J. Appl. Prob. **1**, 177-232 (1964)
Printed in Israel

DIFFUSION MODELS IN POPULATION GENETICS

MOTOO KIMURA, *National Institute of Genetics, Mishima, Japan*

CONTENTS

Page

I. Problems and Methods of Population Genetics

1. Introduction 178
2. Changes of gene frequencies as stochastic processes 178
3. The partial differential equation method 181

II. Random Drift in the Narrow Sense

4. Random drift in a small finite population 191
5. An approximate treatment by the angular transformation 199

III. Random Drift in the Wider Sense

6. A population under linear pressure and random sampling of gametes 201
7. Change of gene frequency under selection and random sampling of gametes 203
   7.1 *Genic selection (Case of no dominance)* 203
   7.2 *Case of overdominance* 208
8. Random fluctuation of selection intensities 211

IV. Genetic Equilibrium; Stationary Distributions and Gene Fixation

9. Gene frequency distribution at equilibrium 215
   9.1 *Stationary distribution* 215
   9.2 *Distribution under irreversible mutation* 218
10. Probability of fixation of mutant genes in a population .. 223
   10.1 *Introductory remarks* 223
   10.2 *Single locus* 223
   10.3 *Multiple loci* 228

References 230

Received in revised form 10 June 1964. Contribution No. 453 from the Nationa Institute of Genetics.

I. Problems and Methods of Population Genetics

1. Introduction

Population genetics is that branch of genetics, whose object is the study of the genetical make-up of natural populations. By investigating the laws which govern the genetic structure of natural populations, we intend to clarify the mechanism of evolution.

In a natural population of sexually reproducing species, with only a hundred loci segregating, the number of possible genotypes may be practically infinite, and the genotype of each individual is quite likely to be unique in the entire history of the species. Thus, as an aggregate of individual genotypes, a population is an enormously complicated system, sometimes too complicated to be treated theoretically. On the other hand, in any reasonably large population, the relative proportion of an allele (a particular form of a gene) within the population changes almost continuously with time. This is because, unlike genotypes, each gene reproduces its own kind with complete fidelity except for the very rare event of mutation. As pointed out by Fisher (1953), "the frequencies with which the different genotypes occur define the gene ratios characteristic of the population, so that it is often convenient to consider a natural population not so much as an aggregate of living individuals as an aggregate of gene ratios. Such a change of viewpoint is similar to that familiar in the theory of gases, where the specification of the population of velocities is often more useful than that of a population of particles." This line of investigation was initiated by Fisher (1922) and later elaborated by him (Fisher, 1930), Haldane (cf. 1932) and especially by Wright (1931, and later publications).

In the present paper, I shall review the theoretical works on population genetics, treating the changes of gene frequencies as stochastic processes, and describing these especially by the use of diffusion equations. Since I started my work in this field as a geneticist, the mathematical sophistication of my approach has been rather limited; I cannot escape from this limitation in the present paper, but I hope it will stimulate mathematicians to work in this fascinating field. Indeed, there is much to be done in the refinement and extension of the mathematical methods involved, as is shown by the works of Feller (1951, 1952) and Moran (cf. 1962).

2. Changes of gene frequencies as stochastic processes

From the standpoint of population genetics, the most elementary step in evo-

lution is the change of gene frequencies. Here, gene frequencies mean the proportions of genes in a population. The simplest mathematical approach to this problem is to regard the process of change as deterministic. Such an approach was first used extensively by Haldane in his series of papers starting in 1924 (Haldane, 1924). Strictly speaking, it applies only if a population is infinitely large and is placed in an environment which remains constant or changes in a deterministic way. There are many circumstances in which this is sufficiently realistic as a first approximation. Furthermore, because of its simplicity, this approach is still the most useful, and is often the only manageable one for many problems. In nature, however, the process of change may not be quite deterministic, because of the existence of factors which produce random fluctuation in gene frequencies, of which two different types may be recognized (Wright, 1949). One is the random sampling of gametes in reproduction. The process of change in gene frequency which is due solely to this factor is often called *drift*. However, the term drift in this context is hardly adequate unless the prefix *random* is also attached, and it may be called the random drift in the narrow sense. This factor becomes prominent in a small population. The other type consists of random fluctuations in what Wright called the systematic evolutionary pressures, of which random fluctuation in selection intensity may be especially important. These two types of factors introduce a random element into the process of change in gene frequencies (random drift in the wide sense). Thus, in the present review, we will regard the process of change as a stochastic process, where this means the mathematical formulation of a chance event evolving in time.

So far as I know, this line of investigation was initiated by Fisher (1922). In his paper, Fisher considered the random sampling of gametes as the factor causing random fluctuation in gene frequency and, assuming no selection, he investigated its effect (*Hagedoorn Effect*) on the decrease of variability in a species. Though his treatment was restricted to a quite simple situation, the paper was important in that he introduced the method of partial differential equations in the study of gene frequency distributions in a population. This method, if properly extended, is equivalent to the approach which makes use of the Fokker-Planck equation, later introduced by Wright (1945). Here Fisher used the transformed gene frequency rather than the gene frequency itself. Also he suggested the method of functional equations to study the probability of fixation of an individual mutant gene. As in many other pioneering works, this 1922 paper was not finished and contained some minor errors and ambiguity. Later a more complete treatment was presented (Fisher, 1930) in which the errors were amended and the results were greatly extended. In my opinion this is one of the most beautiful papers ever written on the mathematical theory of population genetics.

In 1931, Wright published his now classical paper "Evolution in Mendelian populations" in which he studied similar problems by his method of integral

equations (Wright, 1931). Since then, Wright has published a number of important papers on the probability distribution of gene frequencies and on the role of random processes in evolution. He has emphasized the importance of proper balance between the directed and the random processes in relation to population structure in evolution (cf. Wright, 1950).

These two authors, as Feller (1951) has remarked, have studied individual problems with great ingenuity, with the result that many limiting probability distributions have been worked out. However, the problem of constructing a model for the entire process of change in the gene frequencies has not been dealt with by these authors.

In the field of the mathematical theory of probability, progress in our knowledge of stochastic processes has been quite extensive since Kolmogorov's fundamental paper (Kolmogorov, 1931). This is doubtless a result of the growing need for the stochastic treatment of problems in diverse fields of modern science. It is not surprising therefore that pioneering attempts to construct a model for the entire process of change in gene frequencies were made by mathematicians. Malécot (1948) who considered "evolution of the probability law in the course of time", especially for the case of mutation pressure and random sampling of gametes, sketched a method by which the solution might be obtained. Goldberg (1950) in his unpublished thesis studied the same case and succeeded in obtaining two solutions of the diffusion equation involved.

From the mathematical standpoint, various types of stochastic processes arise in population genetics. Of special importance is the Markov process which Kolmogorov called stochastically definite. The process of change in gene frequencies in a very small population consisting of a few individuals with non-overlapping generations will most appropriately be treated as a finite Markov chain. The fate of an individual mutant gene in a large population can be treated by the theory of branching processes.

Generally, however, the processes of organic evolution in nature are very slow and the number of individuals involved per generation is very large, so that they may be treated with advantage as a continuous Markov process in space (gene frequency) and time, as will be explained in the next section. Here the Fokker-Planck equation (Kolmogorov forward equation) plays a fundamental role. Using this approach, the problem of constructing a model for the entire process of change in gene frequencies starting from an arbitrary initial frequency was solved for several genetically interesting cases by the present author (Kimura, 1954, 1955 a, b, c, 1956 a, b, 1957, Crow and Kimura, 1956). Also, it has been shown by the author (Kimura, 1957, 1962) that the Kolmogorov backward equation may be used to obtain the probability of fixation of mutant genes in a population.

Recently, the stochastic theory of gene frequency change was used by Robertson (1960) in his theory of limits in artificial selection. He also studied the problem

of selection for heterozygotes in small populations (Robertson, 1962) based on the mathematical work of Miller (1962) who developed a powerful method for evaluating eigenvalues of the Fokker-Planck equation involved.

Usually, the Fokker-Planck equations which appear in population genetics have singularities at the boundaries and a deep mathematical investigation of these was carried out by Feller (1952) who clarified the nature of boundaries by using semi-group theory.

The diffusion equation approach has been used extensively in the study of gene frequency change because of its extreme usefulness. But it is an approximation based on rather intuitive arguments. Therefore, to investigate the conditions under which such approximation may be valid, is an important task for mathematicians; Moran (1958 a and b) introduced two population models, with overlapping and non-overlapping generations for the study of gene frequency distribution in populations. Watterson (1962) obtained sufficient conditions, concerning the change of gene frequency per unit length of time, under which the diffusion approximation is valid, even if the gene frequency does not necessarily form a Markov process. Recently he applied these conditions to unify Moran's two models (Watterson, 1964). Moran's models have also been investigated by Karlin and McGregor (1962). An important contribution was made by Moran (1961) to the problem of gene fixation in a finite population. Though he was able to treat only a very simple genetical situation, his rigorous treatment is important in giving a case where the diffusion approximation introduced by the author (Kimura, 1957) can be checked.

Accuracy of the diffusion approximation method may also be checked numerically by high-speed computer when the population number is very small; a recent study of Ewens (1963) seems to indicate that the approximation is quite good for populations of reasonable size. Furthermore (Ewens, 1964), formulae for the leading terms of the corrections to diffusion approximations can be found.

3. The partial differential equation method

A natural population which plays a significant role as an evolutionary unit should consist of a large number of individuals, and the gene frequencies for these behave practically as continuous variables. Also any change of gene frequencies must in general be very slow by our ordinary time scale. There are certain cases in which relatively rapid changes were observed in polymorphic characters, such as the spread of industrial melanism in moths. However, the typical rate of evolution shown in fossil records is of the order of one-tenth of a darwin, one darwin standing for the rate of change with a factor of e ($= 2.71\cdots$) per million years (Haldane, 1949); this suggests that the change in gene frequencies involved must be correspondingly slow.

For these reasons, the process of change in gene frequency may be treated as a continuous stochastic process; this means roughly that as the time interval

becomes smaller, so also does the amount of change in gene frequency x during that interval. More strictly, the process is called a continuous stochastic process if for any given positive value ε, the probability that the change in x during the time interval $(t, t+\delta t)$ exceeds ε is $o(\delta t)$, i.e. an infinitesimal of higher order than δt. We will assume also that change in gene frequencies is Markovian, that is, the probability distribution of gene frequencies at a given moment t depends on the gene frequencies at a preceding time $t_0 (t_0 < t)$ but not on the previous history which has led to the gene frequencies at t_0.

For the study of this continuous Markov process, one of the most powerful methods available makes use of the Kolmogorov equations (Kolmogorov, 1931). We will first derive the Kolmogorov forward equation as applied to population genetics. Throughout this article, we will assume, unless otherwise stated, a diploid population consisting of a fixed number N of individuals in each generation. Thus, there are $2N$ genes at each locus.

Consider a pair of alleles A_1 and A_2 with respective frequencies x and $1-x$. Let $\phi(p,x;t)$ be the conditional probability density that the gene frequency is x at time t, given that the initial frequency is p at time $t=0$. This gives the transition probability that the gene frequency moves from p to x after time t. With p fixed, $\phi(p,x;t)$ determines a frequency distribution such that when $1/(2N)$ is substituted for dx, $\phi(p,x;t)dx$ gives an approximation to the frequency of the class with gene frequency x $(0 < x < 1)$ at time t, which when expressed in terms of generations ,we may also roughly refer to as the tth generation; this frequency distribution may be denoted by

$$f(x,t) = \phi(p,x;t)\frac{1}{2N} \qquad (0 < x < 1). \tag{3.1}$$

When p is fixed, it will often be omitted so that $\phi(p,x;t)$ is written as $\phi(x,t)$. Also it should be noted that the above relation (3.1) holds only for *unfixed* classes, i.e. for $0 < x < 1$. Frequencies of classes with $x = 0$ or 1 have to be treated separately.

Let $g(\delta x, x; \delta t, t)$ be the probability density that the gene frequency changes from x to $x + \delta x$ during the time interval $(t, t+\delta t)$. Using this probability density, the assumption of a continuous stochastic process may be expressed as

$$\int_{|\delta x|>\varepsilon} g(\delta x, x; \delta t, t)\, d(\delta x) = o(\delta t), \quad (\delta t \to 0), \tag{3.2}$$

where ε is some arbitrary preassigned positive value.

Furthermore, for the process of change in gene frequency, we have

$$\phi(p,x;t+\delta t) = \int \phi(p, x-\delta x; t)\, g(\delta x, x-\delta x; \delta t, t)\, d(\delta x), \tag{3.3}$$

where the integral on the right is taken over all possible values of δx. The above relation is a natural consequence of the assumption that the process is Markovian. The probability that the gene frequency is x at time $t + \delta t$ is the sum total of the probabilities of cases in which the gene frequency is $x - \delta x$ at time t, and the gene frequency increases by δx during the subsequent time interval $(t, t + \delta t)$, with δx taking all possible values. Actually, the above relation is a special form of the Kolmogorov-Chapman equation in the theory of stochastic processes. In this expression, δx may take any value such that $x - \delta x$ lies between 0 and 1, exclusive of the end points. However, because of (3.2), only values in the range $|\delta x| < \varepsilon$ are of any significance.

Expanding the integrand on the right side of (3.3) in terms of δx, we have

$$\phi(p, x - \delta x; t)\, g(\delta x, x - \delta x; \delta t, t)$$

$$= \phi g - \delta x \frac{\partial}{\partial x}(\phi g) + \frac{(\delta x)^2}{2!} \frac{\partial^2}{\partial x^2}(\phi g) - \frac{(\delta x)^3}{3!} \frac{\partial^3}{\partial x^3}(\phi g) + \cdots,$$

where ϕ and g stand for $\phi(p, x; t)$ and $g(\delta x, x; \delta t, t)$ respectively. Thus (3.3) may be expressed as

$$\begin{aligned} \phi(p, x; t + \delta t) &= \phi \int g d(\delta x) \\ &\quad - \frac{\partial}{\partial x}\{\phi \int (\delta x)\, g d(\delta x)\} \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial x^2}\{\phi \int (\delta x)^2 g d(\delta x)\} \\ &\quad - \cdots. \end{aligned} \tag{3.4}$$

Here we have assumed that the orders of summation, integration and differentiation may be interchanged freely.

Noting that

$$\int g d(\delta x) = 1,$$

and transferring the first term on the right side of (3.4) to the left, we have, after dividing both sides by δt,

$$\frac{\phi(p, x; t + \delta t) - \phi(p, x; t)}{\delta t} \tag{3.5}$$

$$= -\frac{\partial}{\partial x}\left\{\phi(p,x;t)\frac{1}{\delta t}\int(\delta x)\,g(\delta x,x;\delta t,t)\,d(\delta x)\right\}$$
$$+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left\{\phi(p,x;t)\frac{1}{\delta t}\int(\delta x)^2\,g(\delta x,x:\delta t,t)\,d(\delta x)\right\}$$
$$-\frac{1}{3!}\frac{\partial^3}{\partial x^3}\left\{\phi(p,x;t)\frac{1}{\delta t}\int(\delta x)^3\,g(\delta x,x;\delta t,t)\,d(\delta x)\right\}$$
$$+\cdots.$$

Let

$$\lim_{\delta t\to 0}\frac{1}{\delta t}\int(\delta x)\,g(\delta x,x;\delta t,t)\,d(\delta x)=M(x,t), \tag{3.6}$$

$$\lim_{\delta t\to 0}\frac{1}{\delta t}\int(\delta x)^2\,g(\delta x,x;\delta t,t)\,d(\delta x)=V(x,t), \tag{3.7}$$

and assume that

$$\lim_{\delta t\to 0}\frac{1}{\delta t}\int(\delta x)^n g(\delta x,x;\delta t,t)\,d(\delta x)=0 \tag{3.8}$$

for $n\geqq 3$. Then we have

$$\frac{\partial\phi(p,x;t)}{\partial t}=\frac{1}{2}\frac{\partial^2}{\partial x^2}\{V(x,t)\phi(p,x;t)\}-\frac{\partial}{\partial x}\{M(x,t)\phi(p,x;t)\} \tag{3.9}$$

where $M(x,t)$ and $V(x,t)$ refer to the first and second moments of δx during the infinitesimal time interval $(t,t+\delta t)$.

In practice, however, quantities such as mutation rates, rate of migration, intensity of selection, and effect of random sampling of gametes which determine δx are all measured with one generation as a time unit and the limiting rate with $\delta t\to 0$ can only be obtained by extrapolation. So we will replace $M(x,t)$ and $V(x,t)$ in the above equation by $M_{\delta x}$ and $V_{\delta x}$, the mean and variance of the change in gene frequency per generation (δt corresponding to one generation). Thus we obtain

$$\frac{\partial\phi}{\partial t}=\frac{1}{2}\frac{\partial^2}{\partial x^2}(V_{\delta x}\phi)-\frac{\partial}{\partial x}(M_{\delta x}\phi). \tag{3.10}$$

Such an equation, as given in (3.9), is called the Kolmogorov forward equation by mathematicians. It is also called the Fokker-Planck equation by physicists. Actually, Fokker derived the steady state form in 1914 and Planck (1917) later extended it to a quite general form, though rigorous mathematical formuations were first given by Kolmogorov (1931).

The above derivation leading to equation (3.9), is rather formal. More rigorous derivations may be found in the mathematical literature, such as Kolmogorov's paper (1931). On the other hand, the above derivation may be too formal for most biologists to see the physical meaning of the terms involved. A less rigorous but very elementary derivation of the equation was devised by the author, based on the geometrical interpretation of the process involved (Kimura, 1955 c). It was shown that the first and second terms on the right side of (3.10) give the rates of change in the probability distribution due respectively to random fluctuation and systematic pressures. It was also pointed out that the variance

$$E\{(\delta x)^2\} - \{E(\delta x)\}^2$$

rather than the second moment, i.e. $E\{(\delta x)^2\}$ should be used for $V_{\delta x}$ in (3.10), when the equation is applied to actual population genetics problems. This is based on the consideration that (3.10) should give the deterministic process correctly when there is no random fluctuation, i.e. in the limit when $V_{\delta x} = 0$.

Since the gene frequency x lies between 0 and 1 in general, the process of change in gene frequency in a population through time is represented as the stochastic movement of a point x on the closed real interval $[0,1]$. The equation (3.10) can describe this movement at least on the open interval $(0,1)$. We will now show that

$$P(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\{V_{\delta x}\phi(x,t)\} + M_{\delta x}\phi(x,t) \tag{3.11}$$

which satisfies

$$\frac{\partial \phi(x,t)}{\partial t} = -\frac{\partial P(x,t)}{\partial x} \tag{3.12}$$

represents the rate (per generation) of net flow of probability across the point x. First, let us consider the amount of probability which flows over the point x in a positive direction during the time interval of length δt. The contribution of the class with gene frequency ξ to this is (see Figure 1):

$$\phi(p,\xi;t)d\xi \int_{\delta\xi > x-\xi} g(\delta\xi,\xi;\delta t,t)\,d(\delta\xi) \quad (\xi < x),$$

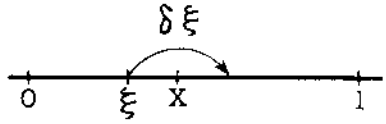


Figure 1
Probability flux across the point x.

and the total amount of probability flow in the positive direction, denoted by $P_+(x)\delta t$, is the sum of all contributions from the class at the left of x:

$$P_+(x)\delta t = \int_{\xi<x} \phi(p,\xi;t)d\xi \int_{\delta\xi>x-\xi} g(\delta\xi,\xi;\delta t,t)\,d(\delta\xi)$$

$$= \int_{\delta\xi>0} d(\delta\xi) \qquad \int_{x-\delta\xi}^{x} \phi(p,\xi;t)\,g(\delta\xi,\xi;\delta t,t)d\xi\,.$$

Let $\xi = x+\eta$ and expand the integrand in terms of η, or $\xi - x$; we have

$$\phi(p,\xi;t)\,g(\delta\xi,\xi;\delta t,t) = \phi g + \eta\frac{\partial(\phi g)}{\partial x} + \cdots,$$

where ϕ and g respectively denote $\phi(p,x;t)$ and $g(\delta\xi,x;\delta t,t)$.

Thus

$$P_+(x)\delta t = \int_{\delta\xi>0} d(\delta\xi)\int_{-\delta\xi}^{0} \left\{\phi g + \eta\frac{\partial(\phi g)}{\partial x} + \cdots\right\} d\eta$$

(3.13)

$$= \int_{\delta\xi>0} \left\{\delta\xi\cdot\phi g - \frac{(\delta\xi)^2}{2}\,\frac{\partial(\phi g)}{\partial x} + \cdots\right\} d(\delta\xi).$$

Similarly, the total amount of probability which flows over the point x in the opposite direction is

$$P_-(x)\delta t = \int_{\xi>x} \phi(p,\xi;t)d\xi \int_{\delta\xi<x-\xi} g(\delta\xi,\xi;\delta t,t)\,d(\delta\xi)$$

$$= \int_{\delta\xi<0} d(\delta\xi) \int_{x}^{x-\delta\xi} \phi(p,\xi;t)\,g(\delta\xi,\xi;\delta t,t)d\xi$$

(3.14)

$$= \int_{\delta\xi<0} d(\delta\xi) \int_{0}^{-\delta\xi} \left\{\phi g + \eta\,\frac{\partial(\phi g)}{\partial x} + \cdots\right\} d\eta$$

$$= \int_{\delta\xi<0} \left\{(-\delta\xi)\phi g + \frac{(\delta\xi)^2}{2}\,\frac{\partial(\phi g)}{\partial x} + \cdots\right\} d(\delta\xi).$$

Thus the net amount of probability which flows past the point x during the time interval $(t, t+\delta t)$ is

$$\begin{aligned} P(x,t)\delta t &= P_+(x)\delta t - P_-(x)\delta t \\ &= \int(\delta\xi)\phi g d(\delta\xi) - \frac{1}{2}\int(\delta\xi)^2\frac{\partial(\phi g)}{\partial x}\,d(\delta\xi) + \cdots \\ &= \phi\int\delta\xi g(\delta\xi,x;dt,t)d(\delta\xi) \\ &\quad - \frac{1}{2}\frac{\partial}{\partial x}\left\{\phi\int(\delta\xi)^2 g(\delta\xi,x;\delta t,t)d(\delta\xi)\right\} + \cdots \end{aligned}$$

and if we note (3.6), (3.7) and (3.8), we obtain, in the limit as $\delta t \to 0$,

$$P(x,t) = M(x,t)\phi(p,x;t) - \frac{1}{2}\frac{\partial}{\partial x}\{V(x,t)\phi(p,x;t)\}. \tag{3.15}$$

In terms of a generation as the unit of time, the above equation becomes (3.11) as was to be shown.

In equations (3.10) and (3.11), $M_{\delta x}$ and $V_{\delta x}$ are in general functions of both x and t, but in most of our applications to population genetics, they are functions of x only, and independent of the time parameter t. Actually in the present paper, we shall be concerned only with the cases where $M_{\delta x}$ and $V_{\delta x}$ are functions of x but independent of t, namely where the process is time homogeneous. However, except for the special case of being identically zero, they can never take constant values, as might be the case in many diffusion problems in physics.

As stated already, our fundamental equation (3.10) can describe the "movement" of the point x representing the gene frequency of a population on the open interval (0,1) and as noted in (3.1), $\phi(x,t)dx$ with $dx = 1/(2N)$ gives the approximate frequency of the class with gene frequency x for $0 < x < 1$. The equation by itself cannot give the rates of change in the relative frequencies of terminal classes. However, these rates can be obtained by utilizing the established relation (3.11); we use the fact that $-P(0,t)$ and $P(1,t)$ respectively represent the rates at which the probability flows into the classes $x = 0$ and $x = 1$, from the open interval (0,1). In the special but important case in which the change in frequencies of these terminal classes ($x = 0$ and $x = 1$) is solely due to such an inflow of the probability, i.e. when boundaries act as "absorbing barriers", we have

$$\frac{df(0,t)}{dt} = -P(0,t), \tag{3.16}$$

$$\frac{df(1,t)}{dt} = P(1,t), \tag{3.17}$$

where $f(0,t)$ and $f(1,t)$ are respectively the frequencies of classes with $x = 0$ and $x = 1$ at the tth generation.

Now, if $\phi(x,t)$ and its first derivative with respect to x are finite at $x = 0$ and if $V_{\delta x}$ and $M_{\delta x}$ vanish there, then

$$-P(0,t) = \frac{1}{2}\left[\frac{d}{dx}V_{\delta x}\right]_{x=0}\phi(0,t).$$

In particular, if the random fluctuation is due solely to random sampling of gametes, $V_{\delta x} = x(1-x)/(2N)$ and $[dV_{\delta x}/dx]_{x=0} = 1/(2N)$. Therefore

$$\frac{df(0,t)}{dt} = \frac{1}{2}\phi(0,t)\frac{1}{2N}. \tag{3.18}$$

The right-hand side of the above equation is approximately equal to half the relative frequency of the subterminal class with $x = 1/(2N)$. This is because, for a large value of N, $\phi(0,t)/(2N)$ must be very near to $f(1/2N,t)$ unless $|\partial\phi/\partial x|$ is very large at the neighborhood of $x = 0$. It should be noted here that if the effective size (N_e) is different from the actual size (N) of the population, we must put $V_{\delta x} = x(1-x)/(2N_e)$ and therefore the right-hand side of (3.18) must be multiplied by the factor of N/N_e. Similarly, we have

$$\frac{df(1,t)}{dt} = \frac{1}{2}\phi(1,t)\frac{1}{2N}. \tag{3.19}$$

The relation (3.11) is also useful in deriving the probability distribution of gene frequencies in the steady state when the distribution curve reaches constancy in form. The distribution in this state may either be obtained from

$$P(x,t) = 0 \qquad \text{(stable distribution)} \tag{3.20}$$

or from

$$P(x,t) = \text{constant} \qquad \text{(steady flux)} \tag{3.21}$$

depending on the circumstances.

The above arguments may readily be extended to the cases of two or more random variables. For the case of two random variables such as appear in the tri-allelic system having three alleles, A_1, A_2 and A_3 with respective frequencies x_1, x_2 and $x_3(=1-x_1-x_2)$, the corresponding differential equation becomes

$$\begin{aligned}\frac{\partial\phi}{\partial t} &= \frac{1}{2}\frac{\partial^2}{\partial x_1^2}\left\{V_{\delta x_1}\phi\right\} + \frac{1}{2}\frac{\partial^2}{\partial x_2^2}\left\{V_{\delta x_2}\phi\right\} \\ &+ \frac{\partial^2}{\partial x_1\partial x_2}\{W_{\delta x_1,\delta x_2}\phi\} - \frac{\partial}{\partial x_1}\{M_{\delta x_1}\phi\} - \frac{\partial}{\partial x_2}\{M_{\delta x_2}\phi\}\end{aligned} \tag{3.22}$$

where $\phi = \phi(p_1,p_2,x_1,x_2;t)$ gives the probability density that the frequencies of A_1 and A_2 become x_1 and x_2 at the tth generation given that their frequencies

are p_1 and p_2 at $t=0$. In the above equation, $W_{\delta x_1 \delta x_2}$ stands for the covariance between δx_1 and δx_2, where δx_1 and δx_2 stand respectively for the rates of change of x_1 and x_2 per generation. Also their mean and variance $M_{\delta x_i}$ and $V_{\delta x_i}$ $(i=1,2)$ are in general functions of x_1 and x_2, as well as of t. More generally, for the case of n segregating loci each having a pair of alleles, if $x^{(i)}$ is the frequency of an allele at the ith locus $(i=1,2,\cdots,n)$ then the Kolmogorov forward equation becomes

$$\frac{\partial \phi}{\partial t} = \frac{1}{2}\sum_i \frac{\partial^2}{\partial^2 x^{(i)}}\{V_{\delta x^{(i)}}\phi\} + \sum_{i<j}\frac{\partial^2}{\partial x^{(i)}\partial x^{(j)}}\{W_{\delta x^{(i)}\delta x^{(j)}}\phi\} - \sum_i \frac{\partial}{\partial x^{(i)}}\{M_{\delta x^{(i)}}\phi\}. \tag{3.23}$$

where $\phi = \phi(x^{(1)},\cdots,x^{(n)};t)$ is the probability density that gene frequencies are $x^{(1)} \sim x^{(1)}+dx^{(1)},\cdots,x^{(n)}\sim x^{(n)}+dx^{(n)}$ in the tth generation. In the above equation $\delta x^{(i)}$ is the rate of change of $x^{(i)}$ per generation and M, V and W respectively denote the mean, variance and covariance of the $\delta x^{(i)}$'s.

The equations for the probability flux in the multivariate case may be derived as in the case of a single variable. Here, I will merely present the equations for the case of two independent random variables, x_1 and x_2:

$$\begin{cases} P(x_1|x_2;t) = -\frac{1}{2}\frac{\partial}{\partial x_1}\{V_{\delta x_1}\phi\} - \frac{1}{2}\frac{\partial}{\partial x_2}\{W_{\delta x_1 \delta x_2}\phi\} + M_{\delta x_1}\phi \\ Q(x_2|x_1;t) = -\frac{1}{2}\frac{\partial}{\partial x_1}\{W_{\delta x_1 \delta x_2}\phi\} - \frac{1}{2}\frac{\partial}{\partial x_2}\{V_{\delta x_2}\phi\} + M_{\delta x_2}\phi, \end{cases} \tag{3.24}$$

where $P(x_1|x_2;t)$ and $Q(x_2|x_1;t)$ are respectively the fluxes at point (x_1,x_2) along the x_1 and x_2 axes. In terms of these quantities, (3.22) is expressed in the form:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial P(x_1|x_2;t)}{\partial x_1} - \frac{\partial Q(x_2|x_1;t)}{\partial x_2}.$$

It should be noted here that the existence of a stable gene frequency distribution at equilibrium does not necessarily mean that

$$P(x_1|x_2;t) = Q(x_2|x_1;t) = 0. \tag{3.25}$$

For example, in a locus with three alleles, A_1, A_2 and A_3, if genes mutate only in the sequence $A_1 \to A_2 \to A_3 \to A_1 \cdots$, a stable non-trivial distribution may be realized under a cyclic flow of probability.

So far, we have treated gene frequencies after t generations as random variables and initial gene frequencies as fixed. For example, in the expression $\phi(p,x;t)$, x is considered as a random variable and p is assumed fixed. This means that

we have considered the process of change in gene frequency in the forward direction in time.

On the other hand, we may regard x as fixed and consider p as a random variable. Namely, we reverse the time sequence and view the process retrospectively. In order to make our argument simpler, we will assume in what follows, that the process is time homogeneous. That is, if $x(t_1)$ and $x(t_2)$ are respectively frequencies of a gene at times t_1 and t_2 $(t_1 < t_2)$ then the probability distribution of $x(t_2)$ given $x(t_1)$, which in general should be a function of t_1 and t_2 separately, depends only on the difference $t_1 - t_2$. Then, we have

$$\phi(p,x;t+\delta t) = \int g(\delta p, p; \delta t)\phi(p+\delta p, x; t)\, d(\delta p). \tag{3.26}$$

The above equation which is a counterpart of (3.3), contains g as a function of three variables only, i.e. δp, p and δt. This is because the probability that the gene frequency changes from p to $p+\delta p$ during the time interval of length δt is the same for any t (generation) due to the assumption of time homogeneity. Expanding $\phi\ (p+\delta p, x; t)$ on the right-hand side of the above equation in terms of δp and using relations (3.6), (3.7) and (3.8) we obtain

$$\frac{\partial \phi(p,x;t)}{\partial t} = \frac{V(x)}{2}\frac{\partial^2 \phi(p,x;t)}{\partial p^2} + M(x)\frac{\partial \phi(p,x;t)}{\partial p}, \tag{3.27}$$

or in terms of one generation as a unit of time, we have

$$\frac{\partial \phi}{\partial t} = \frac{V_{\delta p}}{2}\frac{\partial^2 \phi}{\partial p^2} + M_{\delta p}\frac{\partial \phi}{\partial p}. \tag{3.28}$$

Note here that the initial gene frequency p is the variable and x is assumed to be constant. Mathematically, equation (3.27) is the Kolmogorov backward equation as applied to the time homogeneous case, and it is the adjoint form of (3.9).

When $x = 1$, ϕ in (3.28) gives the probability that the gene whose initial frequency was p becomes fixed in the population by the tth generation. We will denote this probability by $u(p,t)$, for which we have

$$\frac{\partial u(p,t)}{\partial t} = \frac{V_{\delta p}}{2}\frac{\partial^2 u(p,t)}{\partial p^2} + M_{\delta p}\frac{\partial u(p,t)}{\partial p}. \tag{3.29}$$

The probability of fixation by a given time t will then be obtained by solving the above equation with boundary conditions

$$u(0,t) = 0,\ u(1,t) = 1. \tag{3.30}$$

In the present paper we will be especially interested in the ultimate probability of fixation defined by

$$u(p) = \lim_{t\to\infty} u(p,t). \tag{3.31}$$

For this probability,

$$\frac{\partial u}{\partial t} = 0$$

and $u(p)$ satisfies the ordinary differential equation

$$\frac{V_{\delta p}}{2}\frac{d^2u(p)}{dp^2} + M_{\delta p}\frac{du(p)}{dp} = 0 \tag{3.32}$$

with boundary conditions

$$u(0) = 0,\ u(1) = 1. \tag{3.33}$$

Equation (3.29) may readily be extended to multivariate cases: consider n independent loci each with a pair of alleles, a normal and a mutant allele. We will denote by $p^{(i)}$ the initial frequency of the mutant gene at the ith locus ($i = 1,2,\cdots,n$). Let $u(p^{(1)},p^{(2)},\cdots,p^{(n)};t)$ be the probability that all the n mutant genes become fixed in the population by the tth generation, given that their frequencies are $p^{(1)},p^{(2)},\cdots,p^{(n)}$ at $t = 0$. Then $u(p^{(1)},\cdots,p^{(n)};t)$ satisfies

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sum_{i=1}^{n} V_{\delta p^{(i)}}\frac{\partial^2 u}{\partial^2 p^{(i)}} + \sum_{i>j} W_{\delta p^{(i)}\delta p^{(j)}}\frac{\partial^2 u}{\partial p^{(i)}\partial p^{(j)}} + \sum_{i=1}^{n} M_{\delta p^{(i)}}\frac{\partial u}{\partial p^{(i)}}. \tag{3.34}$$

In what follows, we will apply the method of partial differential equations to solve concrete problems arising in the theory of population genetics.

II. Random Drift in the Narrow Sense

4. Random drift in a small finite population

We will start our discussion from the simplest situation where mutation, migration and selection are absent, but the gene frequency fluctuates from generation to generation because of the random sampling of gametes in a finite population. The process of change in gene frequency in this simplified form has attracted considerable attention among evolutionary geneticists, and various names have been given to it. Fisher (1922) called it the "Hagedoorn effect". Since Wright's work (Wright, 1931), the term *drift* has become quite popular among biologists, and terms such as the Wright drift or the Sewall Wright effect have been coined. However, in the mathematical theory of Brownian motion, the term drift originally connotes directional movement of the particle; therefore, to use this term to denote the random process in our context, the adjective *random* should be attached to it.

Let us consider an isolated population of N breeding diploid individuals.

Let A_1 and A_2 be a pair of alleles with respective frequencies x and $1-x$. We assume that mating is at random and that the mode of reproduction is such that N male and N female gametes are drawn as a random sample from the population to form the next generation. The mean and variance in the change of gene frequency x per generation are: $M_{\delta x}=0$ and $V_{\delta x}=x(1-x)/(2N)$, the latter being the binomial variance corresponding to $2N$ genes. If mating is not random, or the distribution of the number of offspring does not follow a Poisson distribution, the effective number N_e may be substituted for the actual number N (cf. Kimura and Crow, 1963).

Substituting the above expressions for $M_{\delta x}$ and $V_{\delta x}$ into (3.10), we obtain the partial differential equation

$$\frac{\partial \phi}{\partial t} = \frac{1}{4N}\frac{\partial^2}{\partial x^2}\{x(1-x)\phi\}, \quad (0<x<1), \tag{4.1}$$

where $\phi=\phi(p,x;t)$ is the probability density that the gene frequency becomes x in the tth generation, given that it is p at $t=0$, i.e.

$$\phi(p,x;0) = \delta(x-p), \tag{4.2}$$

in which $\delta(\cdot)$ represents the Dirac delta function.

To solve (4.1) we try a solution of the form

$$\phi = TX,$$

where T is a function of t only and X is a function of x only. Substituting this into (4.1) and dividing both sides of the equation by TX, we have

$$\frac{1}{T}\frac{\partial T}{\partial t} = \frac{1}{4NX}\frac{\partial^2}{\partial x^2}\{x(1-x)X\}.$$

By assumption, T is a function of t only and hence the left-hand side of the above equation depends only on t, while X is a function of x only, and hence the right-hand side of the equation depends only on x. It follows then that both sides of the equation must equal a constant which we shall designate by $-\lambda$. Thus the above equation can be separated into two ordinary differential equations

$$\frac{dT}{dt} = -\lambda T \tag{4.3}$$

and

$$x(1-x)\frac{d^2X}{dx^2} + 2(1-2x)\frac{dX}{dx} - (2-4N\lambda)X = 0. \tag{4.4}$$

From the first equation (4.3) we have

$$T \propto e^{-\lambda t}$$

The second equation (4.4) is the hypergeometric equation

$$x(1-x)X'' + [\gamma - (\alpha+\beta+1)x]X' - \alpha\beta X = 0 \tag{4.5}$$

in which $\gamma = 2$, $\alpha + \beta = 3$ and $\alpha\beta = 2 - 4N\lambda$.

Thus we have

$$\alpha = \frac{3 + \sqrt{1 + 16N\lambda}}{2} \quad \text{and} \quad \beta = \frac{3 - \sqrt{1 + 16N\lambda}}{2}.$$

Though we cannot impose arbitrary conditions at the boundaries, we require a solution which is finite at the singular points ($x = 0$ and 1).

Among the two independent solutions of (4.5), only one, i.e. $F(\alpha,\beta,\gamma,x)$ is finite at $x = 0$ in this case. In order to find the condition which makes $F(\alpha,\beta,2,x)$ finite at the other singular point ($x = 1$), we make use of the following relation:

$$F(\alpha,\beta,2,x) = \frac{\Gamma(2)\Gamma(2-\alpha-\beta)}{\Gamma(2-\alpha)\Gamma(2-\beta)} F(\alpha,\beta,-1+\alpha+\beta,1-x) + \frac{\Gamma(2)\Gamma(\alpha+\beta-2)}{\Gamma(\alpha)\Gamma(\beta)}(1-x)^{2-\alpha-\beta} F(2-\alpha,2-\beta,3-\alpha-\beta,1-x) \tag{4.6}$$

If we note that $\alpha + \beta = 3$, we see that in order that $\lim_{x\to 1} F(\alpha,\beta,2,x)$ be finite, $2 - \alpha$ must be a negative integer and β must be 0 or a negative integer. Thus the only possible values of λ are represented by

$$\lambda_i = i(i+1)/(4N),$$

where the i's are positive integers ($i = 1,2,3,\cdots$). Corresponding to this eigenvalue λ_i, we have $\alpha_i = 2 + i$ and $\beta_i = 1 - i$. Thus we can write

$$X = F(2+i,1-i,2,x)$$

except that it may be multiplied by a constant. Here it may be convenient to use the Gegenbauer polynomial defined by

$$T^1_{i-1}(z) = \frac{i(i+1)}{2} F\left(i+2,1-i,2;\frac{1-z}{2}\right) \tag{4.7}$$

so that we can put

$$X = T^1_{i-1}(z)$$

where $z = 1 - 2x$. The properties of the polynomial have been thoroughly studied (see for example, Morse and Feshbach, 1953, pp. 782–783).

The complete solution of (4.1) may then be written in the form

$$\phi(p,x;t) = \sum_{i=1}^{\infty} C_i T^1_{i-1}(z) e^{-i(i+1)t/(4N)}. \tag{4.8}$$

The coefficients C_i can be determined by applying the initial condition that the population starts from the gene frequency p, namely from (4.2), that

$$\delta(x-p) = \sum_{i=1}^{\infty} C_i T_{i-1}^1(z).$$

Multiplying by $(1-z^2)T_{i-1}^1(z)$ on both sides of the above equation, and using the orthogonality property

$$\int_{-1}^{1} (1-z^2)T_m^1(z)T_{i-1}^1(z)dz = \delta_{m,i-1}\frac{2(i+1)i}{(2i+1)},$$

where m in the Kronecker $\delta_{m,i-1}$ represents zero or a positive integer, we obtain

$$2\{1-(1-2p)^2\}T_{i-1}^1(1-2p) = C_i\frac{2(i+1)i}{(2i+1)}$$

or

$$C_i = 4p(1-p)\frac{(2i+1)}{i(i+1)}T_{i-1}^1(1-2p).$$

Therefore, the required solution of (4.1) is

$$\phi(p,x;t) = \sum_{i=1}^{\infty} \frac{(2i+1)(1-r^2)}{i(i+1)} T_{i-1}^1(r)\,T_{i-1}^1(z)\,e^{-i(i+1)t/(4N)}, \tag{4.9}$$

where

$r = 1-2p$ and $T_0^1(r) = 1, T_1^1(r) = 3r, T_2^1(r) = \frac{3}{2}(5r^2-1), T_3^1(r) = \frac{5}{2}(7r^3-3r)$, etc.

In terms of the hypergeometric function $F(\cdot,\cdot,\cdot,\cdot)$, equation (4.9) may be expressed in the form

$$\begin{aligned}\phi(p,x;t) = {} & \sum_{i=1}^{\infty} p(1-p)\,i(i+1)(2i+1)\,F(1-i,i+2,2,p) \\ & \cdot F(1-i,i+2,2,x)e^{-i(i+1)t/(4N)} \\ = {} & 6p(1-p)e^{-t/(2N)} + 30p(1-p)(1-2p)(1-2x)e^{-3t/(2N)} + \cdots .\end{aligned} \tag{4.10}$$

For $t > 0$, the series is uniformly convergent in x and p. This may be easily seen if we note that the exponential term approaches zero rapidly.

Based on this solution, the process of change in the probability distribution of gene frequency when the population starts from $p = 0.5$ and 0.1 is illustrated in Figures 2a and 2b. In these figures, the abscissa represents the gene frequency x and the ordinate the probability density ϕ. In discussing such a distribution, it is often convenient to adopt the "frequency interpretation" of probability, regarding the distribution curve as representing relative frequencies of various gene frequency classes in the infinite collection of populations having the same size and subjected to the same conditions. The area under each curve represents

Figure 2a

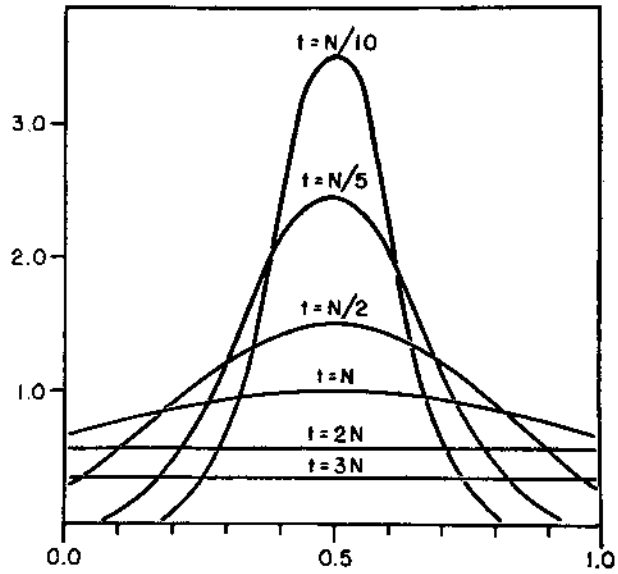
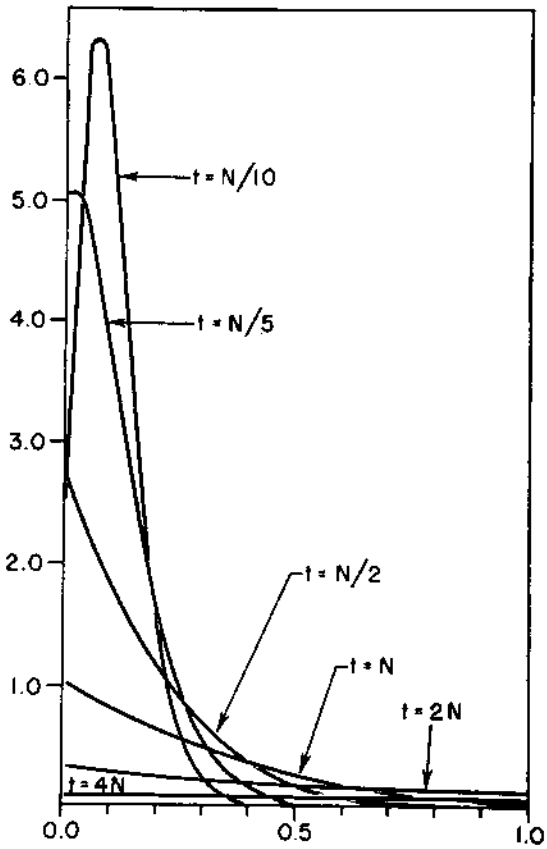


Figure 2b



The process of change in the probability distribution of gene frequency, due to random sampling of gametes in reproduction. It is assumed that the population starts from the gene frequency 0.5 in Figure 2a, and 0.1 in Figure 2b. t indicates time (in generations), and N the effective size of the population. The abcissa is gene frequency, the ordinate is the probability density. (From Kimura, 1955a)

the probability that A_1 and A_2 coexist in the population. It may be seen from the figures that this probability gradually decreases with time. In other words, the frequency of unfixed classes decreases with increasing numbers of generations. This is because if, by random change, the gene frequency becomes either $x = 1$ (fixation of A_1 or loss of A_2 from the population) or $x = 0$ (loss of A_1), it cannot return to intermediate values because of our assumption that no mutation occurs. Namely, genes go irreversibly into fixation (or loss). In the language of probability theory, boundaries at $x = 0$ and $x = 1$ act as absorbing barriers. From (4.10) it can be seen that the probability distribution finally becomes flat and decreases its height at the rate of $1/(2N)$ per generation. This is known as the state of steady decay, and mathematically $1/(2N)$ corresponds to the smallest eigenvalue of the partial differential equation involved. This rate of decay is the most important single quantity used to describe the process of random drift in the narrow sense, and it was first determined by Wright (1931). Thus we have

$$\phi(p,x;t) \sim 6p(1-p)e^{-t/(2N)}, \quad (t \to \infty). \tag{4.11}$$

The number of generations after which this asymptotic formula becomes useful depends on the initial frequency p. For example, with $p = 0.5$, it will be seen from Figure 2a that the distribution curve becomes almost flat after $2N$ generations, and the genes are still unfixed in about 50 per cent of the cases. On the other hand, with $p = 0.1$ (see Figure 2b) it takes $4N$ or $5N$ generations before the distribution curves become practically flat. By that time, however, the genes are fixed in more than 90 per cent of the cases and the asymptotic formula (4.11) may not be as useful as in the case of $p = 0.5$.

The probability that A_1 and A_2 coexist in the population of the tth generation is given by

$$\Omega_t = \int_0^1 \phi(p,x;t)dx = \sum_{j=0}^{\infty} [P_{2j}(r) - P_{2j+2}(r)]e^{-(2j+1)(2j+2)t/(4N)}, \tag{4.12}$$

where $r = 1 - 2p$ and $P(\cdot)$ represents the Legendre polynomials $P_0(r) = 1$, $P_1(r) = r$, $P_2(r) = \frac{1}{2}(3r^2 - 1)$, $P_3(r) = \frac{1}{2}(5r^3 - 3r)$, etc. Thus we have the asymptotic formula

$$\Omega_t \sim 6p(1-p)e^{-t/(2N)}, \quad (t \to \infty). \tag{4.13}$$

The frequency of heterozygotes or the probability that an individual in a population is heterozygous can also be calculated by using (4.10) to obtain

$$H_t = \int_0^1 2x(1-x)\phi(p,x;t)dx = 2p(1-p)e^{-t/(2N)}. \tag{4.14}$$

This shows that heterozygosity decreases exactly at the rate of $1/(2N)$ per genera-

tion. Actually, this holds also for multiallelic cases and is independent of the number of alleles involved.

At this point a few remarks are in order. First, it should be noted that homozygosity or heterozygosity of an individual within a population is a distinct concept from the genetic homogeneity or heterogeneity of the population itself. Wright used the term homallelic or heterallelic; a population is homallelic if it contains only one kind of allele, and is heterallelic if it contains two or more. Secondly, the probability of heterozygosity decreases at the rate of exactly $1/(2N)$ per generation under random mating as shown by (4.14), while as shown in (4.12) the probability of coexistence of both alleles within a population, though continuously diminishing in each generation, does not generally decrease at a constant rate even for a population of constant size N. Its rate of decrease approaches the final value of $1/(2N)$ only asymptotically.

The above treatments do not directly give the probability of absorption, namely the probability of reaching fixation by a given generation t, starting from an intermediate gene frequency p. This may be obtained by the use of the backward equation (3.28) assuming $M_{\delta p}=0$ and $V_{\delta p}=p(1-p)/(2N)$. It turns out to be

$$u(p,t) = p + \sum_{i=1}^{\infty}(2i+1)p(1-p)(-1)^i F(1-i,i+2,2,p)e^{-i(i+1)t/(4N)}\,, \tag{4.15}$$

where this was first obtained from the study of the moments of the distribution (Kimura, 1955 a). In the present case, $u(p,t)$ is equivalent to $f(1,t)$. In terms of Legendre polynomials, (4.15) can be expressed also as

$$f(1,t) = p + \sum_{i=1}^{\infty}\frac{(-1)^i}{2}\{P_{i-1}(r)-P_{i+1}(r)\}e^{-i(i+1)t/(4N)}, \tag{4.16}$$

where $r=1-2p$. Using relation (3.19), it can also be obtained by integrating $\phi(p,1,\tau)/(4N)$ with respect to τ from $\tau=0$ to $\tau=1$. The probability $f(0,t)$ of A_1 being lost or A_2 being fixed by the tth generation is obtained simply by replacing p with $1-p$ and r with $-r$ in the above expressions. It is then possible to show that

$$f(1,t)+\Omega_t+f(0,t)=1\,.$$

So far, we have assumed that the population contains a pair of alleles at the start. If a population contains more than two alleles, the problem becomes much more difficult. For example, suppose that the population contains three alleles A_1, A_2 and A_3 with respective frequencies of x_1, x_2 and $1-x_1-x_2$. Then, the probability density $\phi(p_1,p_2;x_1,x_2,t)$ that the frequencies of A_1 and A_2 become respectively x_1 and x_2 at time t, given that they are p_1 and p_2 at $t=0$, satisfies the following equation (cf. equation 3.22)

$$(4.17)\quad \frac{\partial \phi}{\partial t}=\frac{1}{4N}\frac{\partial^2}{\partial x_1^2}\{x_1(1-x_1)\phi\}-\frac{1}{2N}\frac{\partial^2}{\partial x_1 \partial x_2}\{x_1x_2\phi\}+\frac{1}{4N}\frac{\partial^2}{\partial x_2^2}\{x_2(1-x_2)\phi\},$$

where $0 < x_1 < x_1 + x_2 < 1$.
The solution of this equation with the initial condition

$$\phi(p_1, p_2; x_1, x_2; 0) = \delta(x_1 - p_1)\delta(x_2 - p_2)$$

is (cf. Kimura, 1956 a)

$$\phi(p_1,p_2;x_1,x_2;t) = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} C(n,j)(1-x_3)^n \cdot T_n^1\left(\frac{x_1-x_2}{1-x_3}\right)$$

$$(4.18)\qquad \cdot J_j(2n+5, 2n+4, 1-x_3)\exp\left\{-\frac{(j+n+2)(j+n+3)}{4N}t\right\}$$

$$= 5!\,p_1p_2p_3e^{-3t/(2N)}+\frac{7!}{2!}p_1p_1p_3\left\{\sum_{i=1}^{3}\left(p_i-\frac{1}{3}\right)x_i\right\}e^{-6t/(2N)}+\cdots,$$

where

$$C(n,j) = \frac{4\cdot(j+2n+3)!(j+2n+4)!(2j+2n+5)}{j!(j+1)!(n+1)(n+2)\cdot(2n+2)!(2n+3)!}$$

$$\cdot\, p_1p_2p_3(1-p_3)^n \cdot T_n^1\left(\frac{p_1-p_2}{1-p_3}\right)J_j(2n+5, 2n+4, 1-p_3).$$

Here $x_3 = 1 - x_1 - x_2$, $p_3 = 1 - p_1 - p_2$ and, $T_n^1(\cdot)$ and $J_j(\cdot,\cdot,\cdot)$ denote respectively the Gegenbauer and Jacobi polynomials. The latter is expressed in terms of the hypergeometric function as follows:

$$J_n(a,c,\rho) = F(a+n, -n, c, \rho).$$

In particular $J_0(a, c, \rho) = 1$, $J_1(5, 4, \rho) = 1 - \frac{3}{2}\rho$, etc.

For the general case of an arbitrary number of alleles, the exact solution has not been obtained. Nevertheless, the asymptotic behavior of the processes has been successfully analysed and we have the following result (cf. Kimura, 1955 b).

If we start from a population which contains n alleles, say $A_1, A_2, \cdots, A_n$ with frequencies $p_1, p_2, \cdots, p_n$ respectively ($\sum_1^n p_i = 1$), the probability density that it contains k of them, say $A_1, A_2, \cdots, A_k$ with respective frequencies $x_1, x_2, \cdots, x_k$ ($\sum_1^k x_i = 1$) in the tth generation is given asymptotically by

$$(4.19)\qquad \phi_{1,2,\cdots,k}(x_1, x_2, \cdots, x_{k-1}; t) \sim (2k-1)!\left(\prod_{j=1}^{k} p_j\right)e^{-k(k-1)t/(4N)},$$

where $k \leqq n$. The validity of this formula depends on the assumption that the population size N is sufficiently large as compared with n, the number of alleles in question.

The above result indicates that as the number of coexisting alleles increases,

the rate at which a particular state is eliminated by random drift increases rapidly. In this sense, random drift may be effective in keeping down the number of co-existing alleles in the population.

5. An approximate treatment by the angular transformation

It has been noted by Fisher (1922) that if we transform the gene frequency from x to θ by $\cos\theta = 1 - 2x$, the sampling variance becomes independent of gene frequency. Here θ changes from 0 to π as x changes from 0 to 1. The rate of change of θ per generation, i.e. $\delta\theta$, is related to δx as follows

$$\delta\theta = [x(1-x)]^{-1/2}\delta x - \frac{1}{4}(1-2x)[x(1-x)]^{-3/2}(\delta x)^2 + \cdots . \tag{5.1}$$

Thus, taking $M_{\delta x} = 0$ and $V_{\delta x} = x(1-x)/(2N)$, and neglecting the higher order terms, we get

$$M_{\delta\theta} = -\cot\theta/(4N),\ V_{\delta\theta} = 1/(2N). \tag{5.2}$$

If we start from a fixed gene frequency, the variance in θ after t generations may be given approximately by $V_\theta(t) = t/(2N)$, if t is much smaller than N. It should be noted here, however, that the expected value of $\delta\theta$ is not strictly zero, i.e. the expression $M_{\delta\theta} = 0$ which might be obtained by neglecting the second and following terms on the right-hand side of (5.1) is incorrect. This may not produce any trouble in the treatment of variance for a short period, but will cause a serious error in the gene frequency distribution after a large number of generations. Let $\psi(\theta, t)$ be the probability density of θ at the tth generation. We obtain

$$\frac{\partial\psi}{\partial t} = \frac{1}{4N}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{4N}\frac{\partial}{\partial\theta}(\psi\cos\theta), \tag{5.3}$$

which is the Kolmogorov forward equation in terms of θ for this case (cf. 3.10 and 5.2). The above equation was given by Fisher (1930) as the correct equation to replace the erroneous one which he had given earlier (Fisher, 1922), i.e.

$$\frac{\partial\psi}{\partial t} = \frac{1}{4N}\frac{\partial^2\psi}{\partial\theta^2}. \tag{5.4}$$

The latter equation was obtained by taking $M_{\delta\theta} = 0$ and gave the incorrect value of $1/(4N)$ as the rate of steady decay, rather than the correct value of $1/(2N)$ obtained by Wright (1931). The fact that the sampling variance becomes constant by the angular transformation is nevertheless convenient for treating data on random drift over a relatively short period (cf. Bodmer, 1960).

One of the most interesting applications of this type of transformation was given by Cavalli and Conterio (1960), who analysed the distribution of blood group genes in the Parma River Valley. Their method is based on the concept of "distance" as suggested by Fisher. Consider a locus with multiple alleles

$A_1, A_2, \cdots, A_n$. In order to characterize the genetic constitution of a population, we use n-dimensional Cartesian coordinates with each axis representing the square root of one of the allelic frequencies. A population which contains these alleles with respective frequencies of $x_1, x_2, \cdots, x_n$ may be located on a hypersphere with radius 1. Figure 3 illustrates the case with three alleles. Let $p_1, p_2, \cdots, p_n$ be the corresponding allelic frequencies in some other population. Then the coefficient of genetic distance θ between these two populations may be defined by

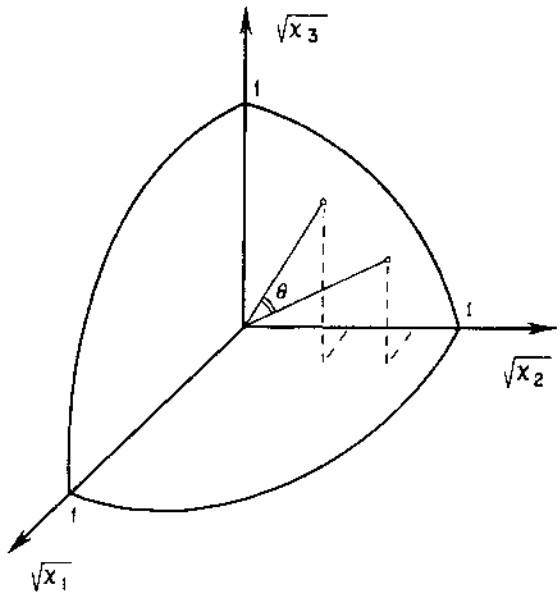


Figure 3
This illustrates the concept of the genetic distance for $n = 3$. (Redrawn from Cavalli and Conterio, 1960, with a slight modification)

$$\cos\theta = \sum_{i=1}^{n} \sqrt{x_i}\sqrt{p_i}. \tag{5.5}$$

Geometrically, θ is the angle made by two vectors $(\sqrt{x_1}, \sqrt{x_2}, \cdots, \sqrt{x_n})$ and $(\sqrt{p_1}, \sqrt{p_2}, \cdots, \sqrt{p_n})$ both radiating from the origin (see Figure 3 where $n=3$). If we take $x_1, x_2, \cdots, x_n$ as the allelic frequencies of a local population which is undergoing random drift, and take $p_1, p_2, \cdots, p_n$ as a set of reference points, such as the initial frequencies or average frequencies over infinitely many local populations, then it can be shown that the variance in the change of θ per generation is given approximately by

$$V_{\delta\theta} = \frac{1}{8N} = \frac{1}{4\,(\text{Number of genes})}, \tag{5.6}$$

where N is the effective size of the population. As in the angular transformation $\theta = \cos^{-1}(1-2x)$, the expected value of $\delta\theta$ in this case is not zero, but this may not cause any serious error in treating the variance of θ, as long as the number of generations involved is much smaller than N. Furthermore, if θ_a is the genetic distance of a population from the general population with respect to the first locus consisting of alleles $A_1, A_2, \cdots, A_n$ and if θ_b is the corresponding distance with respect to the second locus consisting of alleles $B_1, B_2, \cdots, B_m$, then the distance θ_{ab} with respect to these two loci combined is given by $\cos\theta_{ab} = \cos\theta_a \cos\theta_b$.

Using these relations, Cavalli and Conterio studied the regression of θ on population density, village size and "dimensionality". For details, their original paper (Cavalli and Conterio, 1960) should be referred to.

III. Random Drift in the Wide Sense

6. A population under linear pressure and random sampling of gametes

Under the term linear pressure, we include the pressures of gene mutations and of migration. Usually the rate of mutation is so low that although supplying the raw material for evolution, it can hardly determine the course of change in gene frequency. On the other hand, migration between sub-populations may be of considerable significance in determining the gene frequency, as will be found in Wright's theory.

Consider a random mating population of effective size N in which the frequencies of a pair of alleles A_1 and A_2 are x and $1-x$ respectively. Let us suppose that this population exchanges individuals with a random sample taken from the total species at the rate of m per generation. Then the mean and variance of the rate of change in x are given by

$$M_{\delta x} = m(\bar{x} - x), \; V_{\delta x} = x(1-x)/(2N), \tag{6.1}$$

where $\bar{x}$ is the frequency of A_1 in the immigrants. If mutation rates are not negligible, m may be replaced by $m + \mu + \nu$ and $m\bar{x}$ by $m\bar{x} + \nu$, where μ and ν are respectively the mutation rates of A_1 to and from its allele A_2. Though the pressure of selection is intrinsically non-linear, in certain cases, like that of selection acting at the neighborhood of the equilibrium gene frequency, it may be treated as if it were linear with good approximation. However, the range of applicability is quite restricted. The solution of (3.10) when $M_{\delta x}$ and $V_{\delta x}$ are given by (6.1) was obtained by the author through the study of the moments of the distribution (Crow and Kimura, 1956), and it was found that it agrees with the "fundamental solution with flux zero boundary condition" derived by Goldberg (1950). The solution is given by

$$\phi(p,x;t) = \sum_{i=0}^{\infty} X_i(x)\exp\left\{-i\left(m+\frac{i-1}{4N}\right)t\right\} \tag{6.2}$$

where

$$X_i(x) = x^{B-1}(1-x)^{(A-B)-1}F(A+i-1,-i,A-B,1-x)$$
$$\cdot F(A+i-1,-i,A-B,1-p)\,\frac{\Gamma(A-B+i)\Gamma(A+2i)\Gamma(A+i-1)}{i!\Gamma^2(A-B)\Gamma(B+i)\Gamma(A+2i-1)}$$

in which $A=4Nm$ and $B=4Nm\bar{x}$.

Figure 4a

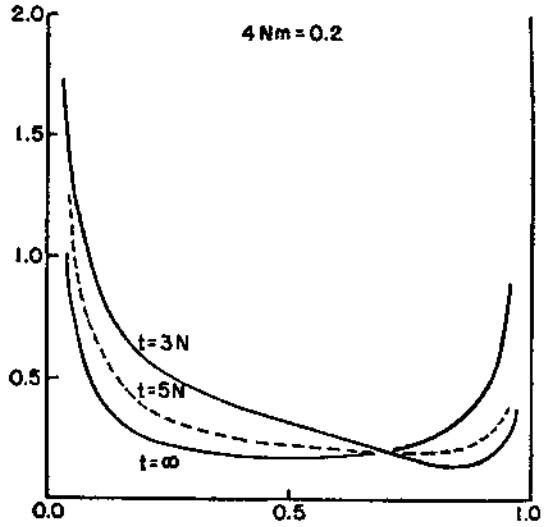
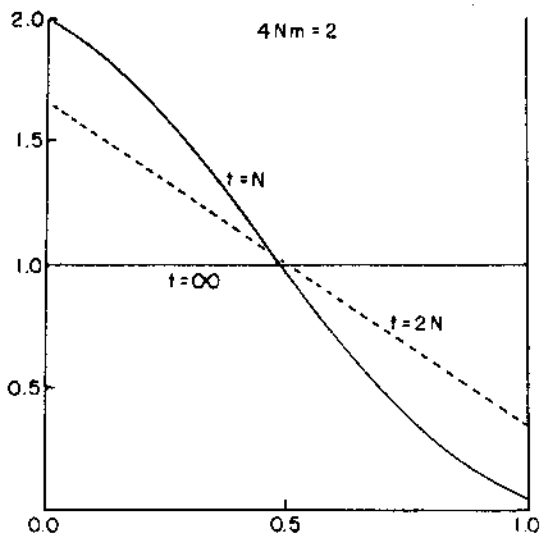


Figure 4b



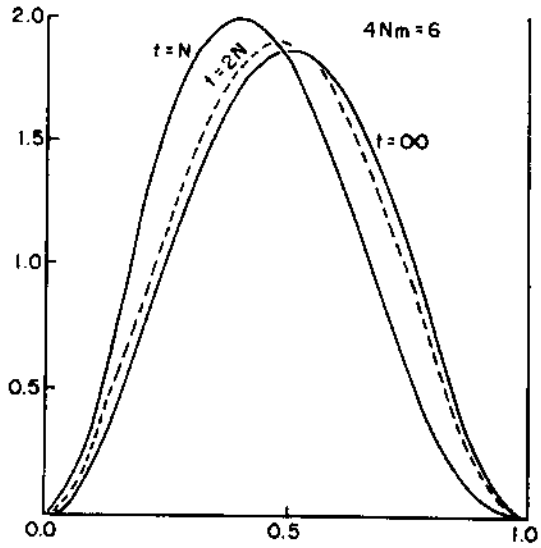


Figure 4c

Asymptotic behavior of the distribution curve for a finite population with migration or other linear pressure. In all three drawings, the gene frequency of the immigrants is assumed to be 0.5 and the initial frequency in the population 0.2. The abscissa is the gene frequency x, the ordinate is the probability density ϕ. N represent population number, and m the rate of migration. (From Crow and Kimura, 1956)

Figures 4a, 4b and 4c show the asymptotic behavior of the distribution curve for three different cases: $4Nm = 0.2$, $4Nm = 2$ and $4Nm = 6$. In all three cases illustrated, the gene frequency $\tilde{x}$ of the immigrants is 0.5 and the initial gene frequency p of the population is assumed to be 0.2.

As $t \to \infty$, our formula (6.2) converges to Wright's well known formula for the steady state gene frequency distribution with migration

$$\phi(x) = \frac{\Gamma(4Nm)}{\Gamma(4Nm\tilde{x})\,\Gamma(4Nm(1-\tilde{x}))}\; x^{4Nm\tilde{x}}(1-x)^{4Nm(1-\tilde{x})-1} .$$

7. Change of gene frequency under selection and random sampling of gametes

7.1. *Genic selection (Case of no dominance).* Since selection, either natural or artificial, is always at work in the process of evolution or breeding, it is extremely important to study the effect of selection under random sampling of gametes. We will start from the simplest case of genic selection and consider random mating population of size N, in which A_1 and A_2 occur with respective frequencies x and $1-x$. Let s be the selective advantage of A_1 over A_2 such that the average rate of change in x per generation is $M_{\delta x} = sx(1-x)$. We take, as before $V_{\delta x} = x(1-x)/(2N)$, which is the variance due to random sampling of gametes. With these expressions, (3.10) becomes

$$\frac{\partial \phi}{\partial t} = \frac{1}{4N}\frac{\partial^2}{\partial x^2}\{x(1-x)\phi\} - s\frac{\partial}{\partial x}\{x(1-x)\phi\}. \tag{7.1}$$

The process of change in gene frequency is analogous to the one studied in Section 4, but here selection is superimposed. The boundaries $x = 0$ and $x = 1$ act as absorbing barriers and the probability that a population contains both alleles A_1 and A_2 gradually decreases with time. Finally, it decreases at a constant rate which is given by the smallest eigenvalue (λ_0) of the above equation. In this state of steady decay, the distribution curve retains constancy of form, but its height decreases at the rate of λ_0 per generation. Probably, the smallest eigenvalue λ_0 is the most important single quantity in the representation of this stochastic process. The above equation has been used to analyse the gene frequency change in a very small experimental population of *Drosophila melanogaster* (Wright and Kerr, 1954). In this paper, Wright devised an ingenious method for analysing the process of steady decay; the complete solution of the above equation has been obtained by the present author (Kimura, 1955c and Crow and Kimura, 1956). In order to solve the equation (7.1), let us put

$$\phi \propto e^{2cx}V(x)e^{-\lambda t}$$

where $c = Ns$ and $V(x)$ is a function of x only. If we substitute this in (7.1), we have

$$x(1-x)\frac{d^2V}{dx^2} + 2(1-2x)\frac{dV}{dx} - \{2 + 4c^2x(1-x) - 4N\lambda\}V = 0. \tag{7.2}$$

Then by the substitution

$$x = (1-z)/2,$$

the above equation (7.2) becomes

$$(1-z^2)\frac{d^2V}{dz^2} - 4z\frac{dV}{dz} + \{(4N\lambda - 2 - c^2) + c^2z^2\}V = 0, \tag{7.3}$$

where $z = 1-2x$ $(-1 < z < 1)$. This type of differential equation is known as the oblate spheroidal equation. We want here the solutions which are finite at the singularities, $z = \pm 1$, and reduce to the Gegenbauer polynomials if there is no selection ($Ns = c = 0$). Such a solution has been studied by Stratton and others (1941) and is expressed in the form

$$V_{1k}^{(1)}(z) = \sum_{n=0,1}{}' f_n^k T_n^1(z), \tag{7.4}$$

where $k = 0, 1, 2, \cdots$ (k here corresponds to l in the notations of Stratton *et al.*). In the above expression, f_n^k's are constants and $T_n^1(z)$'s are the Gegenbauer polynomials defined by (4.7). The primed summation is over even values of n if k is even, odd values of n if k is odd.

The desired solution of (7.2) is given by summing the $V_{1k}^{(1)}(z)$ for all possible values of k, after having multiplied through by $e^{2cx-\lambda_k t}$, where λ_k is the kth eigenvalue; then

$$\phi(p,x;t) = \sum_{k=0}^{\infty} C_k e^{-\lambda_k t + 2cx} V_{1k}^{(1)}(z). \tag{7.5}$$

The coefficient C_k can be determined by the initial condition

$$\phi(p,x;0) = \delta(x-p),$$

using the orthogonal relation

$$\int_{-1}^{1} (1-z^2) V_{1k}^{(1)}(z) V_{1l}^{(1)}(z) dz = \delta_{kl} \sum_{n=0,1}{}' (f_n^k)^2 \frac{(n+2)!}{n!(2n+3)}.$$

Thus we have

$$C_k = \frac{(1-r^2)\, e^{-c(1-r)} V_{1k}^{(1)}(r)}{\sum_{n=0,1}{}' \dfrac{(n+1)(n+2)}{(2n+3)} (f_n^k)^2}, \tag{7.6}$$

where $r = 1-2p$, $c = Ns$ and the primed summation is over even values of n if k is even, over odd values of n if k is odd. The solution (7.5) with coefficients defined by (7.6) gives the probability distribution of the gene frequency among unfixed classes.

As t increases, the exponential terms in (7.5) decrease in absolute value very rapidly, and for large t only the first few terms are important. The numerical values of the first few eigenvalues λ_0, λ_1 and λ_2 can be obtained from the tables of the separation constants $(B_{1,k})$ in the book by Stratton *et al.* (1941), by using the relation

$$4N\lambda_k = c^2 - B_{1,k}.$$

Among them, the smallest eigenvalue λ_0 gives the final rate of decay and has special importance. For small values of c, λ_0 can be expanded into a power series in c,

$$2N\lambda_0 = 1 + \frac{2}{5}c^2 - \frac{2}{5^3 \cdot 7}c^4 - \frac{2^2}{3 \cdot 5^5 \cdot 7}c^6 - \frac{2 \cdot 31}{5^6 \cdot 7^3 \cdot 11}c^8 - \cdots. \tag{7.7}$$

In the new table of spheroidal wave functions by Stratton *et al.* (1956), "t" is tabulated for c (denoted by g in the table) up to $c = 8.0$ (pp. 506–508), from which values of $2N\lambda_0$ may be obtained by the relation

$$2N\lambda_0 = 1 + \left(\frac{2}{5} - \frac{t}{2}\right)c^2.$$

In Figure 5, the relation between $2N\lambda_0$ and Ns is plotted for Ns from 0 to 8.0.

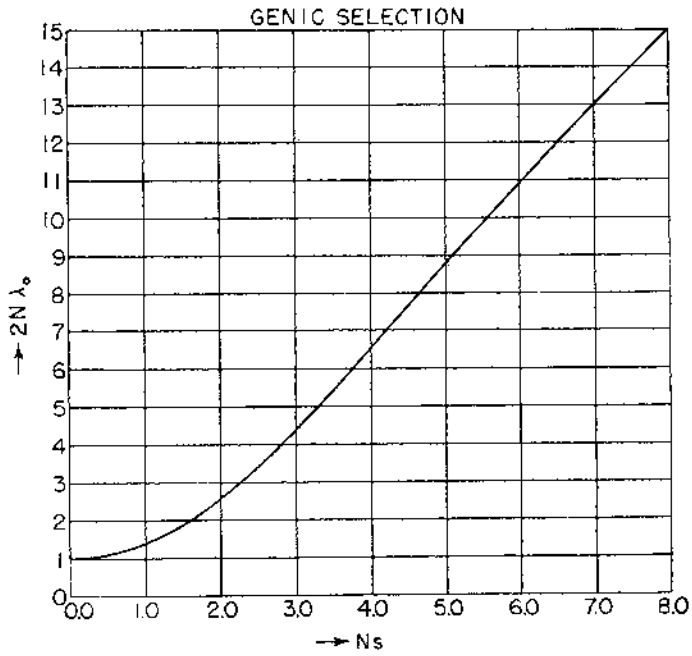


Figure 5

Relation between the rate of steady decay and intensity of selection as illustrated in terms of the relation between $2N\lambda_0$ and Ns, where N is the effective size of the population, λ_0 is the rate of steady decay and s is the selection coefficient relating to genic selection between a pair of alleles. (From Kimura, 1955c)

The eigenfunctions $V_{1k}^{(1)}(z)$ corresponding to the λ_k's are given by (7.4). The coefficients f_n^k corresponding to the first three eigenvalues are found in the table of Stratton *et al.* For $c = 0$, all the formulae given above reduce to the ones for the case of pure random drift studied in Section 4.

The first eigenfunction $V_{10}^{(1)}(z)$ which corresponds to λ_0 is of special significance, since it gives the frequency distribution of unfixed classes in the state of steady decay, when it is multiplied by $e^{c(1-z)}$. It is expressed by

$$V_{10}^{(1)}(z) = f_0^0 T_0^1(z) + f_2^0 T_2^1(z) + f_4^0 T_4^1(z) + \cdots . \tag{7.8}$$

Figure 6 illustrates the shape of the distribution curve in the state of steady decay for three different cases; $Ns = 0$, $Ns = 1.0$ and $Ns = 1.7$. The area under each curve is adjusted so that it is unity.

The frequencies of both terminal classes can be obtained by using the relations (3.19) and (3.18):

$$f(1;t) = e^{2c} \sum_{k=0}^{\infty} \frac{C_k}{4N\lambda_k}(1 - e^{-\lambda_k t} V)_{1k}^{(1)}(-1) \tag{7.9}$$

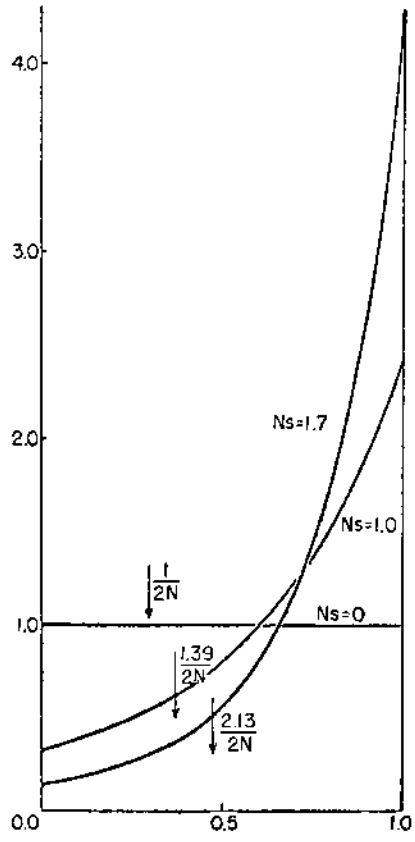


Figure 6

Frequency distribution of unfixed classes at the stage of steady decay is illustrated for three values of Ns. The area under each curve is adjusted so that it is unity. Numbers beside the arrows indicate rates of steady decay. N is the effective size of population, and s is the selection coefficient. (From Kimura, 1955c)

and

$$f(0;t) = \sum_{k=0}^{\infty} \frac{C_k}{4N\lambda_k}(1 - e^{-\lambda_k t}) V_{1k}^{(1)}(1) \tag{7.10}$$

where

$$V_{1k}^{(1)}(-1) = \frac{1}{2} \sum_{n=0,1}{}' (-1)^n (n+1)(n+2) f_n^k$$

and

$$V_{1k}^{(1)}(1) = \frac{1}{2} \sum_{n=0,1}{}' (n+1)(n+2) f_n^k .$$

So far we have considered genic selection in which the gene is additive (i.e. no dominance) with respect to fitness. The case of complete dominance is more

difficult to treat, but the process of steady decay has been worked out for weak selection (Kimura, 1957). Also, for the more general case of an arbitrary degree of dominance, the smallest eigenvalue can be given as a power series for weak selection. Namely, if the selective advantages of A_1A_1 and A_1A_2 over A_2A_2 are s and sh respectively, such that $M_{\delta x} = s[h + (1 - 2h)x]x(1 - x)$, then

$$2N\lambda_0 = 1 + K_1c + K_2c^2 + K_3c^3 + \cdots, \tag{7.11}$$

where

$$K_1 = -\frac{1}{5}D, \quad K_2 = \frac{1}{2\cdot 5} + \frac{2^2\cdot 3}{5^3\cdot 7}D^2, \quad K_3 = \frac{1}{2\cdot 5^3\cdot 7}D - \frac{2^2}{5^5\cdot 7}D^3, \text{ etc.},$$

in which $c = Ns$ and $D = 2h - 1$. It may be noted that for the case of no dominance ($D = 0$), the above series (7.11) agrees with (7.7) provided that $2s$ is used instead of s to express the selective advantage of the homozygous mutants.

7.2. *Case of overdominance.* It has been known since the early work of Fisher (1922) that, in an infinite population, heterozygote superiority in fitness for a pair of alleles leads to a stable polymorphism. Furthermore, a considerable number of claims have been made in recent years stating in effect that overdominance is the major factor for maintaining genetic variability in natural populations. Therefore, investigation of the overdominant case assuming a finite population number will be of interest. Let us assume a pair of overdominant alleles A_1 and A_2 and designate by s_1 and s_2 (both positive) the selection coefficients against the homozygotes A_1A_1 and A_2A_2 respectively, such that the average rate of change in the frequency of A_1 is $M_{\delta x} = [s_2 - (s_1 + s_2)x]x(1 - x)$. The variance of δx is again given by $V_{\delta x} = x(1 - x)/(2N)$.

The partial differential equation corresponding to (3.10) with the $M_{\delta x}$ and $V_{\delta x}$ given above is

$$\frac{\partial\phi}{\partial t} = \frac{1}{4N}\frac{\partial^2}{\partial x^2}\{x(1 - x)\phi\} - \frac{\partial}{\partial x}\{(s_2 - (s_1 + s_2)x)x(1 - x)\phi\}, \tag{7.12}$$

where x is the frequency of A_1.

Let $\bar{s}$ be the average of the two selection coefficients, i.e., $\bar{s} = (s_1 + s_2)/2$, and let $\hat{x}$ be the equilibrium frequency of A_1 that may be expected in an infinitely large population, i.e. $\hat{x} = s_2/(s_1 + s_2)$, then the above equation may be expressed in the form,

$$4N\frac{\partial\phi}{\partial t} = \frac{\partial^2}{\partial x^2}\{x(1 - x)\phi\} - 4N\bar{s}\frac{\partial}{\partial x}\{2(\hat{x} - x)x(1 - x)\phi\}, \tag{7.13}$$

or denoting $2\hat{x} - 1 = \hat{z}$,

$$4N\frac{\partial\phi}{\partial t} = \frac{\partial^2}{\partial x^2}\{x(1 - x)\phi\} + 4N\bar{s}\frac{\partial}{\partial x}\{(2x - 1 - \hat{z})x(1 - x)\phi\}. \tag{7.14}$$

The smallest eigenvalue λ_0 of the above equation has been worked out by Miller (1962).

Without loss in generality, we can take $\hat{x} \geqq 0.5$ or $\hat{z} \geqq 0$. For a large value of $c \equiv N\bar{s} = N(s_1 + s_2)/2$ and for the range $1 > \hat{z} \geqq 0$, Miller obtained the asymptotic expansion

$$2N\lambda_0 = \frac{c}{T(x,\hat{z})}\left[\frac{(1-\hat{z})e^{-c(1-\hat{z})^2}}{S\{c(1-\hat{z})^2\}} + \frac{(1+\hat{z})e^{-c(1+\hat{z})^2}}{S\{c(1+\hat{z})^2\}}\right], \tag{7.15}$$

where

$$S(X) = 1 + \frac{1}{2}\frac{1}{X} + \frac{1\cdot 3}{2^2}\frac{1}{X^2} + \frac{1\cdot 3\cdot 5}{2^3}\frac{1}{X^3} + \cdots,$$

$$T(c,\hat{z}) = \sum_{i=0}^{\infty} \frac{C_{2i}\Gamma(i+\frac{1}{2})}{c^{i+\frac{1}{2}}} = \frac{C_0}{\sqrt{c}}\sqrt{\pi} + \cdots$$

in which the C_i's are given by the recurrence relation

$$(1-\hat{z}^2)C_{i+1} = 2\hat{z}C_i + C_{i-1},$$

starting from

$$C_0 = 1/(1-\hat{z}^2) \text{ and } C_1 = 2\hat{z}/(1-\hat{z}^2)^2.$$

In particular, when $\hat{z} = 0$, that is when $s_1 = s_2$, (7.15) reduces to

$$2N\lambda_0 = \frac{2c^{3/2}e^{-c}}{\sqrt{\pi}\{S(c)\}^2}$$

Also, when $2N(s_1 + s_2)(1-\hat{x})^2$ is large, (7.15) may be replaced by

$$2N\lambda_0 = \sqrt{\frac{N^3(s_1+s_2)^3}{2\pi}}\cdot 4\hat{x}^2(1-\hat{x})^2\left\{\frac{e^{-4c(1-\hat{x})^2}}{\hat{x}} + \frac{e^{-4c\hat{x}^2}}{(1-\hat{x})}\right\}. \tag{7.16}$$

Miller has also obtained λ_0 for various values of $c(\geqq 0)$ up to $c = 12$ by numerical analysis.

It may be noted here that for small values of $N\bar{s}$, the eigenvalue may be calculated from the power series (7.11) by putting $c = N(s_2 - s_1)$ and $cD = N(s_1 + s_2) = 2N\bar{s}$ in it. Figure 7 is constructed on the basis of his numerical results, giving $2N\lambda_0$ as a function of c for the cases of $\hat{x} = 0.5, 0.7, 0.8$ and 0.9. One of the most remarkable features disclosed in the figure seems to be that if s_1 and s_2 differ to such an extent that the equilibrium frequency $\hat{x}$ is higher than 0.8 (or, because of symmetry, less than 0.2), overdominance tends to *accelerate* fixation as compared with the neutral case, rather than retard fixation. This was first pointed out by Robertson (1962) who presented this fact in the form of a graph shown in Figure 8, where the term retardation factor is defined as the reciprocal of $2N\lambda_0$. According to him, selection for a heterozygote is a factor retarding fixation only if the equilibrium frequency lies within the range of

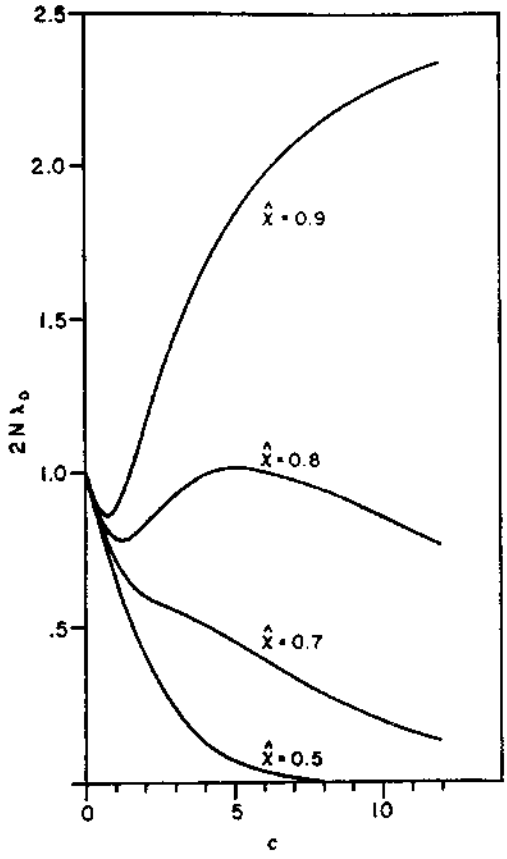


Figure 7
Relation between $2N\lambda_0$ and $c(=N(s_1+s_2)/2)$ is plotted for various equilibrium frequencies $\hat{x}$, when there is overdominance between a pair of alleles.

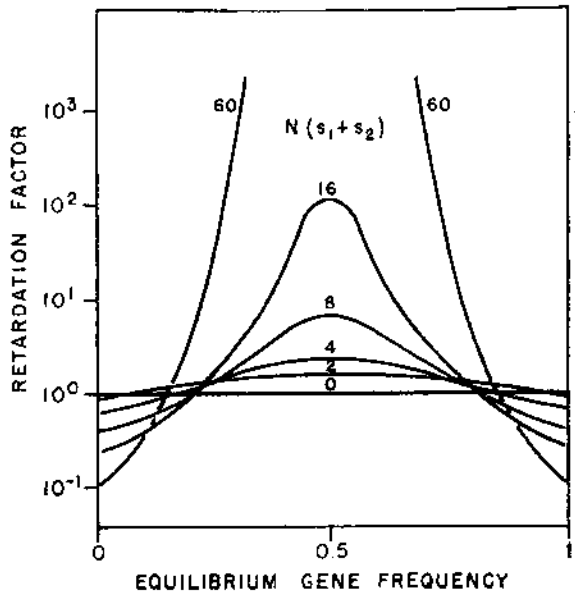


Figure 8
Graphs showing retardation factor as a function of equilibrium gene frequency for various values of $N.(s_1+s_2)$, where N is the effective size of the population, and s_1 and s_2 are the selection coefficients against both homozygotes. (From Robertson, 1962)

0.2–0.8. For equilibrium gene frequencies outside this range, there is a range of values of $N(s_1 + s_2)$ for which heterozygote advantage accelerates fixation, and the more extreme the equilibrium frequency the wider the range. However, for all values of $\hat{x}$ except 0 or 1, an increase in the values of $N(s_1 + s_2)$ eventually leads to retardation.

8. Randon fluctuation of selection intensities

Among the factors which cause random fluctuation in gene frequencies, random fluctuation of selection intensities may often be as important as random sampling of gametes. To single out this factor, we will here assume that the population is infinitely large so that the effect of random sampling may be neglected. Also we will consider the simplest case of genic selection in which a pair of alleles, A_1 and A_2 are involved. Let s be the selective advantage of A_1 over A_2 such that the rate of change in the frequency of A_1 for a fixed value of s is $sx(1-x)$. Let us assume also that A_1 is selectively neutral on the average so that the mean value of s over a long period is zero, and its variance V_s is constant. Then $M_{\delta x} = 0$, $V_{\delta x} = V_s x^2(1-x)^2$ and the partial differential equation (3.10) becomes

$$\frac{\partial \phi}{\partial t} = \frac{V_s}{2} \frac{\partial^2}{\partial x^2}\{x^2(1-x)^2\phi\} \qquad (0 < x < 1). \tag{8.1}$$

To solve this equation, let us put

$$u = \tfrac{1}{2} e^{V_s t/8} x^{3/2}(1-x)^{3/2}\phi$$

and

$$\xi = \log \frac{x}{1-x}.$$

Then we obtain the heat conduction equation

$$\frac{\partial u}{\partial t} = \frac{V_s}{2} \frac{\partial^2 u}{\partial \xi^2} \qquad (-\infty < \xi < \infty).$$

It is known that this equation has a unique solution which is continuous over $-\infty$ to $+\infty$ when $t \geqq 0$, and reduces to $u(\xi,0)$ when $t = 0$. The solution is given by

$$u(\xi,t) = \frac{1}{\sqrt{2\pi V_s t}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(\xi-\eta)^2}{2V_s t}\right\} u(\eta,0)d\eta.$$

Therefore, the required solution of (8.1), when the initial distribution of gene frequency is $\phi(x,0)$, is given by

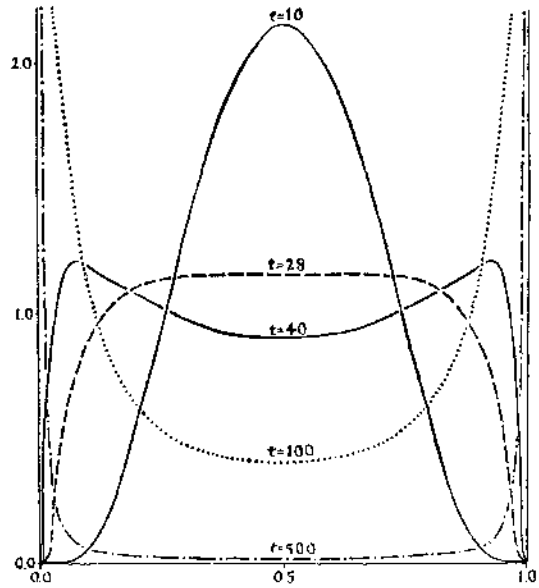


Figure 9

The process of change in the gene frequency distribution under random fluctuation of selection intensities. In this illustration it is assumed that the gene is selectively neutral when averaged over a very long period, that there is no dominance, and $p = 0.5$, $V = 0.0483$. The abscissa is the gene frequency x, the ordinate the probability density ϕ. (From Kimura, 1954)

$$(8.2)\quad \phi(x,t) = \frac{1}{\sqrt{2\pi V_s t}} \cdot \frac{e^{-V_s t/8}}{\{x(1-x)\}^{3/2}} \cdot \int_0^1 \exp\left\{-\frac{\left[\log\frac{x(1-y)}{(1-x)y}\right]^2}{2V_s t}\right\} \cdot \sqrt{y(1-y)}\,\phi(y,0)\,dy.$$

If the initial condition is not a continuous distribution, but a fixed gene frequency p, the above formula reduces to

$$(8.3)\quad \phi(p,x;t) = \frac{1}{\sqrt{2\pi V_s t}} \exp\left\{-\frac{V_s}{8}t - \frac{\left[\log\frac{x(1-p)}{(1-x)p}\right]^2}{2V_s t}\right\} \frac{\sqrt{p(1-p)}}{(\sqrt{x(1-x)})^3}.$$

The process of change in the distribution curve with time is illustrated in Figure 9 assuming $p = 0.5$ and $V_s = 0.0483$. As will be seen in the figure the distribution curve is unimodal for a considerable number of generations (in the case illustrated, 27 generations), after which it becomes bimodal. In the 100th generation, gene frequencies in our example giving maximum probability (peaks in the curve) are approximately 0.0007 and 0.9993. As time goes on the distribution curve becomes more and more U-shaped, though it is not a true U-shaped curve, since its value at either terminal is always 0. This means that as time elapses the gene

frequency shifts towards either terminal of the distribution ($x = 0$ or 1) indefinitely and accumulates in the neighborhood just short of fixation or loss, but never becomes fixed or lost completely (at least theoretically). To distinguish this from the fixation or loss in the case of random drift in small populations, the terms *quasi-fixation* and *quasi-loss* were proposed (Kimura, 1954).

If the genes are not neutral on the average, then $M_{\delta x} = 0$ should be replaced with $M_{\delta x} = \bar{s}x(1-x)$ in the partial differential equation, where $\bar{s}$ is the long term average of s. Unfortunately, the solution of the corresponding partial differential equation has not so far been found.

However, the following approximate treatment may be helpful in obtaining a rough picture of the process involved. If we transform the gene frequency x into its logit ξ by the relation $\xi = \log[x/(1-x)]$, ξ changes continuously from $-\infty$ to $+\infty$ as x changes from 0 to 1. For a small change of ξ, we have

$$\delta\xi = [x(1-x)]^{-1}\delta x + (2x-1)[2x^2(1-x)^2]^{-1}(\delta x)^2 + \cdots. \tag{8.4}$$

Neglecting terms of higher order than the first, and noting $M_{\delta x} = \bar{s}x(1-x)$ and $V_{\delta x} = V_s x^2(1-x)^2$, we obtain

$$M_{\delta\xi} = \bar{s}, \quad V_{\delta\xi} = V_s. \tag{8.5}$$

These expressions indicate that on the logit scale the mean and variance of the gene frequency distribution increase approximately linearly with time. Namely the probability distribution of ξ in the tth generation is given by the normal distribution with mean $\xi_0 + \bar{s}t$ and variance $V_s t$, where ξ_0 is the logit of p, i.e. $\xi_0 = \log[p/(1-p)]$. Figure 10 illustrates the process of change obtained by this method for the case of a pair of alleles with $\bar{s} = 0.1$ and $V_s = 0.0025$ ($\sigma_s = 0.05$).

Actually, the case of genic selection in an infinite population with random fluctuation in selection intensities can most easily be treated by the following discrete model. Consider the multiple alleles, $A_1, A_2, \cdots, A_n$ at a locus and let w_i be the fitness of A_i measured in selective values. If $x_i(t)$ is the frequency of A_i in the tth generation, then

$$x_i(t+1) = w_i x_i(t)/\bar{w}, \tag{8.6}$$

where $\bar{w}$ is the average selective value of the population in the tth generation,

$$\bar{w} = \sum_{i=1}^{n} w_i x_i(t).$$

Here t takes on discrete values, 0, 1, 2, etc.

From (8.6), it follows that, for any pair of alleles, say for A_i and A_j, we have

$$\log[x_i(t)/x_j(t)] = \log(w_i/w_j) + \log[x_i(t-1)/x_j(t-1)],$$

so that

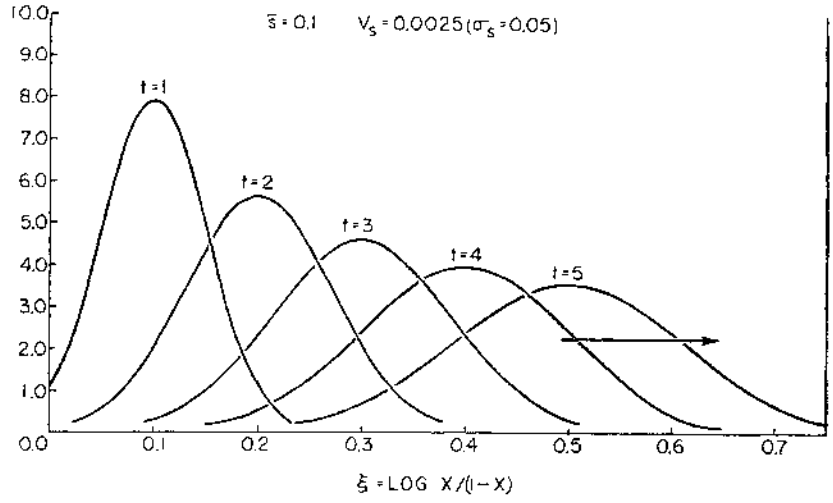


Figure 10

The process of change in gene frequency distribution due to genic selection with random fluctuation of selection intensity. Gene frequency is measured on a logit scale ξ, and $\bar{s}$ and V_s are respectively the mean and variance of the selection intensity. In this figure the initial gene frequency is 0 on the logit scale, i.e. $p = 0.5$, and t stands for time (in generations).
(From Kimura, 1955 c)

$$z_{i/j}(t) = \sum_{\tau=0}^{t-1} s_{i/j}(\tau) + z_{i/j}(0) \qquad (t \geqq 1), \tag{8.7}$$

where $z_{i/j}(t)$ is the logarithmic gene ratio at the tth generation, i.e.

$$z_{i/j}(t) = \log[x_i(t)/x_j(t)],$$

and $s_{i/j}(\tau)$ is the value of $\log(w_i/w_j)$ in the τth generation. Thus, if $s_{i/j}(\tau)$ is distributed normally with mean $\bar{s}_{i/j}$ and variance $\sigma^2_{i/j}$, then $z_{i/j}(t)$ is distributed normally with mean $z_{i/j}(0) + \bar{s}_{i/j}t$ and variance $\sigma^2_{i/j}t$. Furthermore, even if $s_{i/j}$ is not distributed normally, the distribution of $z_{i/j}(t)$ for a large value of t will approach the normal distribution by the central limit theorem. Since (8.7) holds for any pair of i and j, and since

$$\sum_{i=1}^{n} x_i(t) = 1$$

in each generation, the joint distribution of gene frequencies for an arbitrary generation can easily be worked out assuming normality of $z_{i/j}(t)$.

Besides the effect of random fluctuation of selection intensity discussed in this section, that of random fluctuation in migration rate has also been studied. If m is the migration rate which fluctuates randomly from generation to generation with mean $\bar{m}$ and variance V_m, then we have

$$\frac{\partial \phi}{\partial t} = \frac{V_m}{2}\frac{\partial^2}{\partial x^2}\{(x-\xi)^2\phi\} + \bar{m}\frac{\partial}{\partial x}\{(x-\xi)\phi\},$$

where ξ is the frequency of A_1 in the migrants. Assuming that the initial gene frequency p is higher than that of the migrants, the solution of this equation is

$$\phi(p,x;t) = \frac{1}{(x-\xi)\sqrt{\pi V_m t/2}} \exp\left\{-\frac{\left[\log\frac{x-\xi}{p-\xi} + \left(\frac{V_m}{2}+\bar{m}\right)t\right]^2}{2V_m t}\right\} \quad (x \geqq \xi).$$

It is then possible to show that

$$\lim_{t\to\infty} \phi(p,x;t) = \delta(x-\xi),$$

where $\delta(\cdot)$ represents the delta function. For details of this problem, Crow and Kimura (1956) should be consulted.

IV. Genetic Equilibrium; Stationary Distributions and Gene Fixation

9. Gene frequency distribution at equilibrium

9.1. *Stationary distribution.* Random fluctuation of gene frequencies due to the factors so far discussed leads to fixation or loss of genes in a population. On the other hand, reversible mutation or migration tends to restore the intermediate gene frequencies. These two groups of factors will cause a non-trivial stable distribution of gene frequencies at equilibrium.

In the previous sections we have been mainly concerned with $\phi(p,x;t)$, the transition probability density for the gene frequency to become x at the tth generation given that it is p at $t=0$. Now, if we designate by $\phi(x)$ the probability density of such an equilibrium distribution, the existence of a stable distribution $\phi(x)$ means that

$$\lim_{t\to\infty} \phi(p,x;t) = \phi(x),$$

which is independent of the initial frequency p. Detailed knowledge of such a distribution is essential to understand the genetic structure of natural populations. For a single locus with a pair of alleles A_1 and A_2, with respective frequencies x and $1-x$, a quite general formula has been obtained by Wright (1938a, 1949). His formula may be expressed as follows:

$$\phi(x) = \frac{C}{V_{\delta x}} \exp\left\{2\int \frac{M_{\delta x}}{V_{\delta x}} dx\right\}, \tag{9.1}$$

where $M_{\delta x}$ and $V_{\delta x}$ are the mean and the variance of δx. Here $\phi(x)$ stands for the probability density that the gene frequency is x in the population. We can interpret $\phi(x)dx$ as the relative frequency or proportion of populations with gene frequency within the range $x \sim x+dx$ among a hypothetical aggregate consisting of the infinite number of populations satisfying the same conditions.

In the above formula C is a constant which is usually adjusted in such a way that

$$\int_0^1 \phi(x)dx = 1. \tag{9.2}$$

It is remarkable that Wright (1938a) first derived formula (9.1) from the simple consideration that at equilibrium the mean and variance of the frequency distribution are unchanged in successive generations.

From our standpoint it is more natural to derive the formula from (3.11) by imposing condition (3.20), namely that the probability flux must be zero at every point in the open interval $(0, 1)$.

This leads to

$$\frac{1}{2}\frac{d}{dx}(V_{\delta x}\phi) - M_{\delta x}\phi = 0,$$

which, upon integration, gives (9.1). One of the tacit assumptions involved here is that the process is time homogeneous, so that $M_{\delta x}$ and $V_{\delta x}$ are independent of t.

As an example, let us consider the case discussed in Section 6, where

$$M_{\delta x} = m(\bar{x} - x)$$

and

$$V_{\delta x} = x(1-x)/(2N).$$

Substituting these in (9.1), we obtain

$$\phi(x) = Cx^{4Nm\bar{x}-1}(1-x)^{4Nm(1-\bar{x})-1}.$$

The coefficient C as determined from (9.2) is

$$C = \frac{1}{B(4Nm\bar{x}, 4Nm(1-\bar{x}))} = \frac{\Gamma(4Nm)}{\Gamma(4Nm\bar{x})\Gamma(4Nm(1-\bar{x}))}.$$

Thus, the equilibrium distribution is given by

$$\phi(x) = \frac{\Gamma(4Nm)}{\Gamma(4Nm\bar{x})\Gamma(4Nm(1-\bar{x}))} x^{4Nm\bar{x}-1}(1-x)^{4Nm(1-\bar{x})-1}. \tag{9.3}$$

It is important to note that the above equation agrees with (6.3), which is obtained from (6.2) by taking the limit as $t \to \infty$.

For the application of (9.1) to various other cases, the reader may refer to a series of papers by Wright, especially Wright (1938a, b, 1939, 1948), and also to Kimura (1955c).

More recently, an interesting model for a haploid organism with overlapping generations was used by Moran (1958c) to derive an exact distribution under reversible mutation.

So far, we have considered gene frequency distributions containing only one random variable x. Here, Wright's formula (9.1) is fundamental, and it has sufficient generality to be useful for most purposes. However, in order to study

the effect of gene interaction on the frequency distribution involving multiple alleles or multiple loci, formulae for the distribution of more than one random variable are required. Unfortunately, no formulae of comparable generality to (9.1) have been obtained for such cases, but Wright (1937) has obtained an important formula which deals with epistasis under the assumption of random mating and constant fitness of individual genotypes. For the case of two loci each with a pair of alleles, A_1 and A_2 at the first locus and B_1 and B_2 at the second, his formula can be expressed as follows:

$$\phi(x,y) = C\bar{w}^{2N}x^{4Nv_1-1}(1-x)^{4N\mu_1-1}\cdot y^{4Nv_2-1}(1-y)^{4N\mu_2-1}, \tag{9.4}$$

where x and y are respectively the frequencies of A_1 and B_1, and $\bar{w}$ is the average fitness of a population measured in selective values with respect to these two loci. In the above formula, μ_1 is the mutation rate from A_1 to A_2 and v_1 is the rate in the reverse direction, while μ_2 and v_2 are corresponding values for B_1 and B_2. If the fitness is measured in Malthusian parameters (cf. Fisher, 1958, Kimura, 1958), $\bar{w}$ in the above formula should be replaced by $e^{\bar{a}}$, where $\bar{a}$ is the average fitness measured in Malthusian parameters.

Wright (1937) derived the above equation through an ingenious but intuitive argument. From our standpoint, however, it is more natural, as in the case of a single variable, to derive the equation from a consideration of the probability flux. Actually, this enables us to view the problem in a much wider perspective.

Consider a region in two-dimensional Cartesian co-ordinates with $0 \leqq x \leqq 1$ and $0 \leqq y \leqq 1$. The probability flux which passes through the point (x, y) along the x-axis is

$$P(x \mid y;t) = M_{\delta x}\phi(x,y) - \frac{1}{2}\frac{\partial}{\partial x}\{V_{\delta x}\phi(x,y)\} - \frac{1}{2}\frac{\partial}{\partial y}\{W_{\delta x\delta y}\phi(x,y)\}.$$

Similarly, the flux which passes through the same point along the y axis is

$$Q(y \mid x;t) = M_{\delta y}\phi(x,y) - \frac{1}{2}\frac{\partial}{\partial y}\{V_{\delta y}\phi(x,y)\} - \frac{1}{2}\frac{\partial}{\partial x}\{W_{\delta x\delta y}\phi(x,y)\}.$$

At equilibrium, when the flux is zero at every point,

$$P(x \mid y;t) = Q(y \mid x;t) = 0,$$

and we have

$$\begin{bmatrix} V_{\delta x} & W_{\delta x\delta y} \\ W_{\delta x\delta y} & V_{\delta y} \end{bmatrix} \begin{bmatrix} \dfrac{\partial\psi}{\partial x} \\ \dfrac{\partial\psi}{\partial y} \end{bmatrix} = \begin{bmatrix} 2M_{\delta x} - \dfrac{\partial V_{\delta x}}{\partial x} - \dfrac{\partial W_{\delta x\delta y}}{\partial y} \\ 2M_{\delta y} - \dfrac{\partial W_{\delta x\delta y}}{\partial x} - \dfrac{\partial V_{\delta y}}{\partial y} \end{bmatrix} \tag{9.5}$$

where $\psi = \log\phi(x,y)$. Thus, if the above simultaneous equation in $\partial\psi/\partial x$ and $\partial\psi/\partial y$ has a unique solution (ψ_x, ψ_y) and if

$$\psi_x dx + \psi_y dy$$

is the total differential, which we will denote by $d\psi$, then the simultaneous distribution at equilibrium may be given by

$$\phi(x, y) = Ce^{\psi} \tag{9.6}$$

where the constant C is determined by the condition

$$\int_0^1 \int_0^1 \phi(x, y)\, dx dy = 1. \tag{9.7}$$

In the special case in which random sampling of gametes is the sole factor for producing random fluctuation in gene frequencies,

$$V_{\delta x} = \frac{x(1-x)}{2N}, \quad V_{\delta y} = \frac{y(1-y)}{2N} \text{ and } W_{\delta x \delta y} = 0. \tag{9.8}$$

Under the assumption of random mating and constant (but not necessarily equal) fitness of individual genotypes, the mean rates of change in gene frequencies are expressed by

$$\begin{cases} M_{\delta x} = \dfrac{1}{2} x(1-x) \dfrac{\partial \bar{a}}{\partial x} - \mu_1 x + \nu_1 (1-x) \\ M_{\delta y} = \dfrac{1}{2} y(1-y) \dfrac{\partial \bar{a}}{\partial y} - \mu_2 y + \nu_2 (1-y) \end{cases} \tag{9.9}$$

where $\bar{a} = \log \bar{w}$. Substituting (9.8) and (9.9) in (9.5), we get

$$\begin{cases} \psi_x = 2N \dfrac{\partial \bar{a}}{\partial x} - 4N \left(\dfrac{\mu_1}{1-x} - \dfrac{\nu_1}{x} \right) - \dfrac{d}{dx} \log [x(1-x)], \\ \psi_y = 2N \dfrac{\partial \bar{a}}{\partial y} - 4N \left(\dfrac{\mu_2}{1-y} - \dfrac{\nu_2}{y} \right) - \dfrac{d}{dy} \log [y(1-y)], \end{cases}$$

and since $\partial \psi_x / \partial y = \partial \psi_y / \partial x = 2N \partial^2 \bar{a} / \partial x \partial y$, it is evident that $d\psi$ exists. Indeed it is given by

$$\begin{aligned} d\psi = d\{2N\bar{a} &+ (4N\mu_1 - 1)\log(1-x) + (4N\nu_1 - 1)\log x \\ &+ (4N\mu_2 - 1)\log(1-y) + (4N\nu_2 - 1)\log y\}. \end{aligned}$$

Thus, from (9.6), we obtain

$$\phi(x, y) = Ce^{2N\bar{a}} x^{4N\nu_1 - 1} (1-x)^{4N\mu_1 - 1} y^{4N\nu_2 - 1} (1-y)^{4N\mu_2 - 1}, \tag{9.10}$$

where $\bar{a} = \log \bar{w}$. This completes the derivation of (9.4).

9.2. *Distribution under irreversible mutation.* Our formula (3.11) may also be used to obtain the frequency distribution under steady flux. In this case we

assume that the steady state is reached with respect to the distribution of intermediate gene frequencies ($0 < x < 1$), but that there is a constant flow of probability from one terminal class to the other. Such an assumption may be justified if the loss of probability by the donor terminal class is negligible, as in the case of a deleterious mutation steadily reaching fixation in a finite population at an exceedingly low rate under the pressure of irreversible mutation against the strong action of selection.

The steady flux solution may be obtained from (3.11) by applying condition (3.21) as follows. Let D be the probability flux, then

$$D = -\frac{1}{2}\frac{d}{dx}\{V_{\delta x}\phi(x)\} + M_{\delta x}\phi(x)$$

or

$$\frac{1}{2}\frac{d}{dx}\{V_{\delta x}\phi(x)\} - M_{\delta x}\phi(x) + D = 0.$$

The solution of this equation, i. e.

$$\phi(x) = \left\{C - 2D\int G(x)dx\right\}/\{V_{\delta x}G(x)\}, \tag{9.11}$$

where

$$G(x) = \exp\left\{-\int \frac{2M_{\delta x}}{V_{\delta x}}dx\right\} \tag{9.12}$$

gives the steady flux distribution, in which C is a constant. The formula (9.11) above was first obtained by Wright (1945).

In what follows I shall discuss the application of this formula to a more concrete genetical problem, and will also present some extensions of Wright's results on irreversible mutation.

Let us suppose that mutation is irreversible and occurs at a constant rate only in the direction $A_2 \to A_1$. In a finite population, A_2 will eventually be lost from the population even if A_1 is disadvantageous, because random drift may carry A_1 into fixation, and once this occurs A_2 is permanently lost from the population.

Let x be the frequency of A_1 and suppose that mutation is occurring from the class $x = 0$ (i.e. from A_2) at an exceedingly minute rate v, with irreversible fixation in the class $x = 1$. Following Wright (1942), we assume two conditions,

$$f\left(\frac{1}{2N}\right) = 4Nv \tag{9.13}$$

and

$$f\left(1 - \frac{1}{2N}\right) = 0. \tag{9.14}$$

Both of these are approximations. The first condition means that in the neighborhood of $x = 0$, mutation and random extinction balance each other so that the number of new mutations $2Nv$ is half the frequency of the subterminal class

(see equation 3.18). The second condition means that the height of the distribution curve is so low in the neighborhood of $x = 1$ as compared with the neighborhood of $x = 0$, that the frequency of the subterminal class with $x = 1 - 1/(2N)$ may be set equal to zero. Here we assume that the majority of populations contain only A_2, i.e. $f(0) = 1$ approximately.

The selective advantages of A_1A_1 and A_1A_2 over A_2A_2 may be designated by s and sh, both of which may be negative if A_1 is unconditionally deleterious. The mean rate of change in x per generation by selection may be given by

$$M_{\delta x} = s\{h + (1-2h)x\}\, x(1-x).$$

It is more convenient, however, for the subsequent treatment to express $M_{\delta x}$ in the form

$$M_{\delta x} = (s_1 + s_2 x)\, x(1-x),$$

where $s_1 = sh$ and $s_2 = s(1-2h)$. If we combine this with

$$V_{\delta x} = x(1-x)/2N$$

and substitute them in (9.11), we obtain

$$\phi(x) = \frac{2Ne^{4Ns_1x + 2Ns_2x^2}}{x(1-x)} \left\{ C - 2D \int_0^x e^{-4Ns_1\xi - 2Ns_2\xi^2}\, d\xi \right\}. \tag{9.15}$$

The relative frequency, $f(x)$, of the class with gene frequency x (discrete) will be given by $\phi(x)/(2N)$ for x between $1/(2N)$ and $1-1/(2N)$. In the above formula, C and D are constants to be determined by the two conditions (9.13) and (9.14). In the following calculation, we will assume that N becomes infinitely large and $|s_1|$ and $|s_2|$ infinitely small, while $2Ns_1$ and $2Ns_2$ remain finite.

First, from condition (9.13), we get, neglecting higher order terms,

$$\phi\left(\frac{1}{2N}\right)\frac{1}{2N} = 2N\left(C - \frac{2D}{2N}\right) = 4Nv. \tag{9.16}$$

Secondly, from (9.14),

$$\phi\left(1 - \frac{1}{2N}\right)\frac{1}{2N} = 2Ne^{4Ns_1 + 2Ns_2} \left\{ C - 2D \int_0^1 e^{-4Ns_1\xi - 2Ns_2\xi^2} d\xi \right\} = 0. \tag{9.17}$$

From the assumption that $|s_1|$ and $|s_2|$ are very small, it turns out that D/N may be neglected as compared with C in (9.16). Then, we obtain

$$C = 2v, \tag{9.18}$$

$$2D = 2v \Big/ \int_0^1 e^{-4Ns_1\xi - 2Ns_2\xi^2}\, d\xi \tag{9.19}$$

and

$$f(x) = \frac{2ve^{4Ns_1x + 2Ns_2x^2}}{x(1-x)} \cdot \frac{\int_x^1 e^{-4Ns_1\xi - 2Ns_2\xi^2} d\xi}{\int_0^1 e^{-4Ns_1\xi - 2Ns_2\xi^2} d\xi}. \tag{9.20}$$

For the special case of $s_2 = 0$ (no dominance), the above formula reduces to

$$f(x) = \frac{2v}{x(1-x)} \cdot \frac{1 - e^{-4Ns_1(1-x)}}{1 - e^{-4Ns_1}}. \tag{9.21}$$

Fisher (1930) gave the frequency distribution of mutant genes when there is a supply of one mutation in each generation. His formula is

$$df = \frac{2}{1 - e^{-4an}} \{1 - e^{-4an/(1+e^z)}\} dz, \tag{9.22}$$

where df is the frequency element, i.e. $f(x)$ in our notation, a is the selective advantage (our s_1) assumed to be very small, n is the number of breeding individuals in a population (our N) and z is the logit of the mutant gene frequency x. It is not difficult to show that (9.21) agrees with Fisher's formula (9.22) if we note that $n = N, a = s_1, x = \log_e\{x/(1-x)\}, dz = dx/\{x(1-x)\} = 1/\{2Nx(1-x)\}$ and $2Nv = 1$.

The net probability flux D will give us the probability of ultimate fixation of an individual mutant gene if we divide D by the number of mutations per generation, i,e, $2Nv$.

Thus we obtain, from (9.19),

$$u = \left(2N \int_0^1 e^{-4Ns_1x - 2Ns_2x^2} dx\right)^{-1},$$

or putting $s_1 = sh$ and $s_2 = s(1-2h)$,

$$u = \left[2N \int_0^1 e^{-4Nshx - 2Ns(1-2h)x^2} dx\right]^{-1}. \tag{9.23}$$

A more general expression for the distribution under irreversible mutation may be obtained directly from (9.11) by imposing conditions similar to (9.13) and (9.14). In the following treatment, we will take a general form of selection in M_{dx} so that selection coefficients may be gene-frequency dependent. Also, V_{dx} may include the effect of random fluctuation in selection intensities. Since the rate of change in gene frequency by selection contains the factor $x(1-x)$, V_{dx} may be expressed in the form

$$V_{\delta x} = x(1-x)/(2N_e) + x^2(1-x)^2P(x),$$

where $P(x)$ is a poynomial in x and N_e is the effective size of the population. N_e may differ from the actual number N of the individuals.

First, consider the exchange of class frequencies in the neighborhood of $x = 0$. The flux due to the new production of the mutant genes from the terminal class ($x = 0$) is $2Nv$, while the flux towards the opposite direction due to the loss of the mutant genes is

$$\frac{1}{2}\left[\frac{d}{dx}(V_{\delta x}\phi(x))\right]_{x=1/(2N)} = \frac{1}{4N_e}\phi\left(\frac{1}{2N}\right),$$

where the higher order terms are neglected. Equating these two opposite fluxes,

$$f\left(\frac{1}{2N}\right) = 4N_e v. \tag{9.24}$$

Note that the above relation reduces to (9.13) if $N_e = N$. When applied to (9.11) it leads to

$$C = 2v \tag{9.25}$$

if higher order terms are neglected.

Next, we assume as before the condition (9.14). It leads to

$$C = 2D\int_0^1 G(x)dx,$$

or

$$D = v\Big/\int_0^1 G(x)dx. \tag{9.26}$$

Substituting (9.25) and (9.26) in (9.11), we obtain a general formula for the gene frequency distribution under irreversible mutation

$$\phi(x) = \frac{2v}{V_{\delta x}G(x)}\cdot\frac{\int_x^1 G(x)dx}{\int_0^1 G(x)dx}, \tag{9.27}$$

where

$$G(x) = \exp\left\{-\int\frac{2M_{\delta x}}{V_{\delta x}}dx\right\}.$$

Note here that in calculating $M_{\delta x}$ only the effect of selection is assumed. The effect of mutation should not be included in this term for the present calculation.

The probability of fixation of an individual mutant gene may be obtained from D by dividing by $2Nv$, so that

$$u = \left[2N \int_0^1 G(x)dx\right]^{-1}. \tag{9.28}$$

As pointed out by Wright (1938a), the present treatment should have a bearing on the possible evolutionary modification caused by mutation pressure, the eye degeneration and loss of pigment of cave animals being especially suited examples.

10. Probability of fixation of mutant genes in a population

10.1. *Introductory remarks.* In the study of evolutionary genetics, it is important to know the probability of ultimate success (i.e. fixation) of mutant genes, because fixation of an advantageous gene is the key factor in the evolution of the species. Pioneering work has been done by Fisher (1922, 1930) and Haldane (1927) who obtained the approximate (but sufficiently accurate) probability of fixation of an individual mutant gene for the case of genic selection (i.e. no dominance). They made use of the method which is now standard in the treatment of branching processes. Recently, Moran (1961) was able to construct a rigorous theory for the probability of survival of a mutant gene in a finite population of a haploid organism, where complication by dominance is not involved.

Results equivalent to those of Fisher and Haldane have been obtained by Wright (1931) from the study of the frequency distribution under irreversible mutation. Also the probability for a recessive gene was estimated by Haldane (1927) and Wright (1942). Later, a more general result was obtained by the present author (Kimura, 1957) based on a diffusion model which covers any degree of dominance. The probability of eventual fixation $u(p)$ was expressed in terms of the initial frequency p, selection coefficients and the effective population number. This function was used by Robertson (1960) in his theory of selection limits in plants and animal breeding. A still more general, but quite simple expression for $u(p)$ was obtained by the author in terms of the initial frequency, and the mean and variance of the rate of change of gene frequency per generation (Kimura, 1962). It was applied to solve problems where there is a random fluctuation in selection intensity. These results by the author have been obtained by using the method of the Kolmogorov backward equation.

The method is quite far-reaching, and even enables us to obtain the probability of joint fixation of mutant genes at multiple loci under the assumption of random mating and constant (but not necessarily equal) selective values of individual genotypes. In the present article, the result for multiple loci will be presented for the first time. It enables us to study the effect of epistasis on the fixation of genes in a finite population.

10.2 *Single locus.* It was stated in Section 3 that if $u(p,t)$ is the probability of a mutant gene's reaching fixation by the tth generation, given that its frequency

is p at $t = 0$, then $u(p,t)$ satisfies equation (3.29). It was also indicated that the required probability would be obtained by solving this equation with boundary conditions (3.30). In the simplest case of random drift in a finite population of size N with no mutation and selection, the solution of the equation was given by (4.15), but in a more general case the exact solution is rather difficult to obtain.

Now let us consider the ultimate probability of fixation defined by

$$u(p) = \lim_{t \to \infty} u(p,t).$$

From the standpoint of long-term evolution, this may be the most important quantity relating to the fixation of mutant genes. Since $\partial u(p)/\partial t = 0$ for this quantity, equation (3.29) reduces to the ordinary differential equation

$$\frac{V_{\delta p}}{2} \frac{d^2u(p)}{dp^2} + M_{\delta p} \frac{du(p)}{dp} = 0 \tag{10.2}$$

with boundary conditions

$$u(0) = 0, \quad u(1) = 1. \tag{10.3}$$

Fortunately, the pertinent solution of this equation can easily be found and is expressed as follows (Kimura, 1962),

$$u(p) = \int_0^p G(x)dx \Big/ \int_0^1 G(x)dx, \tag{10.4}$$

where

$$G(x) = \exp\left\{ - \int \frac{2M_{\delta x}}{V_{\delta x}} dx \right\},$$

in which $M_{\delta x}$ and $V_{\delta x}$ are the mean and variance of the change in gene frequency x per generation.

The above formula for $u(p)$ is the steady state solution of the Kolmogorov backward equation, and is the counterpart of Wright's formula for $\phi(x)$ (cf. 9.1), which is the steady state solution of the Kolmogorov forward equation. Both formulae have a pleasing simplicity, and are yet of sufficient generality to cover the cases of sexually reproducing haploid, diploid and polyploid organisms, as well as asexually reproducing plants.

The probability of fixation of individual mutant genes in a population of N diploid individuals may then be obtained by taking $p = 1/(2N)$, so that

$$u = u\left(\frac{1}{2N}\right). \tag{10.5}$$

However, caution is necessary when applying the above method to a dioecious population where the number of males, N^*, can be different from that of females,

N^{**}. In such a case, either the initial condition $p = 1/(4N^{*})$ or $p = 1/(4N^{**})$ should be used depending on whether the mutant gene occurred in a male or in a female, as was pointed out by Moran (1961) and Watterson (1962).

In what follows I will discuss a few simple cases, assuming a population of sexually reproducing diploid individuals. We will denote by A_1 the mutant gene whose initial frequency is p.

The simplest case is that of genic selection, in which A_1 has a constant selective advantage s over its alleles in a population of effective size N_e. In this case $M_{\delta x} = sx(1-x)$, $V_{\delta x} = x(1-x)/(2N_e)$ so that $2M_{\delta x}/V_{\delta x} = 4N_e s$, $G(x) = e^{-4N_e sx}$ and we obtain from (10.4)

$$u(p) = (1 - e^{-4N_e sp})/(1 - e^{-4N_e s}). \tag{10.6}$$

For $|2N_e s| < \pi$, the right-hand side of the above equation may be expanded in terms of $4N_e s$ as follows:

$$\begin{aligned} u(p) &= \sum_{i=1}^{\infty} \frac{\phi_i(p)(-1)^{i-1}}{i!}(4N_e s)^{i-1} \\ &= p + 2N_e sp(1-p) + \frac{(2N_e s)^2}{3}p(p-1)(2p-1) + \cdots, \end{aligned} \tag{10.7}$$

where the $\phi_i(p)$'s are Bernoulli polynomials. Thus for a small value of $2N_e s$, $u(p) - p$ is $2N_e$ times $sp(1-p)$. In other words, the total advance is $2N_e$ times the change in the first generation, as was pointed out by Robertson (1960).

If the effective size of the population is equal to the actual size, N may be substituted for N_e. Assuming the sex ratio is unity, the probability of fixation of an individual mutant gene is obtained from formula (10.6) by putting $p = 1/(2N)$. If $|s|$ is small, we obtain

$$u = (2s)/(1 - e^{-4Ns}) \tag{10.8}$$

as a good approximation. This agrees with the result obtained by Fisher (1930) and by Wright (1931) using different methods. For a positive s and very large N, we obtain the well known result that the probability of ultimate survival of an advantageous mutant gene is approximately twice the selection coefficient (Haldane, 1927). If N_e differs from N, this value should be modified by a factor of N_e/N so that

$$u = 2s(N_e/N). \tag{10.9}$$

According to Crow (1954) and also Crow and Morton (1955), estimated values of N_e/N for a few cases are: *Drosophila* $0.48 \sim 0.9$, *Lymnaea* 0.75, Man $0.69 \sim 0.95$.

The above results were obtained on the basis of the method of diffusion approximation, and it is desirable to check some of them by a rigorous treatment. This was done by Moran (1961) who used a population model consisting of

haploid individuals, with offspring number following a negative binomial distribution. He assumed that in each generation, the population consists of exactly M individuals of which a random number k are of one type (say A_1) and $M-k$ of the other type (say A_2). Suppose that the generating function for the probability distribution of the number of offspring is

$$P_2(z) = \left(\frac{a}{1-bz}\right)^{a/b}$$

for gene A_2 and

$$P_1(z) = \left(\frac{a}{1-bz}\right)^{\lambda a/b}$$

for gene A_1, where $b = 1-a$ and $\lambda = 1+s$ with $s = O(M^{-1})$. The mean and variance of the distribution generated by $P_2(z)$ are 1 and a^{-1} respectively. The mean of the distribution generated by $P_1(z)$ is $\lambda = 1+s$, so that s represents the selective advantage of A_1 over A_2. It was then shown by Moran that the probability of ultimate fixation of A_1, say P_1, satisfies the relation

$$\frac{1-e^{-\theta_0 k_0}}{1-e^{-\theta_0 M}} \leqq P_1 \leqq \frac{1-e^{-\theta_1 k_0}}{1-e^{-\theta_1 M}}, \tag{10.10}$$

where k_0 is the initial number of A_1-genes, and that for large M both θ_0 and θ_1 become asymptotically equal to $2sv^{-1}$, where $v = a^{-1}$ is the variance generated by $P_2(z)$. If we denote by X the proportion of A_1 genes in the haploid population ($X = k/M$) and if A_1 were selectively neutral, the variance of X per generation would be $vX(1-X)/(M-1)$. The effective size of the population, say M_e may then be defined by equating this variance with the binomial variance $X(1-X)/M_e$, so that asymptotically $v^{-1} = M_e M^{-1}$. Since both θ_0 and θ_1 become asymptotically equal to $2sv^{-1}$ or $2sM_eM^{-1}$ and since Moran's M and k_0 correspond respectively to our $2N$ and $2Np$, (10.10) shows that (10.6) is asymptotically correct.

Next, we will consider a more general case of zygotic selection. Let s and sh be respectively the selective advantages of the mutant homozygote (A_1A_1) and the heterozygote (A_1A_2) such that

$$M_{\delta x} = sx(1-x)[h+(1-2h)x].$$

For a random mating population of effective size N_e, we have

$$V_{\delta x} = x(1-x)/(2N_e)$$

and therefore

$$G(x) = \exp\{-4N_eshx - 2N_es(1-2h)x^2\}. \tag{10.11}$$

The ultimate probability of fixation in this case is obtained from (10.4) with $G(x)$ given by (10.11). The probability of fixation of an individual mutant gene may then be obtained by putting $p = 1/(2N)$ in $u(p)$. Assuming that $|s|$ and $|sh|$ are small, we obtain

$$u = \left[2N \int_0^1 e^{-4N_e shx - 2N_e s(1-2h)x^2} dx\right]^{-1}. \tag{10.12}$$

This agrees with (9.23) which was obtained in the previous section by a differen method.

If the mutant gene is completely recessive ($h = 0$), but advantageous in the homozygous condition ($s > 0$) and if $N_e s$ is large, we have approximately

$$u = \frac{\sqrt{2N_e s}}{N\sqrt{\pi}}. \tag{10.13}$$

For an idealized situation in which $N_e = N$, the above formula reduces to

$$u = \sqrt{\frac{2s}{\pi N}} \approx 1.128 \sqrt{\frac{s}{2N}} \tag{10.14}$$

(Kimura, 1957). It is interesting to note that the value given in (10.14) above lies between $\sqrt{s/N}$, the value obtained by Haldane (1927) treating this as a branching process, and $\sqrt{s/(2N)}$, obtained by Wright (1942) using his method of integral equations. Wright's numerical approximation, $1.1\sqrt{s/(2N)}$, is very close to (10.14).

Finally I will discuss briefly the effect of random fluctuation of selection intensity on the fixation of mutant genes.

The simplest situation is again genic selection in which the mutant gene A_1 has a selective advantage s over its alleles, but s is now a random variable with mean $\bar{s}$ and variance V_s. Thus

$$\begin{cases} M_{\delta x} = \bar{s}x(1-x) \\ V_{\delta x} = V_s x^2(1-x)^2 + x(1-x)/(2N_e) \end{cases} \tag{10.15}$$

and $G(x)$ in (10.4) is given by

$$G(x) = \left(\frac{\lambda_1 - x}{x - \lambda_2}\right)^{k/\sqrt{1+4r}}$$

where

$$k = 2\bar{s}/V_s$$

and

$$\begin{cases} \lambda_1 = \frac{1}{2}(1 + \sqrt{1+4r}) \\ \lambda_2 = \frac{1}{2}(1 - \sqrt{1+4r}) \end{cases}$$

in which $r = 1/(2N_e V_s)$.

A few special cases have been studied by the author (Kimura, 1962), but the results obtained are not very satisfactory and further study is required.

10.3. *Multiple loci.* Let us consider n independent loci each with a pair of alleles, conveniently called the normal and the mutant genes, and denote by $p^{(i)}$ the initial frequency of the mutant gene in the ith locus ($i = 1, 2, \cdots, n$). Let $u(p^{(1)}, p^{(2)}, \cdots, p^{(n)}; t)$ be the probability that all the n mutant genes become fixed in the population by the tth generation, given that their initial frequencies at $t=0$ are $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ respectively. As in the single locus case, we will be mainly concerned with the ultimate probability of joint fixation defined by

$$u(p^{(1)}, p^{(2)}, \cdots, p^{(n)}) = \lim_{t \to \infty} u(p^{(1)}, p^{(2)}, \cdots, p^{(n)}; t). \tag{10.16}$$

Under the assumption of random mating and constant fitness of individual genotypes, it is possible to show that

$$u(p^{(1)}, p^{(2)}, \cdots, p^{(n)}) = \frac{\int_0^{p^{(1)}} \int_0^{p^{(2)}} \cdots \int_0^{p^{(n)}} e^{-2N\bar{a}} dx^{(1)} dx^{(2)} \cdots dx^{(n)}}{\int_0^1 \int_0^1 \cdots \int_0^1 e^{-2N\bar{a}} dx^{(1)} dx^{(2)} \cdots dx^{(n)}}, \tag{10.17}$$

where $\bar{a} = \bar{a}(x^{(1)}, x^{(2)}, \cdots, x^{(n)})$ is the average fitness of the population measured in Malthusian parameters, and is a function of the gene frequencies $(x^{(1)}, \cdots, x^{(n)})$, in which $x^{(i)}$ $(i = 1, \cdots, n)$ stands for the frequency of the mutant gene at the ith locus in the population. If selective values are used to measure the fitness of individual genotypes (discrete generation time model), $\log \bar{w}$ may be substituted for $\bar{a}$ in the above formula, where w is the relative selective value in the sense used by Wright; w coincides with the relative viability if no fertility differences are involved. The above formula (10.17) can be obtained as the steady state solution of the Kolmogorov backward equation for n variables, assuming random mating, and constant but not necessarily equal fitness of individual genotypes. The formula enables us to study the effect of epistasis in fitness on the chance of joint fixation of mutant genes. To make our discussion simpler, let us choose a population of sexually reproducing haploid organisms and con-

TABLE 1

Genotype	Fitness	Frequency
A_1B_1	$s_1 + s_2 + \varepsilon$	xy
A_1B_2	s_1	$x(1-y)$
A_2B_1	s_2	$(1-x)y$
A_2B_2	0	$(1-x)(1-y)$

sider epistasis involving two loci ($n = 2$) as in Table 1. In this table, A_1 and B_1 denote respectively the mutant genes at the first and second loci, each with respective frequencies x and y. Here, fitness is measured in Malthusian parameters, taking the fitness of A_2B_2 as the standard. Thus ε represents the amount of epistasis in fitness between the mutant genes. If the discrete generation time model is used, the value 1 may be added to each entry of the second column to represent fitness in selective values (= relative viability if no fertility differences are involved). However, as long as the selection coefficients s_1, s_2 and ε are small, there will be no practical difference in the results, whichever model is used to measure fitness.

If p and q are the respective initial frequencies of A_1 and B_1, then from (10.17), we have

$$u(p,q) = \int_0^p \int_0^q e^{-2N\bar{a}(x,y)}dxdy \Big/ \int_0^1 \int_0^1 e^{-2N\bar{a}(x,y)}dxdy, \tag{10.18}$$

where

$$\begin{aligned}\bar{a}(x,y) &= (s_1 + s_2 + \varepsilon)xy + s_1x(1-y) + s_2(1-x)y \\ &= s_1x + s_2y + \varepsilon xy.\end{aligned}$$

Writing $Ns_1 = S_1$, $Ns_2 = S_2$, $N\varepsilon = I$ and assuming that the initial frequencies, p and q, are both very low so that

$$|\bar{a}(p,q)| \ll 1,$$

then we have approximately,

$$u(p,q) = pqJ^{-1}, \tag{10.19}$$

where

$$J = \int_0^1 \int_0^1 e^{-2S_1x - 2S_1y - 2Ixy}\,dxdy. \tag{10.20}$$

In the simplest case where $S_1 = S_2 = 0$ and $I > 0$, namely in the case in which mutant genes are neutral separately but advantageous when combined, the above integral reduces to

$$J = \frac{1}{2I}[\log_e(2I) + \gamma + E_i(2I)], \tag{10.21}$$

where γ is Euler's constant (0.5772...) and $E_i(\cdot)$ denotes the exponential integral defined by

$$E_i(z) = \int_z^\infty \frac{e^{-t}}{t}dt, \tag{10.22}$$

for which fairly extensive tabulation is available. As an example, let us take $N = 500$ and $\varepsilon = 0.01$. Then $2I = 10$, and $E_i(10) \approx e^{-10}/10$ is entirely negligible

as compared with $\log_e 10 \approx 2.30$ and $\gamma \approx 0.58$, giving $J \approx 0.29$ or $J^{-1} \approx 3.5$. Thus, we obtain

$$u(p, q) \approx 3.5\, pq\,.$$

Since the probability of joint fixation of the two mutant genes at two independent loci (if they were completely neutral) is equal to pq, the above result shows that with 1% epistasis, in a population of $N = 500$, this probability is increased by a factor of about 3.5.

Let us now consider the biologically more interesting case in which each mutant gene is deleterious separately ($s_1 < 0$, $s_2 < 0$), but becomes advantageous when combined ($s_1 + s_2 + \varepsilon > 0$).

Writing $S_1 = -S_1'$ and $S_2 = -S_2'$ so that $S_1' > 0$ and $S_2' > 0$, (10.20) can be expressed as follows:

$$(10.23)\quad J = \frac{e^{2S_1'S_2'/I}}{2I}\left[E_i\left(\frac{2S_1'S_2'}{I}\right) + E_i\left(2I\left(1 - \frac{S_1'}{I}\right)\left(1 - \frac{S_2'}{I}\right)\right)\right.$$
$$\left. + \bar{E}_i\left(2S_1'\left(1 - \frac{S_2'}{I}\right)\right) + \bar{E}_i\left(2S_2'\left(1 - \frac{S_1'}{I}\right)\right)\right]\,,$$

where $E_i(\cdot)$ is the exponential integral defined in (10.22), while $\bar{E}_i(\cdot)$ is another exponential integral defined by

$$(10.24)\qquad \bar{E}_i(z) = \text{P.V.}\int_{-\infty}^{-z} \frac{e^{-t}}{t}dt \quad (\text{P.V} = \text{principal value}).$$

For example, in a population of $N = 10^3$, if $s_1 = s_2 = -0.01$ (single mutant, 1% disadvantage) and $\varepsilon = 0.07$ (double mutant, 5% advantage), $S_1' = S_2' = 10$, $I = 70$ and J turns out to be about 4.85×10^8. Thus, the probability of fixation is approximately

$$J^{-1} \approx 2.5 \times 10^{-6}$$

times the corresponding value for neutral genes. This may mean that, in this specific example, the probability is too low for this pair of mutant genes to be of much use in evolution.

A more detailed account of this subject together with the derivation of (10.17) will be published elsewhere.

References

[1] Bodmer, W. F. (1960) Discrete stochastic processes in population genetics. *J.R. Statist. Soc.* B **22**, 218–244.

[2] Cavalli-Sforza, L.L. and Conterio, F. (1960) Analisi della fluttuazione di frequenze geniche nella popolazione della Val Parma. *Atti* A.G.I. **5**, 333–344.

[3] Crow, J.F. (1954) Breeding structure of populations II. Effective population number. *Statistics and Mathematics in Biology*. Kempthorne *et al.* (ed). Iowa State College Press, Ames, Iowa.

[4] Crow, J.F. and Morton, N. (1955) Measurement of gene frequency drift in small populations. *Evolution* **9**, 202–214.

[5] CROW, J. F. AND KIMURA, M. (1956) Some genetic problems in natural populations. *Proc. Third Berkeley Symp. on Math. Statist. and Prob.* **4**, 1–22.

[6] EWENS, W. J. (1963) Numerical results and diffusion approximations in a genetic process. *Biometrika* **50**, 241–249.

[7] EWENS, W. J. (1964) The pseudo-transient distribution and its uses in genetics. *J. Appl. Prob.* **1**, 141–156.

[8] FELLER, W. (1951) Diffusion processes in genetics. *Proc. Second Berkeley Symp. on Math. Statist. and Prob.* 227–246.

[9] FELLER, W. (1952) The parabolic differential equations and the associated semigroup of transformations. *Ann. Math.* **55**, 468–519.

[10] FISHER, R.A. (1922) On the dominance ratio. *Proc. Roy. Soc. Edin.* **42**, 321–341.

[11] FISHER, R.A. (1930) The distribution of gene ratios for rare mutations. *Proc. Roy. Soc. Edin.* **50**, 205–220.

[12] FISHER, R.A. (1953) Population genetics. *Proc. Roy. Soc. London* B **141**, 510–523.

[13] FISHER, R.A. (1958) *The Genetical Theory of Natural Selection* (2nd revised ed.). Dover, New York.

[14] FOKKER, A.D. (1914) Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld. *Ann. d. Phys.* **43**, 810–820.

[15] GOLDBERG, S. (1950) *On a singular diffusion equation.* Ph. D. thesis (unpublished). Cornell University.

[16] HALDANE, J.B.S. (1924) A mathematical theory of natural and artificial selection. Part I. *Trans. Camb. Phil. Soc.* **23**, 19–41.

[17] HALDANE, J.B.S. (1927) A mathematical theory of natural and artificial selection. Part V: Selection and mutation. *Proc. Camb. Phil. Soc.* **23**, 838–844.

[18] HALDANE, J.B.S. (1932) *The Causes of Evolution.* Harper and Brothers, New York.

[19] HALDANE, J.B.S. (1949) Suggestions as to quantitative measurement of rates of evolution. *Evolution* **3**, 51–56.

[20] KARLIN, S. AND MCGREGOR, J. (1962) On a genetics model of Moran. *Proc. Camb. Phil. Soc.* **58**, 299–311.

[21] KIMURA, M. (1954) Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics* **39**, 280–295.

[22] KIMURA, M. (1955 a) Solution of a process of random genetic drift with a continuous model. *Proc. Nat. Acad. Sci.* **41**, 144–150.

[23] KIMURA, M. (1955 b) Random genetic drift in multi-allelic locus. *Evolution* **9**, 419–435.

[24] KIMURA, M. (1955 c) Stochastic processes and distribution of gene frequencies under natural selection. *Cold Spring Harbor Symp.* **20**, 33–53.

[25] KIMURA, M. (1956 a) Random genetic drift in a tri-allelic locus: Exact solution with a continuous model. *Biometrics* **12**, 57–66.

[26] KIMURA, M. (1956 b) *Stochastic processes in population genetics.* Ph. D. thesis (unpublished). Univ. of Wisconsin.

[27] KIMURA, M. (1957) Some problems of stochastic processes in genetics. *Ann. Math. Statist.* **28**, 882–901.

[28] KIMURA, M. (1958) On the change of population fitness by natural selection. *Heredity* **12**, 145–167.

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

[30] KIMURA, M. AND CROW, J.F. (1963) The measurement of effective population number. *Evolution* **17**, 279–288.

[31] KOLMOGOROV, A. (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104**, 415–458.

[32] MALÉCOT, G. (1948) *Les Mathématiques de l'Hérédité*. Masson et Cie, Paris.

[33] MILLER, G.F. (1962) The evaluation of eigenvalues of a differential equation arising in a problem in genetics. *Proc. Camb. Phil. Soc.* **58**, 588–593.

[34] MORAN, P.A.P. (1958 a) A general theory of the distribution of gene frequencies. I. Overlapping generations. *Proc. Roy. Soc. London.* B **149**, 102–112.

[35] MORAN, P.A.P. (1958 b) A general theory of the distribution of gene frequencies. II. Nonoverlapping generations. *Proc. Roy. Soc. London* B **149**, 113–116.

[36] MORAN, P.A.P. (1958 c) Random processes in genetics. *Proc. Camb. Phil. Soc.* **54**, 60–71.

[37] MORAN, P.A.P. (1961) The survival of a mutant under general conditions. *Proc. Camb. Phil. Soc.* **57**, 304–314.

[38] MORAN, P.A.P. (1962) *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford.

[39] MORSE, P.M. AND FESHBACH, H. (1953) *Methods of Theoretical Physics*. McGraw-Hill, New York.

[40] PLANCK, M. (1917) Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie. *Sitz. der. preuss. Akad.* 324–341.

[41] ROBERTSON, A. (1960) A theory of limits in artificial selection. *Proc. Roy. Soc. London* B **153**, 234–249.

[42] ROBERTSON, A. (1962) Selection for heterozygotes in small populations. *Genetics* **47**, 1291–1300.

[43] STRATTON, J.A., MORSE, P.M., CHU, L.J., AND HUTNER, R.A. (1941) *Elliptic Cylinder and Spheroidal Wave Functions*. John Wiley, New York.

[44] STRATTON, J.A., MORSE, P.M., CHU, L.J., LITTLE, J.D.C., AND CORBATÓ, F.J. (1956) *Spheroidal Wave Functions*. Technology Press of M.I.T. & John Wiley, New York.

[45] WATTERSON, G.A. (1962) Some theoretical aspects of diffusion theory in population genetics. *Ann. Math. Stat.* **33**, 939–957.

[46] WATTERSON, G.A. (1964) The application of diffusion theory to two population genetic models of Moran. *J. Appl. Prob.* **1**, 233–246.

[47] WRIGHT, S. (1931) Evolution in Mendelian populations. *Genetics* **16**, 97–159.

[48] WRIGHT, S. (1937) The distribution of gene frequencies in populations. *Proc. Nat. Acad. Sci.* **23**, 307–320.

[49] WRIGHT, S. (1938 a) The distribution of gene frequencies under irreversible mutation. *Proc. Nat. Acad. Sci.* **24**, 253-259.

[50] WRIGHT, S. (1938 b) The distribution of gene frequencies in populations of polyploids. *Proc. Nat. Acad. Sci.* **24**, 372–377.

[51] WRIGHT, S. (1939) The distribution of self-sterility alleles in populations. *Genetics* **24**, 538–552.

[52] WRIGHT, S. (1942) Statistical genetics and evolution. *Bull. Amer. Math. Soc.* **48**, 223–246.

[53] WRIGHT, S. (1945) The differential equation of the distribution of gene frequencies. *Proc. Nat. Acad. Sci.* **31**, 382–389.

[54] WRIGHT, S. (1948) On the roles of directed and random changes in gene frequency in the genetics of populations. *Evolution* **2**, 279–294.

[55] WRIGHT, S. (1949) Adaptation and selection. In *Genetics, Paleontology, and Evolution*. Jepsen *et al.* (ed.), Princeton Univ. Press.

[56] WRIGHT, S. (1950) Population structure as a factor in evolution. In *Moderne Biologie, Festschrift für Hans Nachtsheim*, F. W. Peters, Berlin.

[57] WRIGHT, S. (1952) The theoretical variance within and among subdivisions of a population that is in a steady state. *Genetics* **37**, 312–321.

[58] WRIGHT, S. AND KERR, W.E. (1954) Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*. II. Bar. *Evolution* **8**, 225–240.

References

[4] **Measurement of Gene Frequency Drift in Small Populations**
James F. Crow; Newton E. Morton
Evolution, Vol. 9, No. 2. (Jun., 1955), pp. 202-214.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195506%299%3A2%3C202%3AMOGFDI%3E2.0.CO%3B2-G

[6] **Numerical Results and Diffusion Approximations in a Genetic Process**
W. J. Ewens
Biometrika, Vol. 50, No. 3/4. (Dec., 1963), pp. 241-249.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196312%2950%3A3%2F4%3C241%3ANRADAI%3E2.0.CO%3B2-E

[19] **Suggestions as to Quantitative Measurement of Rates of Evolution**
J. B. S. Haldane
Evolution, Vol. 3, No. 1. (Mar., 1949), pp. 51-56.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28194903%293%3A1%3C51%3ASATQMO%3E2.0.CO%3B2-4

[23] **Random Genetic Drift in Multi-Allelic Locus**
Motoo Kimura
Evolution, Vol. 9, No. 4. (Dec., 1955), pp. 419-435.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195512%299%3A4%3C419%3ARGDIML%3E2.0.CO%3B2-Z

LINKED CITATIONS

[25] **Random Genetic Drift in a Tri-Allelic Locus; Exact Solution with a Continuous Model**
Motoo Kimura
Biometrics, Vol. 12, No. 1. (Mar., 1956), pp. 57-66.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28195603%2912%3A1%3C57%3ARGDIAT%3E2.0.CO%3B2-6

[30] **The Measurement of Effective Population Number**
Motoo Kimura; James F. Crow
Evolution, Vol. 17, No. 3. (Sep., 1963), pp. 279-288.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28196309%2917%3A3%3C279%3ATMOEPN%3E2.0.CO%3B2-U

[54] **On the Roles of Directed and Random Changes in Gene Frequency in the Genetics of Populations**
Sewall Wright
Evolution, Vol. 2, No. 4. (Dec., 1948), pp. 279-294.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28194812%292%3A4%3C279%3AOTRODA%3E2.0.CO%3B2-0

[58] **Experimental Studies of the Distribution of Gene Frequencies in Very Small Populations of Drosophila melanogaster. II. Bar**
Sewall Wright; Warwick E. Kerr
Evolution, Vol. 8, No. 3. (Sep., 1954), pp. 225-240.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195409%298%3A3%3C225%3AESOTDO%3E2.0.CO%3B2-7