Appendix 1

Diffusion Theory

Draft Version 18 March 2014

Exact solutions of the dynamics of many random processes arising in population and quantitative genetics are either unknown or are extremely cumbersome. Nevertheless, starting with Fisher (1922), Wright (1945), and Kimura (1955a,b) the use of **diffusion approximations** to closely approximate the exact dynamics of these processes has proven to be extremely powerful. Useful introductions to diffusion theory with special reference to genetics are given by Ewens (1979, 2004) and Karlin and Taylor (1981), with additional applications being given by Crow and Kimura (1970), Maruyama (1977), Kimura (1983), and Gale (1990). The goal of diffusion theory is to obtain expressions for $\varphi(x,t,p)$, the probability distribution for the random variable x at time t given that the process starts at value p. It is often the case that for sufficiently large time, the probability distribution approaches a stationary value $\varphi(x)$, independent of both time and starting value. Diffusion theory also provides approximations of a number of summary statistics of a particular process, such as probablities and times to fixation for the various boundaries of the process. We consider these issues in turn.

FOUNDATIONS OF DIFFUSION THEORY

Consider a continuous random variable x_t indexed by continuous time t. If $\delta_x = x_{t+\delta_t} - x_t$ (the change in x_t over a very small time interval δ_t) satisfies

$$E\left(\delta_x \,|\, x_t = x\right) = m(x)\delta_t + o(\delta_t) \tag{A1.1a}$$

$$\sigma^2(\delta_x \,|\, x_t = x) = v(x)\delta_t + o(\delta_t) \tag{A1.1b}$$

$$E\left(\,|\delta_x|^k\,\right) = o(\delta_t) \quad \text{for } k \geq 3 \tag{A1.1c}$$

then x_t is said to be a **diffusion process** (provided the additional technical restriction that x_t is a **Markov process** — the transitition probabilities depend only on the current value of the process and no other aspects of its history – is satisfied). The notation $o(\delta_t)$ means that any remaining terms are of order δ_t^2 (or higher), and hence small relative to δ_t. Formally, $\lim_{\delta_t \to 0} o(\delta_t)/\delta_t = 0$, so that terms of order δ_t^2 (or higher) are $o(\delta_t)$, and we can ignore these when δ_t itself is very small.

The Infinitesimal Mean $m(x)$ and Variance $V(x)$

The **infinitesimal mean** $m(x)$ and **infinitesimal variance** $v(x)$ correspond to the mean and variance of the process (given it is at x) over a very small time interval. These are formally defined as

$$m(x) = \lim_{\delta_t \to 0} \frac{E\left(x_{t+\delta_t} - x_t \,|\, x_t = x\right)}{\delta_t} \tag{A1.2a}$$

$$v(x) = \lim_{\delta_t \to 0} \frac{E[\,(x_{t+\delta_t} - x_t)^2 \,|\, x_t = x\,]}{\delta_t} \tag{A1.2b}$$

In words, the diffusion assumptions are that, over a very small time interval, the mean $m(x)$ and variance $v(x)$ around the current position x are sufficient to fully describe the process. Thus, the entire structure of a diffusion is described by the expected mean change $m(x)$ in our variable of interest and its variance $v(x)$ conditioned on the process currently having value x.

Example A1.1. Diffusion processes for allele frequencies are typically obtained by setting $v(x) = x(1-x)/(2N_e)$ (the per generation variance in the change of allele frequencies due to drift, see Chapter 2), where x is the frequency of allele A. The deterministic (i.e., infinite-population) change in allele frequency is subsumed in $m(x)$. For example, since there is no deterministic change under pure drift, the diffusion process is given by

$$m(x) = 0, \qquad v(x) = \frac{x(1-x)}{2N_e} \tag{A1.3}$$

and is defined for $0 < x < 1$.

Now consider additive selection at a diallelic locus (with fitnesses $1 : 1 + s : 1 + 2s$ for the genotypes aa, Aa, and AA, respectively) when $|s|$ is small, so that mean fitness $\overline{W} \simeq 1$. Recalling Equation 5.2, in an infinite population $\Delta p = sp(1-p) + o(s)$. Ignoring terms of $o(s)$ gives a diffusion approximation for the joint effects of drift and selection over $0 < x < 1$ as

$$m(x) = sx(1-x), \qquad v(x) = \frac{x(1-x)}{2N_e} \tag{A1.4}$$

Finally, consider arbitrary selection with constant fitnesses (again with selection sufficiently weak such that $\overline{W} \simeq 1$) and forward and back mutation, such that ν is the mutation rate from A to a and μ is the mutation rate from a to A. Applying Wright's formula (Equation 5.5) and including mutation gives

$$m(x) = \frac{x(1-x)}{2}\frac{d\ln(\overline{W})}{dx} + (1-x)\mu - x\nu, \qquad v(x) = \frac{x(1-x)}{2N_e} \tag{A1.5}$$

In population genetics, diffusion approximations provide an elegant way to rescale a discrete space, discrete time, random variable (usually an allele frequency) to construct a new continuously distributed random variable. For example, consider X_t, the number of copies of allele A in a discrete-generation population of N diploids at generation t. X_t takes on values $0, 1, \ldots, 2N$, and time t is in discrete units of generations. Suppose we construct a new random variable $x_\tau^{(N)} = X_{(\tau)}/(2N)$, where $\tau = t/N$. A unit of time on this transformed corresponds to N generations on the original scale. Taking the limit as N approaches infinity, the limiting process x_τ is a continuous space (in x), continuous time (in τ) process that represents the allele frequency at time τ.

The Kolmogorov Forward Equation

Given $m(x)$, $v(x)$, and initial frequency p, the diffusion approximation for the dynamics of the probability density function for the realized frequency x at time t satisfies the **Kolmogorov forward equation** (or **KFE**),

$$\frac{\partial\,\varphi(x,t,p)}{\partial\,t} = \frac{1}{2}\frac{\partial^2\,v(x)\,\varphi(x,t,p)}{\partial\,x^2} - \frac{\partial\,m(x)\,\varphi(x,t,p)}{\partial\,x} \tag{A1.6}$$

where $\varphi(x,t,p)$ is the probability density for x at time t given the process starts at p, such that

$$\Pr[\, c \le x(t) \le d \,|\, x(0) = p \,] = \int_c^d \varphi(x,t,p)\, dx$$

When they can be found, closed-form solutions of $\varphi(x,t,p)$ are usually complex. The standard approach is to express the solution as a power series,

$$\varphi(x,p,t) = \sum_{i=1}^{\infty} f_i(x,p) e^{-\lambda_i t} \tag{A1.7}$$

The λ_i values are the **eigenvalues** associated with the particular differential equation, and f_i the associated **eigenfunctions**. Crow and Kimura (1970) give exact solutions of φ for a number of population genetic problems.

Example A1.2. The KFE for the diffusion for pure drift ($m(x) = 0$) is

$$\frac{\partial\, \varphi(x,t,p)}{\partial\, t} = \frac{1}{2} \frac{\partial^2\, v(x)\, \varphi(x,t,p)}{\partial\, x^2}$$

Figure 2.2 plots solutions of this partial differential equation for several time points. As mentioned in Chapter 2, Kimura (1955b) obtained the solution for this particular KFE when one starts with the allele at frequency p as

$$\varphi(x,t,p) = \sum_{i=1}^{\infty} p(1-p) i(i+1)(2i+1) g_i(p) g_i(x) e^{-\lambda_i t}$$

where

$$\lambda_i = \frac{i(i+1)}{4N} \quad \text{and} \quad g_i(x) = F(1-i, i+2, 2, x)$$

with F being the hypergeometric function (Abramowitz and Stegun 1972). In the notation of Equation A1.7, the resulting eigenfunctions are

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

This solution is obtained using standard tools from the solution of partial differential equations. As one might expect given the form of the solution, this is not a trivial procedure!

While simple inspection of the above closed form solution provides only a small amount of insight into the process, an immediate one is that $\varphi(x,t,p) \to 0$ as $t \to \infty$ (as $e^{-\lambda_i t} \to 0$ as $t \to \infty$). Thus, for large time, there is vanishingly small probability mass in the open interval $0 < x < 1$. This occurs because drift alone will result in an allele eventually either being fixed ($x = 1$) or lost ($x = 0$). Other forces, such as mutation, are required for non-zero probability mass to remain in the interior space ($0 < x < 1$). For large time, the solution is dominated by the first few eigenvalues,

$$\varphi(x,t,p) \simeq 6p(1-p) e^{-t/(2N)} + 30p(1-2p)(1-2x) e^{-3t/(2N)}$$

Notice that the eigenfunction for the dominant eigenvalue ($1/2N$) is independent of x, so that for large t, the probability density is a constant independent of x (i.e., a uniform over $0 < x < 1$), which decays to zero at rate $\exp[-t/(2N)]$.

Boundary Behavior of a Diffusion

A critical point about the density $\varphi(x,t,p)$ concerns its range of validity. Formally speaking, the diffusion approximation only applies within some *open* interval bounded by two **boundary values** a and b. The behavior *exactly* at the boundaries ($x=a$, $x=b$) is beyond the realm of the approximation. In many cases, x_t does not change value once it reaches a boundary, in which case it is called **absorbing**. For example, in the absence of mutation and migration, once an allele frequency reaches either 0 or 1, it remains there. For this case, both 0 and 1 are absorbing boundaries. Further, a boundary is said to be **accessible** if it can be reached in finite time. When we consider fixation probabilities, we are typically considering absorbing, accessible boundaries. Note that a finite boundary point may not be accessible. Consider a simple mutation-drift equilibrium. If the mutational pressure is sufficiently strong near a boundary (say for the loss of an allele), the resulting boundary may be unaccessible, with the population never being in a state where all copies of an allele are lost.

Derivation of The Kolmogorov Forward Equation

For completeness, we present a derivation of the Equation A1.6, using the diffusion approximations (Equations A1.1a-A1.1c). This section is a bit technical and can be skipped if so desired. Consider the change from the probability distribution $\varphi(x,t,p)$ at time t to a new distribution $\varphi(x,t+\delta_t,p)$ after some very small time interval δ_t. To arrive at x at time $t+\delta_t$ the frequency must have previously been at some value $x-\delta_x$ and then moved by an amount δ_x over the interval δ_t. Let $\phi(\delta_x,x,t)$ be the probability of jumping by an amount δ_x over the time t given the starting point x. Integrating over all possible jump values gives the **Chapman-Kolmogorov equation**,

$$\varphi(x,t+\delta_t,p)=\int \varphi(x-\delta_x,t,p)\,\phi(\delta_x,x-\delta_x,\delta_t)\,d\delta_x \tag{A1.8}$$

To simplify the following derviation, we will write $\varphi(x,t,p)$ as $\varphi(x,t)$, although the dependence on the intial value p should be kept in mind.

To solve the Chapman-Kolmogorov equation, we expand the function in the integral as a Taylor series,

$$\begin{aligned}\varphi(x-\delta_x,t)\,\phi(\delta_x,x-\delta_x,\delta_t) &= \varphi(x,t)\,\phi(\delta_x,x,\delta_t)-\delta_x\frac{\partial\,[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x}\\ &+\frac{\delta_x^2}{2}\frac{\partial^2[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x^2}-\frac{\delta_x^3}{6}\frac{\partial^3[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x^3}+o(\delta_x^3)\end{aligned}$$

Ignoring terms of $o(\delta_x^3)$, i.e., order δ_x^4 and higher, yields

$$\begin{aligned}\int \varphi(x-\delta_x,t)\,\phi(\delta_x,x-\delta_x,\delta_t)\,d\delta_x &= \int \varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,d\delta_x-\int \delta_x\frac{\partial[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x}\,d\delta_x\\ &+\int\frac{\delta_x^2}{2}\frac{\partial^2[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x^2}\,d\delta_x-\int\frac{\delta_x^3}{6}\frac{\partial^3[\,\varphi(x,t)\,\phi(\delta_x,x,\delta_t)\,]}{\partial x^3}\,d\delta_x\end{aligned}$$

Because the integration is with respect of δ_x, partials with respect to x can be moved outside of the integrals, as can functions not involving δ_x, giving

$$\begin{aligned}\int \varphi(x-\delta_x,t)\phi(\delta_x,x-\delta_x,\delta_t)d\delta_x &= \varphi(x,t)\int\phi(\delta_x,x,\delta_t)d\delta_x-\frac{\partial}{\partial x}\left(\varphi(x,t)\int\delta_x\,\phi(\delta_x,x,\delta_t)d\delta_x\right)\\ &+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\varphi(x,t)\int\delta_x^2\,\phi(\delta_x,x,\delta_t)\,d\delta_x\right)-\frac{1}{6}\frac{\partial^3}{\partial x^3}\left(\varphi(x,t)\int\delta_x^3\,\phi(\delta_x,x,\delta_t)d\delta_x\right)\end{aligned}$$

Because $\phi(\delta_x, x, \delta_t)$ is the distribution of moves of size δ_x over the amount δ_t given we start at position x, then

$$\int \phi(\delta_x, x, \delta_t)\, d\delta_x = 1, \quad \int \delta_x\, \phi(\delta_x, x, \delta_t)\, d\delta_x = m(x)\delta_t + o(\delta_t)$$

The first identity follows from the fact that the integral over a probability distribution is equal to one, and the second is just our first diffusion assumption (Equation A1.1a). Next, recalling that $E(x^2) = \sigma_x^2 + [E(x)]^2$,

$$\int \delta_x^2 \phi(\delta_x, x, \delta_t) d\delta_x = \sigma^2(\delta_x) + [E(\delta_x)]^2$$

which using diffusion approximations A1.1a, b reduces to

$$\begin{aligned}\int \delta_x^2 \phi(\delta_x, x, \delta_t) d\delta_x &= v(x)\delta_t + o(\delta_t) + [\, m(x)\delta_t + o(\delta_t)\,]^2 \\ &= v(x)\delta_t + o(\delta_t)\end{aligned}$$

The last step follows because the contribution from the squared change in the mean is $o(\delta_t)$, and we sweep all such terms into a single expression. Finally, under diffusion assumption A1.1c, moments of δ_x^3 and higher are ignored. Substituting these approximations into Equation A1.8 gives

$$\varphi(x, t+\delta_t) = \varphi(x,t) - \frac{\partial\,\{[\, m(x)\delta_t + o(\delta_t)\,]\,\varphi(x,t)\}}{\partial x} + \frac{1}{2}\frac{\partial^2\,\{[\, v(x)\delta_t + o(\delta_t)\,]\,\varphi(x,t)\}}{\partial x^2} \quad \text{(A1.9)}$$

Rearranging, dividing the left side by δ_t, and taking the limit at $\delta_t \to 0$ gives

$$\lim_{\delta_t \to 0} \frac{\varphi(x, t+\delta_t) - \varphi(x,t)}{\delta_t} = \frac{\partial \varphi(x,t)}{\partial t}$$

Likewise, recalling that $\lim_{\delta_t \to 0} o(\delta_t)/\delta_t = 0$, the last two terms of Equation (A1.9) simplify to

$$\lim_{\delta_t \to 0} \frac{1}{\delta_t}\left(\frac{\partial\,\{[\, m(x)\delta_t + o(\delta_t)\,]\,\varphi(x,t)\}}{\partial x}\right) = \frac{\partial[\, m(x)\varphi(x,t)\,]}{\partial x}$$

$$\lim_{\delta_t \to 0} \frac{1}{\delta_t}\left(\frac{\partial^2\{[\, v(x)\delta_t + o(\delta_t)\,]\,\varphi(x,t)\}}{\partial x^2}\right) = \frac{\partial^2[\, v(x)\varphi(x,t)\,]}{\partial x^2}$$

Together, these simplifications imply

$$\frac{\partial\, \varphi(x,t)}{\partial\, t} = -\frac{\partial[\, m(x)\, \varphi(x,t)\,]}{\partial\, x} + \frac{\partial^2[\, v(x)\, \varphi(x,t)\,]}{2\, \partial\, x^2}$$

recovering the Kolmogorov forward equation (Equation A1.6), again recalling that $\varphi(x,t)$ is really $\varphi(x,t,p)$.

Stationary Distributions

At equilibrium, a probability density function does not change over time, i.e.,

$$\frac{\partial\, \varphi(x,t,p)}{\partial\, t} = 0 \quad \text{(A1.10a)}$$

Should such a distribution exist, it is called the **stationary distribution** and denoted by $\varphi(x)$. Stationary distributions are independent of the starting conditions: regardless of where the process starts in the interior of the diffusion, it converges to the same distribution. $\varphi(x,t,p)$ can thus be decomposed into a transient deviation (dependent on t and p) and a stationary expectation (independent of both p and t), $\varphi(x,t,p) = \varphi^*(x,t,p) + \varphi(x)$. The transient deviation satisfies $\lim_{t\to\infty} \varphi^*(x,t,p) = 0$, such that the deviation from equilibrium decays to zero over time.

When Equation A1.10a is satisfied, Equation A1.6 becomes

$$\frac{d\, m(x)\, \varphi(x)}{dx} = \frac{d^2\, v(x)\, \varphi(x)}{2\, dx^2} \tag{A1.10b}$$

and integration of both sides reduces this to the simple differential equation

$$\frac{d\, v(x)\, \varphi(x)}{dx} = 2m(x)\, \varphi(x) \tag{A1.10c}$$

which has solution

$$\varphi(x) = \frac{C}{v(x)\, G(x)} \tag{A1.11}$$

where G (called the **scale function** in the diffusion literature) is defined by the indefinite integral

$$G(x) = \exp\left[-2 \int^x \frac{m(y)}{v(y)} dy\right] \tag{A1.12}$$

and C is a constant such that $\varphi(x)$ integrates to one, making Equation A1.11 a proper probability density function. Wright (1938) first obtained Equation A1.10, using a somewhat different approach. Note that $\int [v(x)\, G(x)]^{-1}\, dx$ may be infinite, in which case no stationary distribution exists. This happens, for example, in the absence of mutation and migration where both boundaries are absorbing.

Example A1.3. Consider the case in which random genetic drift is the only force influencing allele frequency change. From Equation A1.3, $m(x) = 0$ and $v(x) = x(1-x)/(2N_e)$, giving

$$G(x) = \exp\left[-4N_e \int^x \frac{0}{y(1-y)}\, dy\right] = e^{-4N_e 0} = 1 \tag{A1.13}$$

and

$$\varphi(x) = \frac{2N_e C}{x(1-x)} \tag{A1.14}$$

The only valid equilibrium distribution is $\varphi(x) = 0$ (e.g., $C = 0$), as

$$\int_0^1 x^{-1}(1-x)^{-1} dx = \infty$$

As shown in Example A1.2, after sufficient time, without any force to keep allele frequencies in the interior, all alleles eventually reach frequency zero or one. As a result, the equilibrium distribution on the open interval $0 < x < 1$ is zero.

Example A1.4. Compute the stationary distribution for the frequency of an allele at a diallelic locus experiencing selection, mutation and drift. Using $m(x)$ and $v(x)$ from Equation A1.5,

$$\begin{aligned}\int^x \frac{m(y)}{v(y)}\,dy &= 2N_e \int^x \frac{y(1-y)d\ln(\overline{W})/(2\,dy) \;+\; (1-y)\mu \;-\; y\nu}{y(1-y)}\,dy \\ &= N_e \int^x \frac{d\ln(\overline{W})}{dy}\,dy \;+\; 2N_e\mu \int^x \frac{1}{y}\,dy \;-\; 2N_e\nu \int^x \frac{1}{1-y}\,dy \\ &= N_e \ln(\overline{W}) \;+\; 2N_e\mu \ln(x) \;+\; 2N_e\nu \ln(1-x)\end{aligned}$$

Hence,

$$G(x) = \exp\left[-2 \int^x \frac{m(y)}{v(y)}\,dy \right] = \overline{W}^{\,-2N_e}\, x^{-4N_e\mu}\,(1-x)^{-4N_e\nu}$$

and applying Equation A1.11 gives

$$\varphi(x) = C\overline{W}^{\,2N_e}\, x^{4N_e\mu-1}\,(1-x)^{4N_e\nu-1} \quad \text{for } 0 < x < 1, \tag{A1.15}$$

a result first due to Wright (1931).

Example A1.5. A particular application of Equation A1.15 is the case of a deleterious recessive allele maintained by mutation. In an infinite population (Chapter 7), the equilibrium frequency of this allele is approximately $\sqrt{\mu/s}$ (for $s \gg \mu$), where the fitness of the recessive is $1-s$ and the mutation rate from normal to recessive is μ. Here, $\overline{W} = 1 - sp^2$, and because $(1-x)^{N_e} \simeq e^{-xN_e}$,

$$\overline{W}^{\,2N_e} = (1-sp^2)^{\,2N_e} \simeq \exp(-2N_e sp^2)$$

giving the equilibrium distribution as

$$\varphi(x) = Ce^{-2N_e sx^2}\, x^{4N_e\mu-1}\,(1-x)^{4N_e\nu-1} \quad \text{for } 0 < x < 1$$

The Kolmogorov Backward Equation

While the KFE provides both the full solution and, where appropriate, the equilibrium solution, we can obtain much simpler expressions for many quantities of interest when an equilibrium solution does not exist. For example, if one or both boundaries are accessible and absorbing, we can compute the fixation probabilities (the probability that the process eventually reaches a specified boundary), the time to reach the boundary (the time to loss or fixation), and the expected value of many other functions of interest. The key to all of these operations is the **Kolmogorov backward equation**, or **KBE**.

The KBE (which can be derived in a manner similar to the KFE) is

$$\frac{\partial\,\varphi(x,t,p)}{\partial\,t} = m(p)\,\frac{\partial\,\varphi(x,t,p)}{\partial\,p} + \frac{1}{2}\,v(p)\,\frac{\partial^2\varphi(x,t,p)}{\partial\,p^2} \tag{A1.16}$$

Notice that the KBE starts at the current time t and looks *backwards* as to how changes in the initial starting value p influence the current position, hence the name.

DIFFUSION APPLICATIONS IN POPULATION GENETICS

When no stationary distribution exists, useful summary statistics for the diffusion are the probability of fixation of the various boundaries and the expected time to reach a specified boundary. We consider each in turn.

Probability of Fixation

When at least one boundary is absorbing (and accessible), no stationary distribution exists. In such cases, one important descriptor of the process is the probability of reaching one boundary before the other. One can show that the function $u(p,t)$, the probability of fixation by time t given we start at p, satisfies the KBE (we replace $\varphi(x,t,p)$ by $u(p,t)$ in Equation A1.16). We are typically interested in the ultimate probability of fixation $u(p) = \lim_{t\to\infty} u(p,t)$, in which case the partial derivative of $u(p)$ with respect to time is zero and $u(p)$ satisfies

$$0 = m(p)\,\frac{\partial\, u(p)}{\partial\, p} + \frac{1}{2}\,v(p)\,\frac{\partial^2 u(p)}{\partial\, p^2}$$

which has solution (Kimura 1962)

$$u(p) = \frac{\int_0^p G(x)\,dx}{\int_0^1 G(x)\,dx} \tag{A1.17a}$$

where $G(x)$ is defined by Equation A1.12. More generally, for any diffusion (regardless of the nature of the boundaries) the probability that the process reaches b before a, given it starts at p (where $A < a < p < b < B$ with the diffusion defined over $A < x < B$), is

$$u_{b,a}(p) = \frac{\int_a^p G(x)\,dx}{\int_a^b G(x)\,dx} \tag{A1.17b}$$

Example A1.6. Compute the probability of fixation of an allele under drift alone. From Equation A1.13, $G(x) = 1$ giving

$$u(p) = \frac{\int_0^p 1\,dx}{\int_0^1 1\,dx} = \frac{p-0}{1-0} = p$$

The fixation probability of a neutral allele is thus simply its starting allele frequency. Likewise, the probability, starting at $a < p < b$, that we reach b before a is just

$$u_{b,a}(p) = \frac{p-a}{b-a}$$

Now consider a diallelic locus experiencing additive selection and drift. From Equation A1.4,

$$m(x) = sx(1-x), \qquad v(x) = \frac{x(1-x)}{2N_e}$$

implying

$$G(x) = \exp\left[-4N_e s \int^x \frac{y(1-y)}{y(1-y)}dy\right] = e^{-4N_e sx}$$

Thus,

$$u(p) = \frac{\int_0^p e^{-4N_e sx}\,dx}{\int_0^1 e^{-4N_e sx}\,dx} = \frac{1 - e^{-4N_e sp}}{1 - e^{-4N_e s}} \tag{A1.18}$$

as obtained by Kimura (1957, 1962). A case of special interest is an initially rare allele ($p \ll 1$) under strong selection ($4N_e s \gg 1$). Here, the term in the denominator is essentially one, and since $1 - e^{-ax} = ax + o(x)$, we have

$$u(p) \simeq 4N_e sp \quad \text{for} \quad 4N_e s \gg 1 \quad \text{and} \quad p \ll 1$$

In particular, for a new mutation, $p = 1/(2N)$ and the fixation probablity is

$$u\left(\frac{1}{2N}\right) = 4N_e s \cdot \frac{1}{2N} = 2s\frac{N_e}{N}$$

More generally, starting at p, the probablity that we reach b before reaching a is

$$u_{a,b}(p) = \frac{\int_a^p e^{-4N_e sx}\,dx}{\int_a^b e^{-4N_e sx}\,dx} = \frac{e^{-4N_e sa} - e^{-4N_e sp}}{e^{-4N_e sa} - e^{-4N_e sb}}$$

Finally, allowing for dominance, we assign fitness as $1 \,:\, 1 + s(1+h) \,:\, 1 + 2s$. Under weak selection (ignoring terms of s^2 and higher),

$$m(x) = sx(1-x)[\,1 + h(1-2x)\,]$$

giving

$$\begin{aligned} G(x) &= \exp\left[-4N_e s \int^x \frac{y(1-y)[\,1 + h(1-2y)\,]}{y(1-y)}\,dy\right] \\ &= \exp\{-4N_e sx[\,1 + h(1-x)\,]\,\} \end{aligned}$$

and hence

$$u(p) = \frac{\int_0^p \exp\{-4N_e sx[\,1 + h(1-x)\,]\}\,dx}{\int_0^1 \exp\{-4N_e sx[\,1 + h(1-x)\,]\,\}\,dx} \tag{A1.19}$$

Example A1.7. An important special case of Equation A1.19 is a favored, but completely recessive, allele. In this case $h = -1$ and

$$m(x) = 2sx^2(1-x), \quad \text{and} \quad G(x) = \exp\left(-4N_e sx^2\right)$$

giving

$$u(p) = \frac{\int_0^p e^{-4N_e sx^2}\,dx}{\int_0^1 e^{-4N_e sx^2}\,dx}$$

We can express this in terms erf(x), the **error function** (Abramowitz and Stegun 1972), defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt$$

With the change of variables, $t = x\sqrt{4N_e s}$,

$$\int_0^p e^{-4N_e sx^2}\,dx = \frac{1}{\sqrt{4N_e s}} \int_0^{p\sqrt{4N_e s}} e^{-t^2}\,dt = \frac{1}{4}\sqrt{\frac{\pi}{N_e s}}\,\text{erf}\left(p\sqrt{4N_e s}\right)$$

Hence,

$$u(p) = \frac{\text{erf}\left(p\sqrt{4N_e s}\right)}{\text{erf}\left(\sqrt{4N_e s}\right)}$$

Consider strong selection ($4N_e s \gg 1$), so that $\text{erf}\left(\sqrt{4N_e s}\right) \simeq 1$ (as $\text{erf}(x) \to 1$ for $x \gg 1$), and hence $u(p) \simeq \text{erf}\left(p\sqrt{4N_e s}\right)$. We can further simplify this expression. For small x, Abramowitz and Stegun (1972) give the approximation

$$\text{erf}(x) \simeq \frac{2\,x}{\sqrt{\pi}}$$

When the recessive benefical allele is initally rare ($x = p\sqrt{4N_e s} \ll 1$), its fixation probability is

$$u(p) \simeq \frac{2}{\sqrt{\pi}}\, p\sqrt{4N_e s}$$

For the case of a newly arisen favorable recessive mutation, $p = 1/(2N)$, and

$$u\left(\frac{1}{2N}\right) \simeq \frac{1}{N}\sqrt{\frac{4N_e s}{\pi}} \tag{A1.20}$$

This result was obtained by Kimura (1957). For the case of $N = N_e$, this reduces to $\sqrt{4s/N\pi} \simeq 1.13\sqrt{s/N}$. Previously, Haldane (1927) used the theory of branching processes to obtain an approximation of $\sqrt{2s/N} \simeq 1.41\sqrt{s/N}$, while Wright (1942) obtained an approximation of $\sqrt{s/N}$. Note that our expression is slightly different from the common form seen in the literature, as we used the general fitness model given in Example A1.6 (the homozygote fitness is $1 + 2s$), while in models dealing with recessives only the fitness is typically assigned as $1 + s$.

Time to Fixation

The time for a process to reach a specified value (or values) is called the **sojourn time**. While we are often interested in the sojourn time to an absorbing boundary (the fixation time), the more general problem of the time to first reach a specific value within the interval over which the diffusion is defined is also tractable. Thus, it will prove very useful to express the expected time that a diffusion spends in the interval (a, b) as

$$\overline{t}_{a,b} = \int_0^\infty \Pr[a \le x_t \le b]\, dt = \int_0^\infty \int_a^b \varphi(x, t, p)\, dx\, dt \tag{A1.21}$$

If a stationary distribution exists, this time is infinite and not really of any concern. If one or both boundaries are absorbing (and accessible) then $\overline{t}(p)$, the total time the diffusion spends in the interior, is given by evaluating Equation A1.21 taking a and b as the lower and upper limits (respectively) of the diffusion.

A very useful approach to solving Equation A1.21, and more general problems, is to consider $h(x, p)$, the expected amount of time that the process (starting at p) spends in the neighborhood of x before it is eventually lost or fixed. Formally, this is given by

$$h(x, p) = \int_0^\infty \varphi(x, t, p)\, dt$$

Thus, Equation A1.21 can be expressed as

$$\overline{t}_{a,b} = \int_a^b h(x,p)dx \tag{A1.22a}$$

The function $h(x,p)$ is called a **Greens function** and will prove very useful in solving a variety of problems. The general solution for $h(x,p)$ is given in Maruyama (1977),

$$h(x,p) = \begin{cases} \dfrac{2[1-u(p)]}{v(x)G(x)} \displaystyle\int_a^x G(y)\,dy, & \text{for } a < x < p \\ \dfrac{2u(p)}{v(x)G(x)} \displaystyle\int_x^b G(y)\,dy, & \text{for } p < x < b \end{cases} \tag{A1.22b}$$

where the fixation probablity $u(p)$ is given by Equation A1.17a.

One can also obtain modified Greens functions for conditional processes, such as for those leading only to loss or those leading only to fixation. For example, $\overline{t}_F$, the expected time to fix allele A (in those populations where it is fixed) is given by replacing $h(x,p)$ by

$$h_1(x,p) = h(x,p)\,\frac{u(x)}{u(p)} \tag{A1.23a}$$

This follows from standard conditional probability arguments (see Ewens 1979, 2004), with $u(x)/u(p)$ correcting for the fact that we are only considering those sample paths over which A is fixed. Similarly, $\overline{t}_L$, the expected time to lose allele A is obtained by replacing $h(x,p)$ by

$$h_0(x,p) = h(x,p)\,\frac{1-u(x)}{1-u(p)}. \tag{A1.23b}$$

Finally, $\overline{t}$, $\overline{t}_F$, and $\overline{t}_L$ are related by

$$\overline{t}(p) = u(p)\,\overline{t}_F(p) + [1-u(p)]\,\overline{t}_L(p) \tag{A1.24}$$

That is, the expected time to loss or fixation is equal to the expected time to fixation multiplied by the probability of fixation plus expected time to loss multiplied by the probability of loss.

Example A1.8. Compute the conditional and unconditional expected time to loss or fixation for a neutral allele. Under neutrality, $u(x) = x$ and $G(x) = 1$, and hence, $\int_0^x G(y)dy = x$, $\int_x^1 G(y)dy = 1-x$, and Equation A1.22b simplifies considerably to

$$h(x,p) = \begin{cases} 4N_e(1-p)/(1-x) & \text{for } 0 < x < p \\ 4N_e p/x & \text{for } p < x < 1 \end{cases}$$

Thus, the expected amount of time that a neutral allele (with initial frequency p) remains polymorphic is

$$\begin{aligned} \overline{t}(p) &= \int_0^1 h(x,p)\,dx \\ &= 4N_e(1-p)\int_0^p \frac{dx}{1-x} + 4N_e p\int_p^1 \frac{dx}{x} \\ &= -4N_e[\,(1-p)\ln(1-p) + p\ln(p)\,] \end{aligned} \tag{A1.25}$$

Similarly, the conditional fixation times are obtained using Equations A1.23a,b

$$h_0(x,p) = \frac{1-x}{1-p}\,h(x,p) = \begin{cases} 4N_e & \text{for } 0 < x < p \\ \\ 4N_e \dfrac{p(1-x)}{x(1-p)} & \text{for } p < x < 1 \end{cases} \tag{A1.26a}$$

for those paths leading to loss, and

$$h_1(x,p) = \frac{x}{p}\,h(x,p) = \begin{cases} 4N_e \dfrac{x(1-p)}{p(1-x)} & \text{for } 0 < x < p \\ \\ 4N_e & \text{for } p < x < 1 \end{cases} \tag{A1.26b}$$

for those paths leading to fixation. Integration yields the expected conditional time to fixation as

$$\begin{aligned} \overline{t}_F(p) &= 4N_e \frac{1-p}{p} \int_0^p \frac{x}{1-x}\,dx \;+\; 4N_e \int_p^1 dx \\ &= -4N_e\,\frac{1-p}{p}\,\ln(1-p) \end{aligned} \tag{A1.27}$$

and the expected conditional time to loss as

$$\begin{aligned} \overline{t}_L(p) &= 4N_e \int_0^p dx \;+\; 4N_e \frac{p}{1-p} \int_p^1 \frac{1-x}{x}\,dx \\ &= -4N_e\,\frac{p}{1-p}\,\ln(p) \end{aligned} \tag{A1.28}$$

These results were first obtained by Kimura and Ohta (1969a,b). For the special case of a neutral allele introduced as a single copy, $p = 1/2N$, and Equations A1.27 and A1.28 reduce further to

$$\overline{t} \simeq \overline{t}_L \simeq \frac{2N_e}{N}\,\ln(2N), \qquad \overline{t}_F \simeq 4N_e$$

Lastly, one can show that Equation A1.24 is satisfied by noting

$$\begin{aligned} u(p)\overline{t}_F(p) + [1-u(p)]\overline{t}_L(p) &= -4N_e\left[\,p\left(\frac{1-p}{p}\ln(1-p)\right) + (1-p)\left(\frac{p}{1-p}\,\ln(p)\right)\right] \\ &= -4N_e[\,(1-p)\ln(1-p) + p\ln(p)\,] = \overline{t} \end{aligned}$$

Expectations of More General Functions

The sojourn times are examples of taking expected values along a sample path, and more complex functions can be evaluated in a similar manner. Let x_t, residing in the open interval $(0,1)$, denote the values of our random variable along a particular sample path, and suppose that we wish to compute the integral (over time) of some function g of x_t,

$$I_g(p) = \int_0^\infty g(x_t)\,dt$$

Note that x_t is the realization at a particular time point, and hence the distribution of x_t (and thus ultimately I_g) is a function of p, the starting value of our process. Because x_t is a random variable, so is the integral $I_g(p)$. Its expected value is given by

$$E[I_g(p)] = \int_0^\infty \int_0^1 g(x)\varphi(x,t,p)dt\,dx \tag{A1.29a}$$

Fortunately, we do not have to solve for $\varphi(x,t,p)$, as the general solution is given by using the Greens function described by Equation A1.22b,

$$E[I_g(p)] = \int_0^1 g(x)\,h(x,p)\,dx \tag{A1.29b}$$

This makes sense, as $h(x,p)dx$ is the expected amount of time that the process (starting at p) spends in the neighborhood of x before it is eventually lost or fixed. Thus, integration over all possible neighborhoods gives the expected value of the function over all sample paths.

Example A1.9. Again assuming an allele under drift only, consider two different functions, the total number of copies of the allele and total number of homozygotes involving this allele along those paths that lose the allele. These give rise to functions of $g(x) = 2Nx$ for the total number of copies of the allele during the course of its history and (assuming perfect Hardy-Weinberg) $g(x) = Nx^2$ for total number of homozygotes. To obtain the expected values of each, we use Equation A1.26a, which gives the Greens function for pure drift on paths that are ultimately lost,

$$\begin{aligned} E(g,p) &= \int_0^1 g(x)h_0(x,p)dx = 4N_e\int_0^p g(x)dx + 4N_e\frac{p}{1-p}\int_p^1 \frac{1-x}{x}g(x)dx \\ &= 4N_e\int_0^p g(x)dx + 4N_e\frac{p}{1-p}\left(\int_p^1 \frac{g(x)}{x}dx - \int_p^1 g(x)dx\right) \end{aligned}$$

First considering the total number of alleles $g(x) = 2Nx$, the above expression becomes

$$E(2Nx) = 8N_eN\left[\int_0^p x\,dx + \frac{p}{1-p}\left(\int_p^1 dx - \int_p^1 x\,dx\right)\right]$$

The integrals in the square brackets are easily solved to give

$$\frac{p^2}{2} + \frac{p}{1-p}\left((1-p) - \frac{1-p^2}{2}\right) = \frac{p}{2}$$

Hence, the expected total number of copies of the allele, starting at frequency p, during its entire existance is

$$8N_eNp/2 = 4N_eNp$$

Note that there are initially $2Np$ copies of the allele at the start of the process, so that the total number of copies over the entire life of the conditional process leading to loss is just $2N_e$ times the initial number. Hence, a single new mutation destined to become lost leaves an average of $2N_e$ copies over its lifetime. Of course, the distribution of the actual number of copies is heavily skewed, with a long right tail. Most mutations are quickly lost, but a few become very successful (and hence leave many copies) before eventually dying out.

Turning to the expected total number of homozygotes, $g(x) = Nx^2$, giving

$$E(Nx^2) = 4N_eN\left[\int_0^p x^2dx + \frac{p}{1-p}\left(\int_p^1 x\,dx - \int_p^1 x^2dx\right)\right]$$

$$\simeq 4N_eN(p/6) = (2/3)N_eNp \quad (\text{for } p \ll 1)$$

Again starting with a single copy, $E(Nx^2) \simeq N_e/3$.

MULTIVARIATE DIFFUSIONS: MULTIPLE ALLELES AND TWO LOCI

Our previous focus has been on univariate diffusions. However in many cases, such as a locus with multiple alleles or two diallelic loci (e.g., Chapter 5), more than one random variable is required to describe the system. A multivariate diffusion process involving n variables needs to consider $n(n-1)/2$ covariances in addition to the n infinitesimal means and the n infinitesimal variances. The multivariate diffusion assumption is that means, variances, and covariances are still of the form $f(x)\delta_t + o(\delta_t)$. While it is fairly straightforward to write down the KFE or KBE for common multivariate diffusions that arise in population genetics, their analysis is far from trivial. Thus, in contrast to the rather rich theory and results for univariate diffusions, very little work has been done with multivariate diffusions.

As an example of such a diffusion, consider k alleles at a single locus. Here $\varphi(\mathbf{x}, \mathbf{p}, t)$ is the density function with $\mathbf{x}$ denoting the vector of allele frequencies and $\mathbf{p}$ the vector of their starting values. Under pure drift, the KFE becomes

$$\frac{\partial\,\varphi(\mathbf{x},\mathbf{p},t)}{\partial\,t} = \frac{1}{2}\sum_{i=1}^{k-1}\frac{\partial^2}{\partial\,x_i^2}\left[\varphi(\mathbf{x},\mathbf{p},t)\,\frac{x_i(1-x_i)}{N_e}\right] - \sum_{i<j}\frac{\partial^2}{\partial\,x_i\,\partial\,x_j}\left[\varphi(\mathbf{x},\mathbf{p},t)\,\frac{x_i\,x_j}{N_e}\right] \quad \text{(A1.30)}$$

where the first sum involves the allele-frequency variances, and the second involves the covariances between allele frequencies ($\text{Cov}(x_i, x_j) = -x_i\,x_j$ for a multinomial distribution). If there are directional forces, then the term

$$-\sum_{i=1}^{k-1}\frac{\partial\,m(\mathbf{x})\,\varphi(\mathbf{x},\mathbf{p},t)}{\partial\,x_i}$$

appears on the right hand side of Equation A1.30. The KBE for this process is similarly defined.

Now consider two diallelic loci under strict drift. In this case, the additional force of recombination needs to be considered. Here the dynamics of the system are described by three variables, the frequencies x and y of the alleles at the two loci and the disequilibrium D between them (Chapter 2). Thus the diffusion needs to follow the mean change in D (under drift, the allele frequencies, on average, remain unchanged but D can change due to recombination), the variances of x, y, and D, and the appropriate covariances between these. See Ohta and Kimura (1969) for details.

APPLICATIONS IN QUANTITATIVE GENETICS

While we have focused on population-genetic applications of diffusion processes, this approach is also very useful for solving a number of problems in quantitative genetics and, we

conclude by considering a few important examples. When attention shifts from individual alleles to a quantitative character, diffusions typically follow mean phenotypes instead of allele frequencies. Two well-studied diffusions, **Brownian motion** and the **Ornstein-Uhlenbeck process**, are especially useful.

Brownian Motion Models

For Brownian motion (also called the **Wiener process**), the diffusion over $-\infty < x < \infty$ is given by

$$m(x) = a \qquad v(x) = b \tag{A1.31a}$$

where $b > 0$. The general diffusion solution under Brownian motion starting at x_0 is that the distribution of x_t is normal, with mean $x_0 + at$ and variance $\sigma_t^2 = bt$,

$$x_t \sim \mathrm{N}(x_0 + at, bt) \tag{A1.31a}$$

There is no equilibrium solution, as the process converges to a normal with infinite variance (and infinite mean if $a \neq 0$).

Example A1.10. Lande (1976) used the Brownian-motion model to approximate the change in the phenotypic mean of a neutral character with constant additive genetic variance. In this case there is no directional force to change the mean, so $a = 0$. Assuming the character is strictly additive, the per generation sampling variance in the mean is σ_A^2/N_e (Chapter 12), which is used for b. Hence, at generation t, the distribution of phenotypic means is approximately normal with expected mean μ_0 (the initial mean) and variance $\sigma_t^2 = t\sigma_A^2/N_e$. One measure of how quickly phenotypic means drift is given by the minimum number of generations required for a random population to have at least a 50% probability of being more that K standard deviations from its initial mean. This is expressed as $\Pr(|x_t - \mu_0| \geq K\sigma_z) = 0.5$, where x_t is the mean of a randomly-drawn replicate population and σ_z^2 the phenotypic variance. Assuming Brownian motion, $(x_t - \mu_0)/\sigma_t$ is a unit normal random variable, hence

$$\Pr(|x_t - \mu_0| \geq K\sigma_z) = \Pr\left[\frac{|x_t - \mu_0|}{\sigma_t} \geq \frac{K\sigma_z}{\sigma_t}\right] = \Pr\left[|U| \geq \frac{K\sigma_z}{\sigma_t}\right] = 0.5$$

For a unit normal U, $\Pr(|U| \geq 0.675) = 0.5$, giving $K\sigma_z/\sigma_t = K\sigma_z/(\sigma_A\sqrt{t/N_e}) = 0.675$. Upon rearranging and substituting $h^2 = \sigma_A^2/\sigma_z^2$,

$$t = \frac{K^2\, N_e}{h^2\, 0.675^2} \simeq 2\, N_e\, \frac{K^2}{h^2}$$

Thus, for $N_e = 10$, a neutral character with heritability $h^2 = 0.5$ requires $2 \times 10 \times 9/0.5 = 360$ generations until half the populations have phenotypic means more than three standard deviations ($K = 3$) from their initial value. For $N_e = 10^5$, this time is 3.6 million generations!

The above analysis considers a single population over time (such as in a fossil sequence). Another common situation is where an initial population splits into two distinct isolated populations. Here, the divergence between the two means is the sum of the two individual variances, so that if the initial difference is 0, and d denotes the difference in mean,

$$d_t \sim \mathrm{N}\left[0, t \cdot \left(\frac{\sigma_A^2(1)}{N_e(1)} + \frac{\sigma_A^2(2)}{N_e(2)}\right)\right]$$

where $\sigma_A^2(i)$ and $N_e(i)$ are the additive genetic variance and effective population size for population i.

Example A1.11. The careful reader may recall that drift also changes σ_A^2, with the assumption of a constant σ_A^2 being reasonable only for $t < N$ (Chapter 11). In Example A1.10, we fixed the genetic variance but allowed N to vary. An alternative is to assume that the population has been at its current size sufficiently long enough that additive the variance is at its mutation-drift equilibrium value $\widehat{\sigma_A^2} = 2N_e\sigma_m^2$ (Equation 11.19b). The distribution of means now has expected variance

$$\sigma_t^2 = 2tN_e\sigma_m^2/N_e = 2t\sigma_m^2$$

and thus the expected number of generations until 50% of the means exceed K standard deviations is obtained from $K\sigma_z/\sqrt{t\,2\sigma_m^2} = 0.675$, or

$$t = \frac{K^2\sigma_z^2}{2 \cdot 0.675^2 \cdot \sigma_m^2} = 1.1\,\frac{K^2\sigma_z^2}{\sigma_m^2} \simeq \frac{K^2\sigma_z^2}{\sigma_m^2}$$

Because $\sigma_z^2 = \sigma_A^2 + \sigma_E^2 = 2N_e\sigma_m^2 + \sigma_E^2$,

$$t \simeq K^2\frac{\sigma_z^2}{\sigma_m^2} = K^2(2N_e + 1/h_m^2)$$

where $h_m^2 = \sigma_m^2/\sigma_E^2$ is the mutational heritability. From LW Table 9.1, h_m^2 has an approximate average value of 0.006. Taking this value, and repeating the calculations from Example A1.10 (e.g., $N_e = 10$ and $K = 3$) gives $t = 9 \times (20 + 1/0.006) = 1680$ generations. The reason for the huge increase in time (1680 vs. 360) relative to the fixed variance example above is that additive variance is much smaller due to the small population size. However, when $N_e = 10^5$, the time is 1.8 million generations, half the value obtained in Example A1.10.

Example A1.12. The Brownian motion model easily extends to multiple traits. Here, $\mathbf{x}$ is a vector of traits, and we assume an infinitesimal mean vector of zero and denote the infinitesimal covariance matrix by $\mathbf{V}$. If the process starts at $\mathbf{x}_0$, then

$$\mathbf{x}_t \sim \text{MVN}(\mathbf{x}_0, t\mathbf{V}) \tag{A1.32a}$$

where MVN denotes the multivariate normal. Lande (1979) examined the change in a vector of trait means under drift and mutation. The per-generation sampling variance in the mean is $\mathbf{G}/N_e$, the genetic covariance matrix divided by the effective population size. If $\mathbf{u}_0$ is the starting vector of means, then

$$\mathbf{u}_t \sim \text{MVN}(\mathbf{u}_0, (t/N_e)\mathbf{G}) \tag{A1.32b}$$

A key feature of Equation A1.32b is that we expect the most divergence along the major axes of variation of $\mathbf{G}$. Thus, if drift dominates, we expect to see the divergence between two populations to largely be in the direction of the first few principal components (eigenvalues) of $\mathbf{G}$.

Lande (1979) assumed that over long periods of time, the genetic covariance matrix would approach its mutation-drift equilibrium value. Let $\mathbf{U}$ denote the matrix of mutational variances and covariances, so that $\mathbf{U}_{ii}$ corresponds to the mutational (additive-genetic) variation in trait i generated each generation (σ_m^2 in the univariate case), while $\mathbf{U}_{ij}$ is the mutational additive-genetic covariance between traits i and j. For effective population size N_e, drift decays the

current additive-genetic variance by an amount $1/(2N_e)$ each generation, but this is countered by new mutation, giving the expected change in the average covariance matrix as

$$\Delta\mathbf{G} = \mathbf{G}/(2N_e) + \mathbf{U}$$

which implies an equilibrium value of

$$\widehat{\mathbf{G}} = 2N_e\mathbf{U}$$

Hence, the sampling variance of the mean becomes $\widehat{\mathbf{G}}/N_e = 2\mathbf{U}$, reducing Equation A1.32b to

$$\mathbf{u}_t \sim \text{MVN}(\mathbf{u}_0, 2t\mathbf{U}) \tag{A1.32c}$$

Here the neutral divergence between two populations follows the leading eigenvalues of $\mathbf{U}$, rather than $\mathbf{G}$.

Ornstein-Uhlenbeck Models

The Ornstein-Uhlenbeck process is an extension of the Brownian motion model to include a linear restoring force back to the origin. The diffusion process for $-\infty < x < \infty$ is given by

$$m(x) = -ax \qquad v(x) = b \tag{A1.33a}$$

with $a, b > 0$. As with Brownian motion, the distribution of x_t (given the starting condition x_0) is also normal, with mean and variance

$$\mu_t = x_0 e^{-at} \qquad \sigma_t^2 = \frac{b}{2a}\left(1 - e^{-2at}\right) \tag{A1.33b}$$

See Karlin and Taylor (1981) for a derivation. The resulting stationary distribution is normal with mean zero and variance $b/(2a)$.

Example A1.13. Lande (1976) examined the distribution of phenotypic means under drift and stabilizing selection, using the Gaussian fitness function (**nor-optimal selection**) as a model of stabilizing selection,

$$W(z) = C\, e^{-z^2/(2\omega)}$$

where the optimal phenotype is $z = 0$ and the strength of selection is given by ω, with smaller values of ω corresponding to stronger selection. Under nor-optimal selection, if phenotypes before selection are normally distributed with mean μ_t and phenotypic variance σ_z^2, they remain normal after selection, with new mean $\mu_t + S$, where

$$S = -\mu_t \frac{\sigma_z^2}{\sigma_z^2 + \omega}$$

and variance

$$\sigma_z^2 - \frac{\sigma_z^4}{\sigma_z^2 + \omega}$$

We will assume sufficiently weak selection ($\omega \gg \sigma_z^4$) so that the variance remains effectively unchanged after selection. As expected, selection moves the mean in the direction of the optimum (decreasing the current mean if it is positive, while increasing it if negative). Let

x_t be the mean in generation t of a randomly-drawn replicate population. Assuming the breeder's equation, $R = h^2 S$, defines the response to selection, the distribution of means can be approximated by an Ornstein-Uhlenbeck process, with

$$a = h^2 \, \frac{\sigma_z^2}{\sigma_z^2 + \omega} = \frac{\sigma_A^2}{\sigma_z^2 + \omega}, \qquad b = \frac{\sigma_A^2}{N_e}$$

where a follows from the change in mean $\Delta\mu = h^2 S$ and using the above value for S under nor-optimal selection. Hence, the distribution of phenotypic means in generation t is normal, with mean

$$\mu_t = \mu_0 \, \exp\left(-\, t \, \frac{\sigma_A^2}{\sigma_z^2 + \omega} \right) \tag{A1.34a}$$

and variance

$$\sigma_t^2 = \frac{\sigma_z^2 + \omega}{2N_e} \left[1 - \exp\left(-\, 2t \, \frac{\sigma_A^2}{\sigma_z^2 + \omega} \right) \right] \tag{A1.34b}$$

At equilibrium, the distribution of the means of replicate populations is normal with mean zero (centered at the fitness optimum) and variance $(\sigma_z^2 + \omega)/(2N_e)$.

Stationary Distributions for Mean Phenotype

Example A1.13 shows how diffusion approximations can give the stationary distribution under the joint actions of drift and stabilizing selection. More general forms of selection can be consider as well, as the following example illustrates.

Example A1.14. Suppose our concern is simply the change in the mean of a trait under arbitrary selection. Provided phenotypes are normally distributed, then from Equation 13.27a the change in mean is

$$\Delta\mu = \sigma_A^2 \frac{d\,[\,\ln(\overline{W}(\mu))\,]}{d\mu}$$

Here the mean population fitness $\overline{W}(\mu)$ is a function of the population mean μ. Since we made no assumptions about the nature of selection, any function for mean fitness can be used, *provided* the distribution of offspring phenotypes remains roughly normal. Hence, an approximating diffusion for the behavior of the mean under selection and drift is to let x = the current mean and set

$$m(x) = \sigma_A^2 \frac{d\ln(\overline{W})}{dx} \quad \text{and} \quad v(x) = \frac{\sigma_A^2}{N_e} \tag{A1.35}$$

Substituting Equation A1.35 into Equation A1.12 gives

$$G(x) = \exp\left[-2N_e \int^x \frac{d\ln(\overline{W})}{dy} dy \right] = \exp[-2N_e \ln(\overline{W})] = \overline{W}(x)^{-2N_e}$$

Thus, Equation A1.11 gives the equilibrium distribution as

$$\varphi(x) = \frac{C}{v(x)G(x)} = \frac{CN_e}{\sigma_A^2} \, \overline{W}(x)^{2N_e} \propto \overline{W}(x)^{2N_e} \tag{A1.36}$$

This result, due to Lande (1976), shows that as the effective population size increases, the probability of the population mean being near a local maximum in fitness also increases. This

follows since $\varphi(x)$ becomes increasingly peaked around local maxima relative to other parts of the fitness surface as we increase N_e. Equations A1.35-36 rely on the assumptions that the phenotypic and additive genetic variances remain constant as the mean changes, and (as mentioned above) that the phenotypic distribution remains normal.

Literature Cited

Abramowitz. M., and I. A. Stegun (Eds). 1972. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover, New York. [A1]

Crow, J. F., and M. Kimura. 1970. *An introduction to population genetics theory.* Harper & Row, New York. [A1]

Ewens, W. J. 1979. *Mathematical population genetics.* Springer, Berlin. [A1]

Ewens, W. J. 2004. *Mathematical population genetics. 1. Theoretical introduction*, 2nd Edition. Springer, Berlin. [A1]

Fisher, R. A. 1922. On the dominance ratio. *Proc. Royal Society Edinburgh* 42: 321–341. [A1]

Gale, J. S. 1990. *Theoretical population genetics.* Unwin Hyman, London. [A1]

Haldane, J. B. S. 1927. A mathematical theory of natural and artificial selection. Part V. Selection and mutation. *Proc. Cambridge Philos. Soc.* 23: 838–844. [A1]

Karlin, S., and H. M. Taylor. 1981. *A second course in stochastic processes*. Academic Press, New York. [A1]

Kimura, M. 1955a. Stochastic process and distribution of gene frequencies under natural selection. *Cold Spring Harbor Symp. on Quant. Biol.* 20: 33–53. [A1]

Kimura, M. 1955b. Solution of a process of random genetic drift with a continuous model. *Proc. Natl. Acad. Sci. USA* 41: 144–150. [A1]

Kimura, M. 1957. Some problems of stochastic processes in genetics. *Annals of Mathematical Statistics* 28: 882–901. [A1]

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47: 713 – 719. [A1]

Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge Univ. Press, Cambridge, UK. [A1]

Kimura, M., and T. Ohta. 1969a. The average number of generations until fixation of an individual mutant gene in a finite population. *Genetics* 61: 763–771. [A1]

Kimura, M., and T. Ohta. 1969b. The average number of generations until extinction of an individual mutant gene in a finite population. *Genetics* 63: 701–709. [A1]

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334. [A1]

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain-body allometry. *Evolution* 33: 402–416. [A1]

Maruyama, T. 1977. *Stochastic problems in population genetics*. Springer-Verlag, Berlin. [A1]

Ohta, T. and M. Kimura. 1969. Linkage disequilibrium due to random genetic drift. *Genetics* 63: 299–238. [A1]

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159. [A1]

Wright, S. 1938. The distribution of gene frequencies under irreversible mutation. *Proc. Natl. Acad. Sci.* 24: 253–259.[A1]

Wright, S. 1942. Statistical genetics and evolution. *Bull. Amer. Math. Soc.* 48: 223–246. [A1]

Wright, S. 1945. The differential equation of the distribution of gene frequencies. *Proc. Natl. Acad. Sci. USA* 31: 382-389. [A1]