider a system where there are two subpopulations each containing three individuals with selection acting. The fixation probabilities for this system are given by equation 7.22. The approximation being used here for the fixation probvability in the population as a whole is the weighted average of the fixation probability in each subpopulation considered separately. In Figure 7.10, the fixation probabilities for this system are compared with the approximation. Clearly the approximation is not very good except when selection is small. The approximation rapidly deteriorates as selection becomes stronger in either direction. The reason that the approximation does not work very well is that for fixation of the non-selected allele to occur in the subdivided population, this allele would have to be fixed in both of these populations. When selection is acting against fixation of A_1 in the whole population, the probability of fixation of in the whole population is closer to the product of the probabilities of fixation in each subpopulation separately than the weighted average of probabilities of fixation in each subpopulation separately. This is because fixation in the whole population is most likely to occur by fixation happening in each subpopulation separately and independently, and the probability of two independent events both occurring is the product of the probabilities of the two events. When selection is acting against fixation of A_1 and one subpopulation becomes fixed for allele A_2 then it is more likely that a migrant from the subpopulation fixed for A_2 will become fixed in the other subpopulation fixed for A_1 than a migrant from the population fixed for A_1 becoming fixed in the subpopulation fixed for A_2.

7.10 Mean absorption times

The general formula derived for the mean absorption time in chapter 4 is still applicable here with the state space and transition probabilities appropriately modified. The state

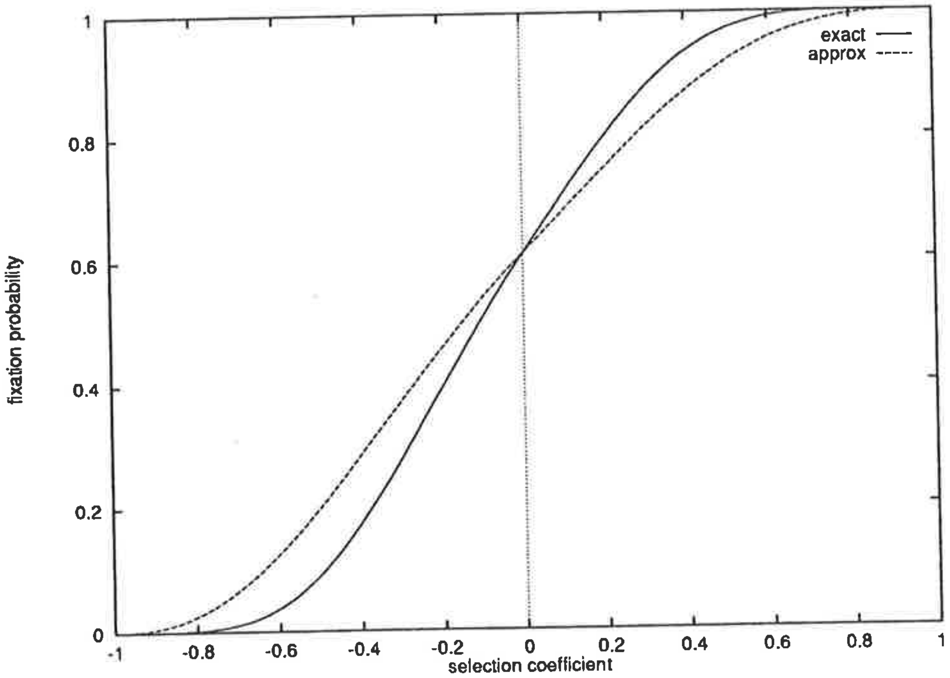


Figure 7.10: Comparison of exact solution of two population fixation probability with approximation from section 7.9 where $x_0 = 1$, $y_0 = 2$, $\mu_{12} = 0.1$, $\mu_{21} = 0.5$. Approximation only works for very small values of s.

space is the set of all pairs of states (x_t, y_t) with the absorbing states removed. This means there are $(N_1+1)(N_2+1)-2$ states in the state space. The formula which was derived in chapter 4 is

$$\mathbf{T} = (\mathbf{I} - \mathbf{P})^{-1}\mathbf{1}$$

where

$$\mathbf{P} = \begin{pmatrix} P_{(0,1)\to(0,1)} & P_{(0,1)\to(0,2)} & \cdots & P_{(0,1)\to(N_1,N_2-1)} \\ P_{(0,2)\to(0,1)} & P_{(0,2)\to(0,2)} & \cdots & P_{(0,2)\to(N_1,N_2-1)} \\ \vdots & \vdots & & \vdots \\ P_{(N_1,N_2-1)\to(0,1)} & P_{(N_1,N_2-1)\to(0,2)} & \cdots & P_{(N_1,N_2-1)\to(N_1,N_2-1)} \end{pmatrix}. \quad (7.23)$$

As in the case of the fixation probabilities, it is generally not possible to solve these equations analytically. They can be solved numerically and again, the case $N_1 = N_2 = 3$, $\mu_{12} = \mu_{21} = 0.2$ is used as an example and compared with the subgle population case with $N = 6$.

From figures 7.11 and 7.12 it is clear that there are dramatically different mean absorption times when the population is subdivided compared with a single population. In general, mean absorption times are greater for the subdivided population than for the equivalent single population. This is because in a subdivided population each subpopulation may be fixed for a different allele for some time until a migrant passes between the two subpopulations and provides the opportunity for absorption in the whole popululation. When migration is lower then this effect is even more noticeable.

If the initial frequency of the A_1 allele in the whole population is less than half, then the mean absorption time is greatest when there is moderate selection in favour of that allele. This is true regardless of the population structure. This is because moderate selection decreases the chance that the population will quickly become fixed for A_2 the most common allele in the population, and increases the chance that the allele will be fixed rather than lost. Because selection is only moderate it will still take a long

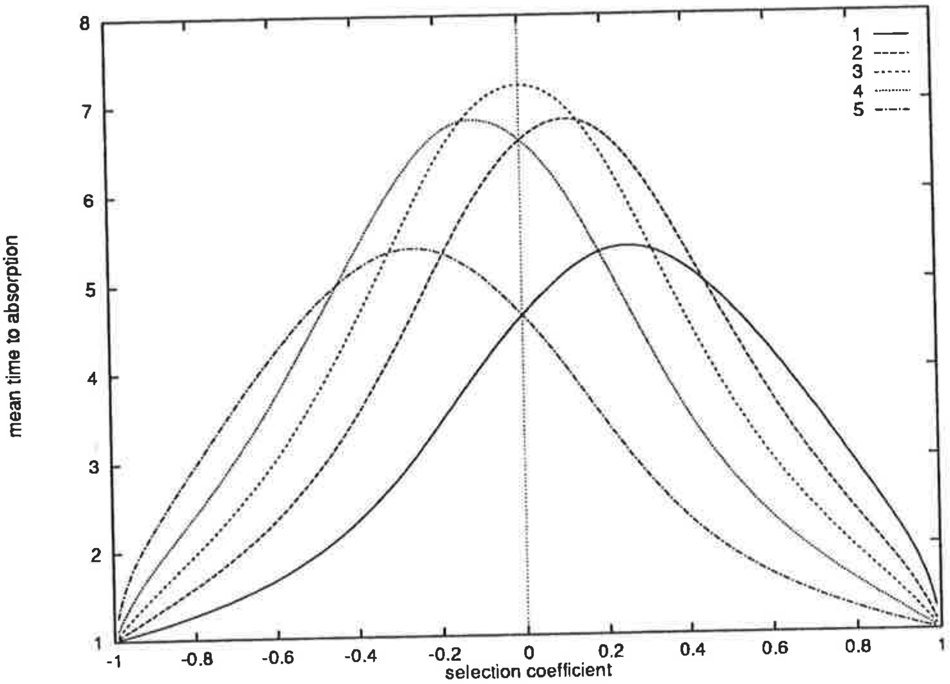


Figure 7.11: Plot of mean absorption time as a function of the selection coefficient for $N = 6$.

time for the allele to become fixed. When selection is strong in favour of the A_1 allele, the mean absorption time is not as high because the allele is more likely to be fixed but the time to absorption will be short as A_2 alleles will quickly be removed from the population by selection.

One interesting phenomenon which occurs when the population is subdivided but not when it is a single population is the mean absorption time having a minimum of 6 generations for initial states where the A_1 allele is initially absent from one of the subpopulations but selection favours that allele.

To show more clearly the effect that subdividing the population has on mean absorption time, figure 7.13 shows the mean absorption time for initial states in the two different population structures with the same overall initial frequency. Generally for high initial frequencies of the A_1 allele, weak selection against this allele will increase the mean absorption time. For both of the situations shown in figure 7.13 the max-

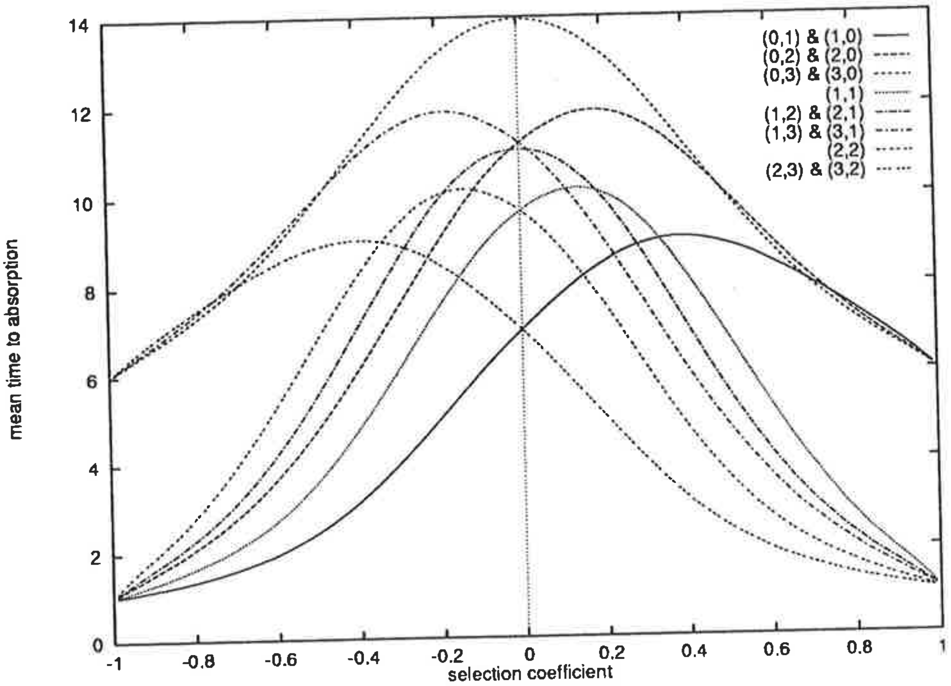


Figure 7.12: Plot of mean absorption time as a function of the selection coefficient for $N_1 = N_2 = 3$ and $\mu_{12} = \mu_{21} = 0.2$.

imum mean absorption time occurs when $s \approx -0.3$. For larger negative values of s the system will very quickly move to loss of the A_1 allele in those cases where it is not fixed by chance in the first generation. The mean absorption time when the population is subdivided will decrease to 6 generations whereas in the single population case it will decrease to 1 generation. This is because in the subdivided population case where s is close to -1, the subpopulation which initially contains the A_2 allele will quickly become fixed for the A_2 allele. It will then take on average 5 generations for a migrant to go from this population to the other population where again it will quickly become fixed. If in the intervening time there is a migrant from the population fixed for A_1 to the population fixed for A_2 this migrant will be selected against and will disappear very quickly.

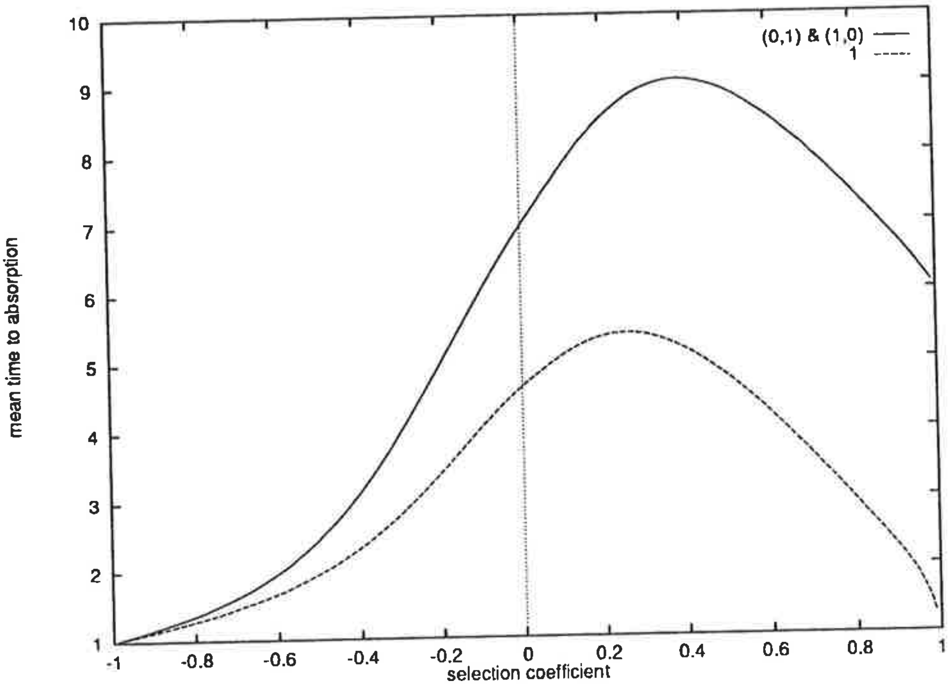


Figure 7.13: Plot of mean absorption time as a function of the selection coefficient for initial frequency 1/6 in each of the single and subdivided population cases.

7.10.1 Comparison with other results

It would be more difficult to compare mean absorption times than it was to compare fixation probabilities as the papers referred to in the fixation probability section (Maruyama, 1970a; Tachida and Iizuka, 1991; Slatkin, 1981) all consider the mean absorption time rather than the mean absorption time which I have considered here. A comparison could be made by suitably adjusting the probability matrix **P** to only consider those situations which result in absorption.

Chapter 8

Variations in population size

8.1 Introduction

A major limitation of the models considered in the preceding chapters is that they all require constant population size. In this chapter I consider cyclical variation in population size – an example of this is seasonal variation in insect numbers, although there are many other examples of populations that vary to a lesser extent. Stochastic variation in population size is more difficult to combine in a population genetic model. Most authors use an average effective population size or use Monte-Carlo simulation to model populations which vary in size over time. Here I will use an exact Wright-Fisher model to test the appropriateness of this model.

In this chapter I will first consider cyclical variation in population size before looking briefly at arbitrary variations in population size. The problem of cyclic variation in population size has generally been dealt with by other authors by calculating an effective population size, N_e which is equivalent to the cycling population "averaged" over time. This concept was first introduced by Wright (1931). The effective population size, N_e, is the size that a population of constant size would have to be in order to have the same fixation probabilities and mean fixation times as the population which varies between two different sizes. The effective population size in this case is equal to the harmonic mean of the two population sizes in the population that varies in size. The formula for the effective population size is given in section 1.8 in the literature review and is

$$N_e^* = 1/\left[\frac{1}{k}\sum_{t=1}^{k}\frac{1}{(N_e)_t}\right]$$

The concept of effective population size also works for cycles of more than two generations.

I consider a single population with a two generation cycle. Two examples are used to compare the results that I obtain with the results obtained using the effective population size method. I will then consider cycles of more than two generations and give equations for the fixation probability and mean absorption time which are analogous to the equations found when the population was of constant size.

Finally I discuss other types of variation in the population size and consider how fixation probabilities may be found for these cases. A formula is given for the fixation probability and mean absorption time in this case which is analogous to the equations found for the case when the population was of constant size. In general it is very difficult to include population dynamics and population genetics in one model without resorting to simulation. I show that in some situations it is possible to obtain matrix equations for the fixation probability when there is stochastic variation in the population size. These equations can be solved for any particular case although it is not possible to find a general interpretation from these equations.

8.2 Description of the model

The model used in this chapter is a single population version of the model described in chapter 2 with population size varying over time. There is one population which cycles between size N_1, at odd numbered generations, and N_2, at even numbered generations. Selection is allowed in the model and is parameterised by the variable $s \in (-1, \infty)$. Selection acts in essentially the same way as in previous chapters (see section 7.3). It is assumed that there are many juveniles born with the ratios different from the current generation because of the action of selection. The breeding stock for the next generation is then sampled from these juveniles. If there are i individuals in the population with the A_1 allele at time t then A_1 juveniles are born in the ratio $(1+s)i : N_t - i$. These juveniles are sampled binomially to form the breeding stock. The function used to determine the probability of an A_1 allele being born in an even numbered generation is

$$\sigma_1(i, s) = \frac{(1+s)i}{N_1 + si}$$

and the function used to determine the probability of an A_1 allele being born in an odd numbered generation is

$$\sigma_2(j, s) = \frac{(1+s)j}{N_2 + sj}.$$

The variable of interest is the number of A_1 alleles present in the population in generation t and this will be represented by X_t.

The transition probabilities in the model are given by

$$X_{t+1} | X_t = j \sim Bin(N_2, \sigma_1(i, s)) \qquad \text{when } t \text{ is odd} \tag{8.1}$$

and

$$X_{t+1} | X_t = i \sim Bin(N_1, \sigma_2(j, s)) \qquad \text{when } t \text{ is even.} \tag{8.2}$$

So

$$\begin{aligned} P_1(i \to j; s) &= \binom{N_1}{j} \sigma_1(i,s)^j (1-\sigma_1(i,s))^{N_1-j} && (8.3) \\ P_2(j \to k; s) &= \binom{N_2}{k} \sigma_1(j,s)^k (1-\sigma_1(j,s))^{N_1-k}. && (8.4) \end{aligned}$$

8.3 Finding the fixation probabilities

Define

$\alpha(i) = P(\text{fixation of } A_1 \mid \text{start with } i \text{ individuals in an odd numbered generation})$

and

$\beta(j) = P(\text{fixation of } A_1 \mid \text{start with } j \text{ individuals in an even numbered generation}).$

Then the Kolmogorov equations are:

$$\begin{aligned} \alpha(i) &= \sum_j P_1(i \to j; s)\beta(j) \\ \beta(j) &= \sum_i P_2(j \to k; s)\alpha(k) \end{aligned}$$

and we know $\alpha(0) = \beta(0) = 0$, $\alpha(N_1) = \beta(N_2) = 1$. The Kolmogorov equations can then be expressed in matrix form as

$$\alpha = \mathbf{P}_1\beta + P_{1N_1} \qquad (8.5)$$

$$\beta = \mathbf{P}_2\alpha + P_{2N_2} \qquad (8.6)$$

and substituting equation 8.6 into equation 8.5 we get

$$\begin{aligned} \alpha &= \mathbf{P}_1(\mathbf{P}_2\alpha + P_{2N_2}) + P_{1N_1} \\ &= (\mathbf{P}_1\mathbf{P}_2)\alpha + (\mathbf{P}_1 P_{2N_2} + P_{1N_1}). && (8.7) \end{aligned}$$

Defining $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2$ and $P_{N_1} = \mathbf{P}_1 P_{2N_2} + P_{1N_1}$, the problem is mathematically the same as the problems discussed in chapter 7 with $\mathbf{P}$ a $N_1 - 1 \times N_1 - 1$ matrix. The matrix P_{N_1} can can be interpreted as the probability that the population goes to fixation in the next two generations. It is the probability that fixation occurs on or before the next generation that the population reaches the same point in the cycle, in this case the next odd numbered generation. Equation 8.7 becomes

$$\alpha = \mathbf{P}\alpha + P_{N_1}. \tag{8.8}$$

This can be solved by first rearranging equation 8.8 and substituting $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ to get

$$\mathbf{Q}\alpha = P_{N_1} \tag{8.9}$$

which has solution

$$\alpha(i) = \frac{\sum_k p_{k,N_1} |Q_{k,i}|}{|\mathbf{Q}|} \tag{8.10}$$

where p_{k,N_1} is the kth element of the vector P_{N_1} and $Q_{k,i}$ is the matrix $\mathbf{Q}$ with row k and column i removed.

The fixation probabilities for even generations, β can be found by multiplying the α by the appropriate matrix, $\mathbf{P}_2$ (see equation 8.6).

8.4 Calculating mean absorption times

The method for finding the mean absorption times is very similar to the method for finding the fixation probabilities.

Define

$T_1(i)$ = mean absorption time starting with i indivduals in an odd numbered generation and

$T_2(j)$ = mean absorption time starting with j indivduals in an even numbered generation.

Then the equations which define the mean absorption time are:

$$\begin{aligned} T_1(i) &= \sum_j P_1(i \to j; s) T_2(j) \\ T_2(j) &= \sum_i P_2(j \to k; s) T_1(k) \end{aligned}$$

for $i, k = 1, \ldots, N_1 - 1$, $j = 1, \ldots, N_2 - 1$ and $P_1(i \to j; s)$ and $P_2(j \to k; s)$ are as defined in equations 8.3 and 8.4. The equations can then be expressed in matrix form as

$$\mathbf{T}_1 = \mathbf{P}_1 \mathbf{T}_2 + \mathbf{1} \tag{8.11}$$

$$\mathbf{T}_2 = \mathbf{P}_2 \mathbf{T}_1 + \mathbf{1} \tag{8.12}$$

and substituting equation 8.12 into equation 8.11 we get

$$\begin{aligned} \mathbf{T}_1 &= \mathbf{P}_1(\mathbf{P}_2 \mathbf{T}_1 + \mathbf{1}) + \mathbf{1} \\ &= (\mathbf{P}_1 \mathbf{P}_2)\mathbf{T}_1 + (\mathbf{P}_1 \mathbf{1} + \mathbf{1}). \end{aligned} \tag{8.13}$$

Defining $\mathbf{P} = \mathbf{P}_1 \mathbf{P}_2$ and $P_f = \mathbf{P}_1 \mathbf{1} + \mathbf{1}$, the problem is mathematically the same as the problems discussed in chapter 7 with $\mathbf{P}$ a $N_1 - 1 \times N_1 - 1$ matrix. Equation 8.13 becomes

$$\mathbf{T}_1 = \mathbf{P}\mathbf{T}_1 + P_f. \tag{8.14}$$

This can be solved by first rearranging equation 8.14 and substituting $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ to get

$$\mathbf{Q}\mathbf{T}_1 = P_f \tag{8.15}$$

which has solution

$$T_1(i) = \frac{\sum_k P_{f,k} |Q_{k,i}|}{|\mathbf{Q}|} \tag{8.16}$$

where $P_{f,k}$ is the kth element of the vector P_f and $Q_{k,i}$ is the matrix $\mathbf{Q}$ with row k and column i removed.

The mean absorption times for even generations, $T_2(j)$, can be found by multiplying the $T_1(i)$ by the appropriate matrix, that is $\mathbf{T}_2 = \mathbf{P}_2\mathbf{T}_1 + \mathbf{1}$.

8.5 Some examples

Example 1: Suppose there is a situation where $N_1 = 12$ and $N_2 = 4$. The fixation probability and mean absorption time are shown in figures 8.1 and 8.2 respectively. They are compared with populations of constant size 4, 6 and 12 and in the fixation probability case with the diffusion approximation solution for the fixation probability. The effective population size, N_e, calculated according to Wright's formula (see section 1.8) is $N_e = 6$ so the constant population with $N = 6$ should in theory be a good approximation to the fluctuating population.

Figure 8.1 shows the fixation probability for the cyclical population that varies between 4 and 12 individuals, compared to a population of constant size 4, a population of constant size 12 and a population of constant size 6 ($= N_e$). Each of these populations is considered with half the initial population having A_1 alleles. The fixation probability for the cyclical population and the population of size 6 both fall inbetween the fixation probability for the population of size 4 and the population of size 12 as would be expected. The diffusion approximation is also generally fairly close except when $s < -0.5$. The diffusion approximation also has a slightly different shape to all of the Wright-Fisher model graphs. The fixation probability for the cyclical population is not exactly equal to the fixation probability for a population of size N_e and behaves like a population slightly smaller than 6 individuals. It appears from this example that N_e gives a useful approximation of fixation probabilities for a varying population, at least

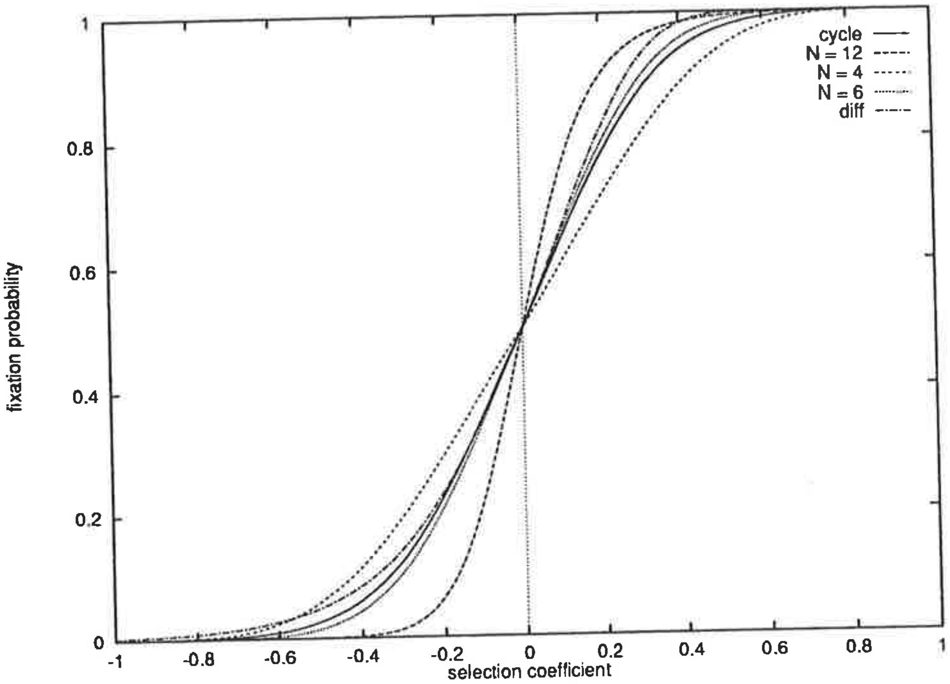


Figure 8.1: Comparison of fixation probabilities of a cyclical population with some constant size populations and a diffusion model with population size 6. In each of the populations, half of the individuals initially have A_1 alleles.

when the population size varies cyclically.

Figure 8.2 shows the mean absorption times for the same size populations as before. This figure shows that using a population of constant size N_e gives a mean absorption time very close to the time obtained when a cyclic population is used. The approximation is better when selection is strong but is still very close when there is no selection. As with fixation probability, it appears that N_e gives a useful approximation of the mean absorption time when the population size varies cyclically.

Example 2: Suppose now that the population varies between $N_1 = 90$ and $N_2 = 10$. Here the magnitude of the cycle is much greater as $N_1 = 9N_2$ whereas in example 1, $N_1 = 3N_2$.

Figure 8.3 shows the fixation probabilities for the cyclic population with $N_1 = 90$ and $N_2 = 10$, the fixation probabilities for populations of constant size 10 and 90,

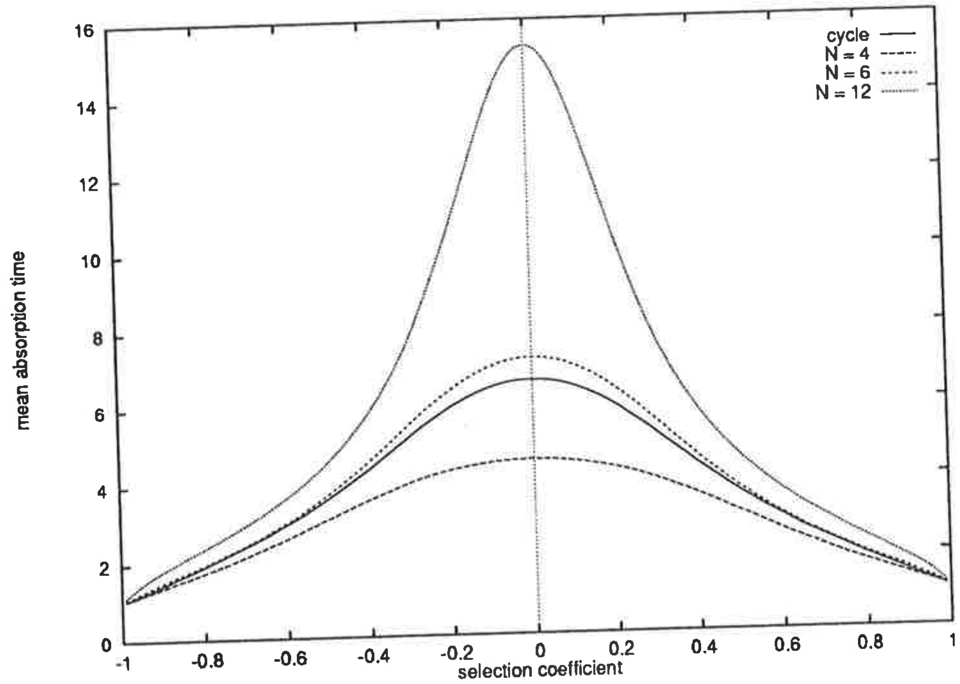


Figure 8.2: Comparison of mean absorption times of a cyclical population with some constant size populations. In each of the populations, half of the individuals initially have A_1 alleles.

the fixation probability for a population of constant size $N_e = 18$ and the diffusion approximation to the fixation probability for a population of size $N_e = 18$. Half of the inidividuals in each model initially have the A_1 allele.

The results here are qualitatively the same as the results for example 1. The fixation probability for the cyclic population is close to the fixation probability for the population of constant size $N_e = 18$. The cyclic population has fixation probability equivalent to a population of constant size slightly smaller than 18. The fixation probability for the diffusion approximation is also close to the fixation probability for the population of constant size N_e but is equivalent to a population of constant size slightly larger than N_e.

Figure 8.4 shows the mean absorption time for the cyclic population and for constant size populations of size 10 and 90, and $N_e = 18$. Once again the mean absorption time for the cyclic population is very close to the mean absorption time for the population of constant size N_e. When selection is weak the cyclic population has mean absorption times slightly less than the population of constant size N_e. So the population is equivalent to a population slightly smaller than 18 ($=N_e$).

This shows that N_e is a reasonable approximation regardless of the relative magnitude of the population cycles or the absolute size of the population involved. However, the examples considered here show that the effective population size is a slight overestimate of the size a population would have to be in order to have the same fixation probability and mean absorption time as a population which varies in size.

8.6 Cycles with a longer period

So far only a cycle of two generations has been considered. This is now generalised to allow cycles of three or more generations. The Kolmogorov equations, given in

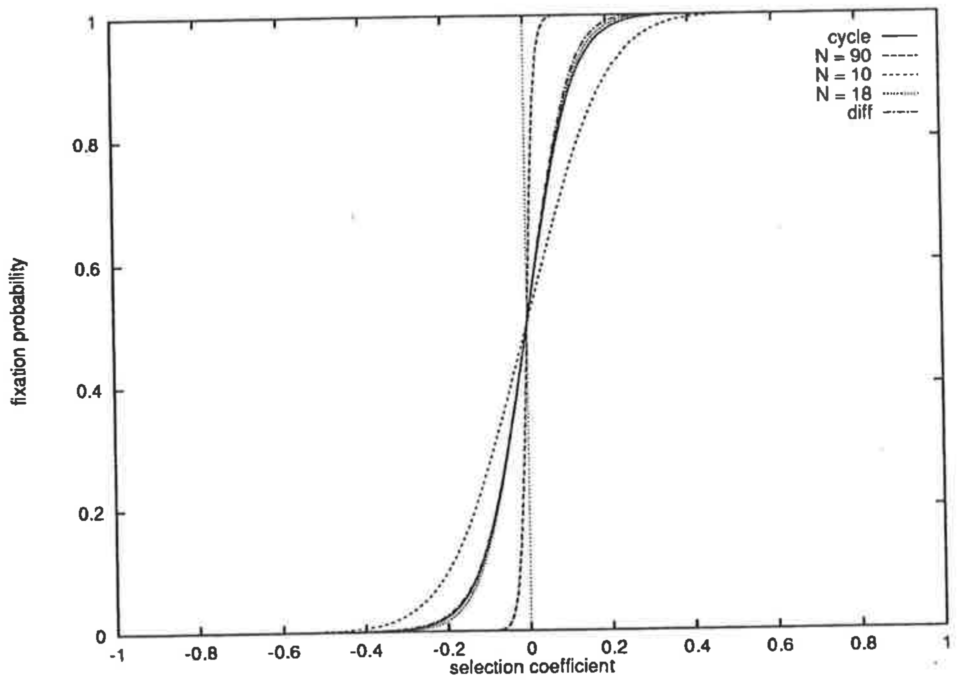


Figure 8.3: Comparison of fixation probabilities of a cyclical population with some constant size populations and a diffusion model with population size 18. In each of the populations, half of the individuals initially have A_1 alleles.

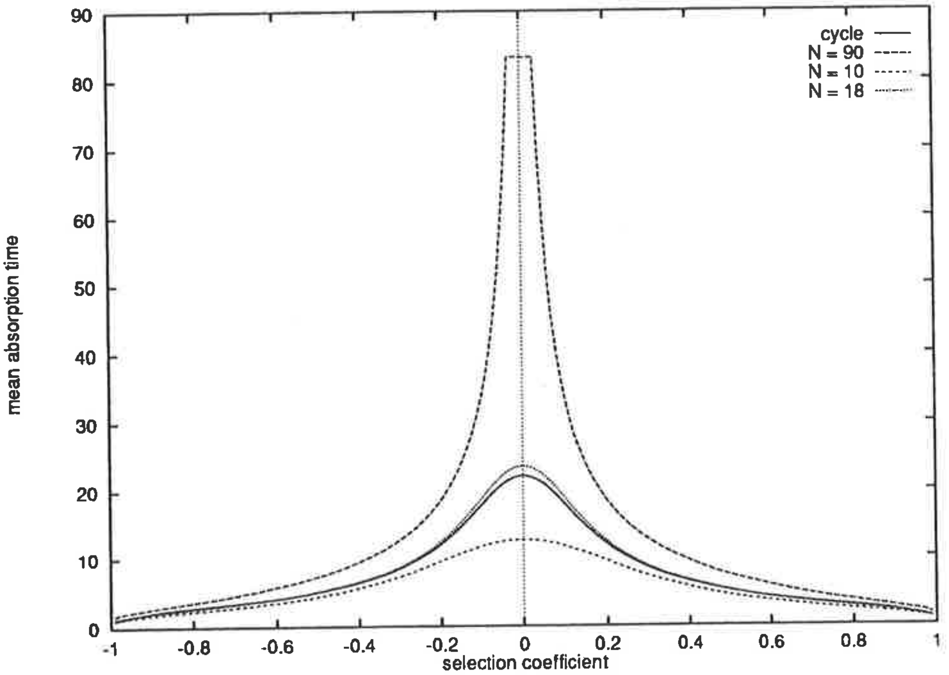


Figure 8.4: Comparison of mean absorption times of a cyclical population with some constant size populations. In each of the populations, half of the individuals initially have A_1 alleles.

matrix form, which must be satisfied by the fixation probabilities when the cycle is n generations are

$$\begin{aligned}
\alpha_1 &= \mathbf{P}_1\alpha_2 + P_{1N_1} \\
\alpha_2 &= \mathbf{P}_2\alpha_3 + P_{2N_2} \\
&\vdots \\
\alpha_{n-1} &= \mathbf{P}_{n-1}\alpha_n + P_{n-1N_{n-1}} \\
\alpha_n &= \mathbf{P}_n\alpha_1 + P_{nN_n},
\end{aligned}$$

where α_t is the vector of fixation probabilities from all of the possible states at time t, $\mathbf{P}_t$ is transition matrix from generation t to generation $t+1$ and P_{t,N_t} is the vector of probabilities from each state in the next generation. These are then solved by substituting the equation for α_n into the equation for α_{n-1} and so on until the equation for α_2 is substituted into the equation for α_1. This equation is of the form

$$\alpha_1 = \mathbf{P}^*\alpha_1 + P^*_{N_1} \tag{8.17}$$

where $\mathbf{P}^* = \mathbf{P}_1\mathbf{P}_2\ldots\mathbf{P}_n$, and $P^*_{N_1} = \sum_{k=1}^{n}\mathbf{P}_1\ldots\mathbf{P}_{k-1}P_{k,N_k}$. Equation 8.17 is a set of $N_1 - 1$ simultaneous equations in $N_1 - 1$ unknowns which can be solved using standard techniques for solving sets of linear equations. I have not given an example here as there is nothing new mathematically which would be demonstrated which was not demonstrated in the case of a two generation cycle.

The derivation for the mean absorption time is analogous to the derivation of fixation probability and gives a set of equations of the form

$$\mathbf{T}_1 = \mathbf{P}^*\mathbf{T}_1 + P^*_f.$$

where $\mathbf{P}^* = \mathbf{P}_1\mathbf{P}_2\ldots\mathbf{P}_n$ and $P^*_f = \sum_{k=1}^{n}\mathbf{P}_1\ldots\mathbf{P}_{k-1}\mathbf{1}$. Again this is a set of $N_1 - 1$ simultaneous equations in $N_1 - 1$ unknowns which can be solved to find the mean absorption times for a specific example.

8.7 Other variations in population size

In this chapter I have considered populations with deteministically cycling population size. In nature populations are more likely to fluctuate randomly in size. These fluctuations occur because of the stochastic nature of birth and death (demographic stochasticity) and year to year changes in conditions (environmental stochasticity including catastrophes such as fire). To incorporate randomly fluctuating population sizes means that there are two random processes in the model. There is the random process which determines the population size in each generation and once the population size has been determined there is the random process that decides how many of these individuals will have the A_1 allele.

It is very difficult to find results for a model like this but there are some useful things that can be found about the fixation probabilities. It is not my aim here to discuss population dynamics and the different types of process that cause variation in population numbers so I will ignore the details of the process which determines population size and use a general unspecified process. Let the number of individuals in the population at time t be N_t and as this is independent of the proportion of A_1 alleles I will assume that the N_t are predetermined. Of these N_t individuals X_t have the A_1 allele.

Let $\alpha_t(i)$ be the probability that the population will eventually go to fixation of the A_1 allele given that there are i A_1 alleles in generation t. The probability that there are j A_1 individuals in generation $t+1$ given there were i A_1 individuals in generation t is

$$P_t(i,j) = \binom{N_{t+1}}{j} \left(\frac{(1+s)i}{N_t+si}\right)^j \left(1 - \frac{(1+s)i}{N_t+si}\right)^{N_{t+1}-j}.$$

Then the $\alpha_t(i)$ satisfy the following equations:

$$\alpha_t(i) = P_t(i,j)\alpha_{t+1}(j) + P_t(i,N_{t+1}) \text{ for } i = 1,\ldots,N_t,\ j = 1,\ldots,N_{t+1}-1.$$

or in matrix notation

$$\alpha_t = \mathbf{P}_t \alpha_{t+1} + P_{t,N} \tag{8.18}$$

where α_t is the vector $(\alpha_t(1), \ldots, \alpha_t(N_t-1))'$, $P_{t,N}$ is the vector $(P_t(1, N_{t+1}), \ldots, P_t(N_t - 1, N_{t+1}))'$ and $\mathbf{P}_t$ is the matrix with (i,j)th element $P_t(i,j)$ and is $N_t - 1 \times N_{t+1} - 1$.

Equation 8.18 can then be used to find a solution for α_1 given a series of N_t or a process for determining such a series. Using repeated substitutions a formula can be found for α_1 in terms of a matrix $\mathbf{P}_n^*$ and a sum of products of terms involving probability matrices $\mathbf{P}_t$ and vectors $P_{t,N}$ for $t = 1, \ldots, n$. The equation is of the form

$$\alpha_1 = \mathbf{P}_n^* \alpha_n + P_{n,N}^*$$

where $\mathbf{P}_n^* = \mathbf{P}_1\mathbf{P}_2 \ldots \mathbf{P}_n$, and $P_{n,N}^* = \sum_{k=1}^{n} \mathbf{P}_1 \ldots \mathbf{P}_{k-1} P_{k,N} = \sum_{k=1}^{n} \mathbf{P}_{k-1}^* P_{k,N}$. If n is much larger than the mean absorption time then $\mathbf{P}_n^*$ will be very close to zero and so α_n is irrelevant to the probability of fixation from time period 1 which will largely be determined by $P_{n,N}^*$. Using any vector such as $(1/N_n, \ldots, 1/N_n)'$ for α_n will provide a reasonable approximation to α_1 as α_n is multiplied by a matrix which is close to zero. This process can be performed using a computer program with matrix multiplication. Again the derivation for the mean absorption time is almost identical to the derivation of the fixation probability and the mean absorption time is given by the formula

$$\mathbf{T}_1 = \mathbf{P}_n^* \mathbf{T}_n + P_{n,f}^*.$$

where $\mathbf{P}_n^* = \mathbf{P}_1\mathbf{P}_2 \ldots \mathbf{P}_n$ and $P_{n,f}^* = \sum_{k=1}^{n} \mathbf{P}_1 \ldots \mathbf{P}_{k-1}\mathbf{1} = \sum_{k=1}^{n} \mathbf{P}_{k-1}\mathbf{1}$. This can be solved to find the mean absorption time for any particular example.

8.8 Conclusions

In this chapter I have examined the effect of variations in population size. Initially I considered a population where the population size oscillates over two generations.

Chapter 9

Summary of results

The aim of this thesis is to build up a model with as few assumptions as possible and, in doing so, to find how each assumption affects the fixation probability and the mean absorption and mean fixation times. To achieve this end I investigate how population structure and migration rates between subpopulations affect the standard measures of loss of heterozygosity. I consider fixation probabilities and mean absorption and fixation times of an allele at a single locus for a haploid organism under several different sets of assumptions. In each case there are two alternative alleles, A_1 and A_2 at the locus of interest. I consider single populations and subdivided populations. In some cases the alleles are selectively neutral and in other models there is selection acting on the allele in question. Finally I consider populations which are not of fixed size.

In chapter 2 the assumptions underlying all of the models given in this thesis are detailed and a physical justification of these assumptions is given.

Initially, in chapters 3 and 4, I consider the case where the population is subdivided into two subpopulations of size N_1 and N_2 with migration between the two subpopulations and no selection acting on either allele. Migration is assumed to be asymmetric.

Using matrix theory I show that in this initial model the fixation probability of the A_1 allele is a weighted sum of its initial frequencies in the two subpopulations. I prove that the fixation probability only depends on the ratio, $\gamma = \mu_{21}/\mu_{12}$, of the average number of migrants and not on the absolute number of migrants. If the initial frequencies of the A_1 allele in subpopulations 1 and 2 are x_0 and y_0 respectively then the fixation probability of the A_1 allele is $\alpha(x_0, y_0) = (x_0 + \gamma y_0)/(N_1 + \gamma N_2)$. When migration is symmetric $\gamma = 1$ and the fixation probability is equal to the initial frequency in the whole population. The result for the case of symmetric migration is compared with the results of other authors who have considered the problem using various other models. I show that the symmetric migration result obtained here is qualitatively consistent with the results of Maruyama (1970a; 1970b; 1972; 1974) using diffusion models and Moran models, Slatkin (1981) using a diffusion model to find upper and lower bounds, and Tachida and Iizuka (1991) using an exact model with low migration and selection. The results of Barton (1993) using diffusion models are not consistent either with my results or the results of the other authors mentioned here.

Calculation of the mean absorption and mean fixation times under the same assumptions is more complicated. This is because quantities involving times to absorption or fixation depend on the absolute value of the average number of migrants μ_{12} and μ_{21} and not just on their ratio. In chapter 4 I find a general set of matrix equations which can be solved to find the mean time to absorption or fixation for any specific case although it is not possible to find a general analytic formula as it is in the case of the fixation probability. These equations are solved analytically for a simple case ($N_1 = N_2 = 2$) with symmetric migration. A larger example is also considered numerically. The examples show that mean absorption and fixation time are largely independent of the level of migration except when migration is at a very low level. When migration is at a low level the mean fixation and absorption times both rise dramatically. I compare these results with the low migration limit discussed by Slatkin (1981). Slatkin showed that in the limit of low migration the mean fixation and absorption time both behave as $1/m$ where m is the migration rate.

I consider an example of asymmetric migration and find that mean absorption times are generally fairly low when there is strong migration in one or both directions. The mean absorption time rises dramatically only when there is low migration in both directions. I also note in chapter 4 that the results obtained for the mean fixation time and the mean absorption time are qualitatively similar to each other.

In chapter 5 I consider the case of three or more subpopulations, still with no selection. Initially, the three and four subpopulation models are analysed using matrix theory. The result can not be presented as simply as in the two population case as the parameter space of the model is larger, but the fixation probability is still a weighted average of the initial frequency in each subpopulation with the weights summing to 1. I consider the effects of some different population structures and migration strengths. The main conclusions, as for the case of two subpopulations, are:

- The fixation probability of the A_1 allele is always a weighted average of its initial frequency in each subpopulation with the weights summing to 1.
- This is equivalent to saying that the fixation probability is a weighted average of the fixation probability in each subpopulation considered separately.
- When migration is symmetric the fixation probability is equal to the inital frequency in the whole population regardless of the population structure.
- The most productive populations (that is, those which produce the greatest excess of migrants over immigrants) have the highest weight in determining the fixation probability. This is intuitively what would be expected.

- It is not just the direct migration, but rather the “net” migration from a subpopulation which determines its influence on fixation probability. So, for example, if a subpopulation sends out only a small number of migrants but these all go to a subpopulation which is a source for the rest of the population, and this forms a large proportion of the immigrants to the source subpopulation, then this subpopulation will have a strong influence on the fixation probability in the whole population.

- A central subpopulation does not necessarily have a greater influence on fixation probability than edge subpopulations, unless it is also a source subpopulation. This result is counter-intuitive as it would seem that central subpopulations should have a greater impact. However, although the central subpopulations generally send migrants to more subpopulations, they also receive immigrants from more subpopulations. This is true unless, for example, the central subpopulation is of higher quality habitat than edge subpopulations in which case the central subpopulation is more productive than edge subpopulations and hence will have a greater influence on fixation probabilty.

Mean fixation and absorption time are not considered for the three and four subpopulation cases.

In chapter 6 I consider the general problem for an arbitrary number of subpopulations, n. Using graph theory I find a general analytic formula for the fixation probability for any population structure. The fixation probability is again a weighted average of the initial frequency in each subpopulation with the weights summing to 1. As in the specific cases discussed in earlier chapters, this can be thought of as the weighted average of the fixation probability for the A_1 allele in each subpopulation considered separately. I prove for the general case that when migration is symmetric the fixation probability is equal to the initial frequency in the whole population, and give a geometrical interpretation of the result when migration is not symmetric using graph theory.

I find formulae for the fixation probabilities in several examples of general population structures including the situation where the subpopulations are arranged with the minimal number of links between the subpopulations (for example in a straight line). As the number of subpopulations increases the state space of the model increases and it becomes increasingly difficult to interpret the results. However it can be seen that when migration is not symmetric the populations which produce the most migrants will have the greatest effect on determining the fixation probability because of the structure of the weights. The geographical location of a subpopulation in the population structure, that is whether they are on the outside or the inside of the structure, is not important in determining the fixation probability. The position of a subpopulation relative to the most productive subpopulations is important in determining the fixation probability. Net exporters of individuals (sources) are most important in determining the fixation probability, but if a subpopulation provides a large proportion of the total immigrants to a subpopulation which is a net exporter to the rest of the population then it will be important even if it is not a large exporter of juveniles itself. This is because although this population is not directly spreading individuals widely, it is indirectly having its individuals spread widely through the source subpopulation, increasing its importance to the fixation probability.

In chapter 7 I introduce selection into the model with only a single panmictic population of size N. When selection is introduced into the model the fixation probability is no longer equal to the initial frequency and the problem of finding the fixation probability of an allele becomes mathematically much more difficult.

Initially I consider populations of size two and three and find explicit analytic solutions for the fixation probability in terms of the model parameters and display these graphically. Although it is theoretically possible to find explicit analytic solutions for any value of N it becomes mathematically very difficult for larger populations. A general matrix formula for the fixation probability is found although it is less explicit than the formula found in the case of no selection. This formula can be solved numerically for any given value of N.

I also find a matrix formula for the mean absorption time. I solve these explicitly for the cases $N = 2$ and $N = 3$ and give solutions in terms of the model parameters and display these graphically. Once again it is theoretically possible to find an explicit formula for any value of N. A general formula is given which can be solved numerically for any value of N.

These results are compared with other models which include selection – two approximate Wright-Fisher models (Moran, 1960), a diffusion model (Kimura, 1962) and a Moran model (Moran, 1958). The Moran model approach gives very accurate results for the fixation probability even for small values of N. This model and the Moran model both give results which are intermediate between the two approximate Wright-Fisher models.

In the second half of chapter 7, I modify the subdivided population model of chapters 3 and 4 and combine this with the selection model developed in the first half of chapter 7 to create a model which includes selection in a subdivided population model. I derive general matrix formulae which the fixation probability and the mean absorption and mean fixation time must satisfy, then simplify these equations and find a general matrix formula for the fixation probability and the mean absorption time. Due to the large number of variables involved it is difficult to give a general interpretation of these equations. The matrix forumulae are solved numerically for some examples and compared with the results of Slatkin (1981), Maruyama (1970a) and

Tachida and Iizuka (1991). Each of these authors gives results which are qualitatively very similar. The results obtained in this chapter are intermediate between the results of the other authors mentioned.

I also solve the equations for mean absorption time numerically for some interesting cases. As in the case of no selection, mean absorption times are generally higher for the subdivided population than for the single population of the same size. The graphs of mean absorption times for the subdivided and the single population show similar qualitative behaviour. The main difference is that when there is strong selection in favour of an allele which is initially absent from one of the populations there is a delay in fixation caused by having to wait for that allele to appear in the population which it is currently absent from. This shows on the graph as an asymptote to $1 + 1/m$ where m is the migration rate. This is another example of Slatkin's (1981) low migration limit.

I show that when there is selection, the fixation probability is not equal to the weighted sum of the fixation probabilities in each subpopulation considered seperately as it is in the case of no selection.

In chapter 8 I remove the assumption that populations must be constant in size. I consider a single population although this model could be extended to two or more subpopulations. Initially I allow the population to vary cyclically between two sizes N_1 and N_2. General matrix equations are given for the fixation probabilities for this case. These equations are solved numerically for two specific cases and these results are compared to the results that would be obtained if a population of constant size equal to the effective population size, N_e, (Wright, 1931) was used.

Similarly for the mean absorption time, I solve the equations for two specific examples and compare these results to the results that would be obtained if a population of constant size equal to the effective population size was used.

I find that an exact Wright-Fisher model with the effective population size, N_e, is a reasonable approximation to the results obtained when the cyclic population is used. The effective population size given by Wright (1931) appears to be an underestimate of the true effective population size.

Following this I give general equations for the fixation probability and mean absorption and mean fixation times for populations with cycles longer than two generations. No attempt is made to solve these equations although they could be solved numerically for specific examples.

I also consider a population which fluctuates in size as a consequence of demographic stochasticity (or environmental variation) where the variation is not cyclic but is governed by some general random process. I assume that the gene frequency has no effect on the process that determines the population size, so effectively the population size at each generation is determined without reference to the genetic process. This is because the aim of this thesis is to consider the effect of genetic processes, and the population dynamic process at work is beyond the scope of this thesis. I find general matrix formulae for the fixation probability and the mean absorption and mean fixation time. No attempt is made to solve these equations for any specific cases, however, these equations could be solved numerically for any given situation of interest.

9.1 Conclusions

In conclusion, I show the following results in this thesis. When symmetric migration is allowed between the subpopulations, the fixation probabilty is independent of the population structure and is equal to the initial frequency in the population considered as a whole. When asymmetric migration is allowed between subpopulations, the structure of the population is important in determining the fixation probabilities. The absolute magnitude of migration between populations is not important in determining the fixation probability, rather it is the relative magnitudes of the level of migration between the subpopulations that is important in determining the fixation probability. The fixation probability is equal to the weighted average of the initial frequencies in each of the subpopulations, or equivalently, the weighted average of the fixation proabability on each subpopulation considered separately.

In general, those populations which are sources (defined here to be net exporters of individuals) have a greater effect on the fixation probability than those which are sinks (defined here to be net importers of individuals). If a subpopulation is close to a source subpopulation then it can have a larger impact on the fixation probability than it otherwise would if it is a large supplier of immigrants to the source subpopulation. With this in mind, a subpopulation is not inherently more important in determining the fixation probabaility just because it is a central subpopulation to the whole subpopulation structure. Note that some authors use the terms source and sink to have a more specific meaning – a source is a net exporter of individuals because of high fecundity (see Pulliam, 1988; Tuck and Possingham, 1994).

When there is selection acting in the population, the situation can not be so easily summarised. The fixation probability is no longer equal to a weighted sum of the initial frequency in each subpopulation, nor is it equal to a weighted sum of the fixation probability in each subpopulation considered separately. The results for the fixation probability in a single subpopulation are consistent with the results from Moran's model. Subdividing the population slightly reduces the effect that selection has on the fxation probability. That is, as selection is strengthened, the fixation probability for the single population is more extreme than for the subdivided population. Subdivision of a population also increases the mean time to absorption, approximately doubling it in the example given where there are six individuals in the population.

When the size of a population varies over time, the effective population size, N_e, is a reasonable approximation of the size that a population would have to be if it was constant in size and had the same mean fixation times as the varying population. In each of the examples given here it was a slight overestimate (of approximately 5-10%) when compared to the exact model.

Each of these factors – subdivision, selection and variation in size – has been examined individually to see how it affects the fixation probaility and the mean absorption and fixation times in a population. It would be virtually impossible to combine all of these in a single model without resorting to simulation. By examining each of these factors separately, it has been possible to qualitatively assess what might happen in a model which combines all of these factors.

9.2 Implications for genetic management in remnant populations

These conclusions have several important consequences for management of remnant subpopulations in order to maintain genetic diversity when migration between the remnant subpopulations is not symmetric. The implications for managers can be summarised in the following four rules of thumb:

- A subdivided population loses neutral genes at a slower rate than a single population with the same number of individuals;
- In a subdivided population, selection has less effect on fixation probabilities than in a single population of the same size, even when the habitat is homogeneous;
- If a population varies in size then it is the effective size of the population which is important. The effective population size is generally very close to the minimum

population size;

- Source subpopulations dominate the genetics of a subdivided population. This means that to conserve the genetic diversity from sink subpopulations, it may be necessary to transfer individuals from the sink subpopulation to the source subpopulation in order to even out the migration and reduce the domination of the source subpopulation.

Bibliography

Barton, N. H. (1993). The probability of fixation of a favoured allele in a subdivided population, *Genetical Research* **62**, 149–157.

Barton, N. H. and Whitlock, M. C. (1997). The evolution of metapopulations. In Hanski, I. and Gilpin, M. E., editors, *Metapopulaiton biology: ecology genetics and evolution* Academic Press.

Burton, R. S., Feldman, M. W., and Curtsinger, J. W. (1979). Population genetics of *tigriopus californicus* (Copepoda: Harpaticoida): I. Population structure along the California couast, *Marine Ecology Progress Series* **1**, 29 –39.

Campbell, N. J. H., Geddes, M. C., and Adams, M. (1994). Genetic variation in yabbies, *cherax destructor* and *c. albidus* (Crustacea: Decapoda: Parastacidae), indicates the presence of a single, highly sub-structured species, *Australian Journal of Zoology* **42**, 745–760.

Chen, W. (1971). "Applied Graph Theory," North-Holland Publishing Company, Amsterdam.

Darwin, C. (1859). "On The Origin Of Species By Means Of Natural Selection, Or The Preservation Of Favoured Races In The Struggle For Life," John Murray, London.

Dawkins, R. (1976). "The selfish gene," Oxford University Press.

Dawkins, R. (1982). "The extended phenotype," W. H. Freeman, Oxford.

Downes, S. J., Handasyde, K. A., and Elgar, M. A. (1997). The use of corridors by mammals in fragmented Australian eucalypt forests, *Conservation Biology* **11**, 718 – 726.

Ewens, W. J. (1963). The mean time for absorption in a process of genetic type, *Journal of the Australian Mathematical Society* **3**, 375–383.

Ewens, W. J. (1967). The probability of survival of a new mutant in a fluctuating environment, *Heredity* **22**, 438–443.

Ewens, W. J. (1982). On the concept of the effective population size, *Theoretical Population Biology* **21**, 373–378.

Falk, D. A. (1990). Integrated strategies for conserving plant genetic diversity, *Annals of the Missouri Botanical Gardens* **77**, 38–47.

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

Gale, J. S. (1990). "Theoretical population genetics," Unwin Hyman, London.

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

Hanski, I. and Gilpin, M. (1991). Metapopulation dynamics: brief history and conceptual domain, *Biological Hournal of the Linnean Society* **42**, 3–16.

Hastings, A. M. and Harrison, S. (1994). Metapopulation dynamics and genetics, *Annual Review of Ecology and Systematics* **25**, 167 – 188.

Hickey, R. J., Vincent, M. A., and Guttman, S. I. (1991). Genetic variation in running buffalo clover (*trifolium stoloniferum*, Fabaceae), *Conservation Biology* **5**, 309–316.

Hill, W. G. (1974). Prediction and evaluation of response to selection with overlapping generations, *Animal Production* **18**, 117–140.

Jensen, L. (1973). Random selection advantages of genes and their probability of fixation, *Genetical Research* **21**, 215–219.

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

Kimura, M. (1983). “The neutral theory of molecular evolution,” Cambridge University Press.

Kimura, M. and Ohta, T. (1969a). The average number of generations until extinction of an individual mutant gene in a finite population, *Genetics* **63**, 701–709.

Kimura, M. and Ohta, T. (1969b). The average number of generations until fixation of a mutant gene in a finite population, *Genetics* **61**, 763–771.

Lamberson, R. H., McKelvey, K., Noon, B. R., and Voss, C. (1992). A dynamic analysis of northern spotted owl viability in a fragmented forest landscape, *Conservation Biology* **6**, 505 – 512.

Lande, R. (1988). Genetics and demography in biological conservation, *Science* **241**, 1455–1460.

Lande, R. and Barrowclough, G. F. (1987). Effective population size, genetic variation, and their use in population management. In Soule, M. E., editor, *Viable populations for conservation* pages 87 – 124 Cambridge University Press.

Leary, R. F., Allendorf, F. W., and Forbes, S. H. (1993). Conservation genetics of bull trout in the Columbia and Klamath River drainages, *Conservation Biology* **7**, 856–865.

Margules, C. R. and Austin, M. P. (1990). “Nature conservation: cost effective biological surveys and data analysis,” CSIRO.

Maruyama, T. (1970a). On the fixation probability of mutant genes in a subdivided population, *Genetical Research* **15**, 221–225.

Maruyama, T. (1970b). On the rate of decrease of heterozygosity in circular stepping stone models of populations, *Theoretical Population Biology* **1**, 101–119.

Maruyama, T. (1972). Some invariant properties of a geographically structured population: distribution of heterozygotes under irreversible mutation, *Genetical Research*

20, 141–149.

Maruyama, T. (1974). A simple proof that certain quantities are independent of the geographical structure of population, *Theoretical Population Biology* **5**, 148–154.

Maruyama, T. (1977). "Stochastic problems in population genetics," Springer-Verlag.

Maynard-Smith, J. (1982). "Evolution and the theory of games," Cambridge University Press.

McKelvey, K., Noon, B. R., and Lamberson, R. H. (1992). Conservation planning for species occupying fragmented landscapes: the case of the northern spotted owl. In Kareiva, P. M., Kingsolver, J. G., and Huey, R. B., editors, *Biotic Interactions and global change*, pages 424 – 450 Sinauer.

Meffe, G. K. and Vrijenhoek, R. C. (1988). Conservation genetics of desert fishes, *Conservation Biology* **2**, 157 – 169.

Moran, P. A. P. (1958). Random processes in genetics, *Proceedings of the Cambridge Philological society* **54**, 60–71.

Moran, P. A. P. (1960). The survival of a mutant gene under selection. II, *Journal of the Australian Mathematical Society* **1**, 485–491.

Moran, P. A. P. (1962). "The statistical processes of evolutionary theory," Oxford University Press.

Nei, M. and Murata, M. (1966). Effective population size when fertility is inherited, *Genetical Research* **8**, 257–260.

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

Pollak, E. (1966). On the survival of a gene in a subdivided population, *Journal of Applied Probability* **3**, 142–195.

Possingham, H. P. and Roughgarden, J. (1990). Spatial population dynamics of a marine organism with a complex life cycle, *Ecology* **71**, 973–985.

Pulliam, H. R. (1988). Sources, sinks and population regulation, *American Naturalist* **132**, 652 – 661.

Richards, S. A., Possingham, H. P., and Noye, B. J. (1995). Larval dispersion along a straight coast with tidal currents: complex distribution patterns from a simple model, *Marine Ecology Progress Series* page In Press.

Sampson, J. F., Hopper, S. D., and James, S. H. (1988). Genetic diversity and the conservation of *eucalyptus crucis* Maiden, *Australian Journal of Botany* **36**, 447–460.

Slatkin, M. (1980). The distribution of mutant alleles in a subdivided population, *Genetics* **98**, 503–523.

Slatkin, M. (1981). Fixation probabilities and fixation times in a subdivided population, *Evolution* **35**(3), 477–488.

Tachida, H. and Iizuka, M. (1991). Fixation probability in spatially changing environments, *Genetical Research* **58**, 243–251.

Templeton, A. R., Shaw, K., Routman, E., and Davis, S. K. (1990). The genetic consequences of habitat fragmentation, *Annals of the Missouri Botanical Garden* **77**, 13–27.

Tuck, G. N. and Possingham, H. P. (1994). Optimal harvesting strategies for a metapopulation, *Bulletin of Mathematical Biology* **56**, 107–128.

Walters, J. R. (1991). Application of ecolocical principles to the management of endangered species: the case of the red-cockaded woodpecker, *Annual Review of Ecology and Systematics* **22**, 505 – 523.

Watterson, G. A. (1961). Markov chains with absorbing states: A genetic example, *Annals of Mathematical Statistics* **32**, 716–729.

Winter, K., Escalante, P., Rappole, J. H., Ramos, M. A., Oehlenschlager, R. J., and Warner, D. W. (1997). Periodic migration and lowland forest refugia in a "sedentary" neotropical bird, Wetmore's bush-tanager, *Conservation Biology* **11**, 692 – 697.

Wright, S. (1931). Evolution in Mendelian populations, *Genetics* **16**, 97–159.

Wright, S. (1939). "Statistical genetics in relation to evolution," Hermann, Paris.

Wright, S. (1969). "Evolution and the genetics of populations," volume **2** University of Chicago Press, Chicago.