Appendix 5

Diffusion Theory

Draft Version 1 December 2000, ©Dec 2000. B. Walsh and M. Lynch

Please e-mail any comments to: **jbwalsh@u.arizona.edu**

Exact solutions of the dynamics of most of the random processes arising in population and quantitative genetics are either unknown or are extremely cumbersome. Starting with Fisher (1922), the use of *diffusion approximations* to model the exact dynamics has proven to be extremely powerful. Useful introductions to diffusion theory with special reference to genetics are given by Ewens (1979) and Karlin and Taylor (1981). Specific applications to population genetics are covered by these authors as well as by Crow and Kimura (1970), Maruyama (1977), Kimura (1983), and Gale (1990). We start by reviewing, without being overly technical, the idea behind obtaining a diffusion approximation. We then present analytic expressions for time to fixation, the distribution at equilibrium, and time to loss or fixation. While most of our initial emphasis is on using diffusion theory to approximate the frequency of an allele at a diallelic locus, we conclude by examining applications to quantitative characters.

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

$$E\left(\delta_x \mid x_t = x\right) = m(x)\delta_t + o(\delta_t)$$

$$\sigma^2(\delta_x \mid x_t = x) = v(x)\delta_t + o(\delta_t)$$

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

then x_t is said to be a *diffusion process* (provided the additional technical restriction that x_t is a Markov process is satisfied). $o(\delta_t)$ means that the error term is small relative to δ_t (formally, $\lim_{\delta_t \to 0} o(\delta_t)/\delta_t = 0$), while $m(x)$ and $v(x)$ correspond to the mean and variance of the process over a very small time interval. The entire structure of a diffusion is described by $m(x)$ and $v(x)$, allowing analytic expressions for quantities of interest to be obtained.

Diffusion processes are of special importance in that a discrete space, discrete time, random variable can often be closely approximated by a diffusion. This is done by appropriately scaling of space and time to construct a new 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. Constructing a new random variable $x_{\tau_N}^{(N)} = X_{(\tau_N)}/(2N)$, where $\tau_N = t/N$ — a single generation increments the new time scale by $1/N$. Taking the limit as N approaches infinity, the limiting process x_τ is a continuous space, continuous time process that represents the allele frequency at time τ (e.g., $\lim_{N\to\infty} x_{\tau_N} = p_\tau$, the frequency at time τ). If scaling is correctly done, x_τ is now a diffusion process and provides a good approximation to the behavior of X_t. One immediate consequence of this scaling argument is that *the process proceeds at a rate proportional to* N. Hence, the larger N, the slower the rate.

The Kolmogorov Forward Equation

The *infinitesimal mean*, $m(x)$, and the *infinitesimal variance*, $v(x)$, are formally defined by

$$m(x) = \lim_{\delta\to 0} \frac{E\left(x_{t+\delta} - x_t \,|\, x_t = x\right)}{\delta} \tag{A5.1a}$$

$$v(x) = \lim_{\delta\to 0} \frac{E[\,(x_{t+\delta} - x_t)^2 \,|\, x_t = x\,]}{\delta} \tag{A5.1b}$$

Example 1. A diffusion process examining changes in allele frequencies in a diploid is 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 21) and using the deterministic change in allele frequency for $m(x)$. An diffusion process approximating pure drift can be constructed by use of a suitable time scale (see Karlin and Taylor 1981 for details) to give

$$m(x) = 0, \qquad v(x) = \frac{x(1-x)}{2N_e} \tag{A5.2}$$

where $x = \text{freq}(\mathbf{A})$. Similarly, consider additive selection when $|s|$ is small, so that $\overline{W} \simeq 1$. For Example 1 in Chapter 24, $\Delta p \simeq sp(1-p)$, suggesting as an approximating diffusion over $0 < x < 1$

$$m(x) = sx(1-x), \qquad v(x) = \frac{x(1-x)}{2N_e} \tag{A5.3}$$

Finally, consider arbitrary selection with constant fitnesses (with selection sufficiently weak such that $\overline{W} \simeq 1$) and forward and back mutation. Let ν be mutation

rate from $\mathbf{A}$ to $\mathbf{a}$, μ the mutation rate from $\mathbf{a}$ to $\mathbf{A}$, and $p = \text{freq}(\mathbf{A})$. Applying Wright's formula (modified to include mutation)

$$m(x) = \frac{x(1-x)}{2}\frac{d\ln(\overline{W})}{dx} + (1-x)\mu - x\nu$$

and again

$$v(x) = \frac{x(1-x)}{2N_e} \tag{A5.4}$$

Given $m(x)$, $v(x)$, and initial frequency p_0, the probability density function for x_t satisfies the *Kolmogorov forward equation*:

$$\frac{\partial\,\varphi(x,t,p_0)}{\partial\,t} = \frac{1}{2}\frac{\partial^2\,v(x)\,\varphi(x,t,p_0)}{\partial\,x^2} - \frac{\partial\,m(x)\,\varphi(x,t,p_0)}{\partial\,x} \tag{A5.5}$$

where $\varphi(x,t,p_0)$ is the probability density for x at time t given the process starts at p_0, such that

$$\Pr[a \le x_t \le b] = \int_a^b \varphi(x,t,p_0)\,dx$$

Solving this partial differential equation gives $\varphi(x,t,p_0)$. For exact solutions of φ for a number of population genetic problems, see Crow and Kimura (1970) and Maruyama (1977). For example, the curves given in Figure 4, Chapter 21 represent $\varphi(x,t,p_0)$ for the pure drift diffusion given by Equation A5.2. These curves (each representing $\varphi(x,t,p_0)$ for a particular time) were obtained by solving

$$\frac{\partial\,\varphi(x,t,p_0)}{\partial\,t} = \frac{1}{4N_e}\frac{\partial^2\,x(1-x)\,\varphi(x,t,p_0)}{\partial\,x^2}$$

subject to the initial conditions of either $p_0 = 0.5$ (left half of Figure 4) or $p_0 = 0.1$ (right half of Figure 4).

A final point about the density $\varphi(x,t,p_0)$ concerns its range of validity. Suppose the diffusion occurs over the interval (a,b). It is important to realize that $\varphi(x,t,p_0)$ is valid only for $a < x < b$. The diffusion approximates only what happens inside the boundaries, the behavior exactly at the boundaries being beyond the realm of the approximation. In many cases, x_t does not change value once it reaches a boundary. Such a boundary is called an *absorbing boundary*. For example, in the absence of mutation and migration, once an allele frequency reaches either 0 or 1, it stays there. For this case, both 0 and 1 are absorbing boundaries.

Stationary Distributions

At equilibrium, the probability density function does not change over time, e.g.,

$$\frac{\partial \varphi(x,t,p_0)}{\partial t} = 0 \tag{A5.6}$$

Such a distribution (if it exists) is called the *stationary distribution* and is denoted by $\varphi(x)$. The stationary distribution is independent of the starting conditions: regardless of where the process starts in the interior of the diffusion, it converges to the same distribution. $\varphi(x,t,p_0)$ can thus be decomposed into a transient and a stationary part, $\varphi(x,t,p_0) = \varphi^*(x,t,p_0) + \varphi(x)$, where the transient part satisfies $\lim_{t\to\infty} \varphi^*(x,t,p_0) = 0$. Hence, that part of the distribution depending on the initial starting conditions decays away over time leaving only the stationary distribution.

Solving Equation A5.5 subject to Equation A5.6 gives

$$\varphi(x) = \frac{C}{v(x)\,G(x)} \tag{A5.7}$$

where G is defined by the indefinite integral

$$G(x) = \exp\left[-2\int^x \frac{m(y)}{v(y)}dy\right] \tag{A5.8}$$

C is a constant such that $\varphi(x)$ integrates to one, so Equation 30A.7 is a proper probability density function. 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 2. Consider pure drift. From Equation A5.2, $m(x) = 0$ and $v(x) = x(1-x)/(2N_e)$, giving

$$G(x) = \exp\left[-4N_e\int^x \frac{0}{y(1-y)}\,dy\right] = e^0 = 1 \tag{A5.9}$$

and

$$\varphi(x) = \frac{2N_eC}{x(1-x)} \tag{A5.10}$$

The only valid equilibrium distribution is $\varphi(x) = 0$ (e.g., $C = 0$), as $\int_1^0 x^{-1}(1-x)^{-1}dx = \infty$. This makes sense, as after sufficient time all populations are either at $x = 1$ or $x = 0$, with no populations showing $0 < x < 1$ (see Figure 4, Chapter 21). The resulting equilibrium distribution on this interval is zero.

Example 3. Compute the stationary distribution for the frequency of an allele at a diallelic locus experiencing selection, mutation and drift. From Equation A5.4,

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

Hence,

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

Applying Equation A5.7 gives

$$\varphi(x) = C\overline{W}^{\,2N_e}\, x^{4N_e\mu-1}\,(1-x)^{4N_e\nu-1} \quad \text{for } 0<x<1 \qquad \text{(A5.11)}$$

a result first due to Wright (1931). Using interactive mathematical graphics packages (such as Mathematica) is a very fruitful way to explore the behavior of Equation A5.11, as well as other stationary distributions.

Probability of Fixation

When at least one boundary is absorbing, no stationary distribution exists. In such cases, one important descriptor of the process is the probability of reaching one boundary before the other. A companion equation to the forward equation (the Kolmogorov backward equation, or KBE) can be solved to obtain an expression for $u(p_0)$, the probability of fixation of an allele, given initial frequency p_0:

$$u(p_0) = \frac{\int_0^{p_0} G(x)\,dx}{\int_0^1 G(x)\,dx} \qquad \text{(A5.12a)}$$

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_0 (where $A \le a \le p_0 \le b \le B$ with the diffusion defined over $A < x < B$), is

$$u_{b,a}(p_0) = \frac{\int_a^{p_0} G(x)\,dx}{\int_a^b G(x)\,dx} \qquad \text{(A5.12b)}$$

Example 4. Compute the probability of fixation of an allele at a diallelic locus experiencing additive selection and drift. From Equation A5.2,

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

implying

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

Thus,

$$u(p_0) = \frac{1 - e^{-4N_e sp_0}}{1 - e^{-4N_e s}}$$

Likewise, for pure drift, $G(x) = 1$ (Equation A5.9), giving

$$u(p_0) = \frac{\int_0^{p_0} 1\, dx}{\int_0^1 1\, dx} = p_0$$

Finally, allowing for dominance, from Equation 29.4a,

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

giving

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

and hence

$$u(p_0) = \frac{\int_0^{p_0} \exp\left[-4N_e sx(1+h(1-x)\right]\, dx}{\int_0^1 \exp\left[-4N_e sx(1+h(1-x)\right]\, dx}$$

Time to Fixation

The expected time a diffusion spends in the interval (a, b) is given by

$$\overline{t}_{a,b} = \int_0^\infty \Pr[a \le x_t \le b]\, dt = \int_0^\infty \int_a^b \varphi(x, t, p_0)\, dx\, dt \tag{A5.13}$$

If a stationary distribution exists, this time is infinite and not really of any concern. If one or both boundaries are absorbing then $\overline{t}(p_0)$, the total time the diffusion spends in the interior, is given by evaluating Equation A5.13 taking a and b as the lower and upper limits (respectively) of the diffusion. Using the KBE, we can avoid the problem of first having to solve for $\varphi(x, t, p_0)$. Instead the integral in Equation A5.13 can be expressed as

$$\overline{t}_{a,b} = \int_a^b h(x, p_0) dx \tag{A5.14a}$$

where

$$h(x, p_0) = \begin{cases} \dfrac{2[1 - u(p_0)]}{v(x)G(x)} \displaystyle\int_a^x G(y)\, dy, & \text{for} a \le x \le p_0 \\ \dfrac{2u(p_0)}{v(x)G(x)} \displaystyle\int_x^b G(y)\, dy, & \text{for} p_0 \le x \le b \end{cases} \tag{A5.14b}$$

This can be modified to obtain expressions for conditional times. For example, $\overline{t}_F$, the expected time to fix allele **A** (in those populations where it is fixed) is given by replacing $h(x, p_0)$ by

$$h_F(x, p_0) = h(x, p_0)\, \frac{u(x)}{u(p_0)} \tag{A5.15a}$$

This follows from standard conditional probability arguments (see Ewens 1979) and $u(x)/u(p_0)$ is a weighting factor 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 $h(x, p_0)$ by

$$h_L(x, p_0) = h(x, p_0)\, \frac{1 - u(x)}{1 - u(p_0)} \tag{A5.15 b}$$

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

$$\overline{t}(p_0) = u(p_0)\, \overline{t}_F(p_0) + [1 - u(p_0)]\, \overline{t}_L(p_0)$$

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 5. Compute the conditional and unconditional expected time to loss or fixation for a neutral allele. For a neutral allele, $u(x) = x$ and $G(x) = 1$.

Hence, $\int_0^x G(y)dy = x$, $\int_x^1 G(y)dy = 1 - x$, and Equation A5.14b simplifies considerably to

$$h(x, p_0) = \begin{cases} 4N_e(1-p_0)/(1-x), & \text{for } 0 \le x \le p_0 \\ \\ 4N_e p_0/x & \text{for } p_0 \le x \le 1 \end{cases}$$

Thus the expected amount of time a neutral allele (with initial frequency p_0) remains polymorphic is

$$\begin{aligned} \overline{t}(p_0) &= \int_0^1 h(x, p_0)\, dx \\ &= 4N_e(1-p_0) \int_0^{p_0} \frac{dx}{1-x} + 4N_e p_0 \int_{p_0}^1 \frac{dx}{x} \\ &= -4N_e[\,(1-p_0)\ln(1-p_0) + p_0 \ln(p_0)\,] \end{aligned}$$

Similarly, the conditional fixation times are obtained using

$$h_0(x, p_0) = \frac{1-x}{1-p_0}\, h(x, p_0) \qquad \text{and} \qquad h_1(x, p_0) = \frac{x}{p_0}\, h(x, p_0)$$

giving the expected conditional time to fixation as

$$\begin{aligned} \overline{t}_F(p_0) &= 4N_e \frac{1-p_0}{p_0} \int_0^{p_0} \frac{x}{1-x}\, dx \, + \, 4N_e \int_{p_0}^1 dx \\ &= -4N_e \, \frac{1-p_0}{p_0} \, \ln(1-p_0) \end{aligned}$$

and the expected conditional time to loss as

$$\begin{aligned} \overline{t}_L(p_0) &= 4N_e \int_0^{p_0} dx \, + \, 4N_e \frac{p_0}{1-p_0} \int_{p_0}^1 \frac{1-x}{x}\, dx \\ &= -4N_e \, \frac{p_0}{1-p_0} \, \ln(p_0) \end{aligned}$$

These results were first obtained by Kimura and Ohta (1969a ,b). For the special case of a neutral allele introduced as a single copy, $p_0 = 1/2N$ and these reduce further to

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

Green's Functions: Exceptations of More General Expressions.

Material still to be written

Applications to Quantitative Characters

When we shift our attention to from individual alleles to the phenotype of a quantitative character under drift, diffusions now follow the mean phenotype instead of the frequency of an allele. Two well studied diffusions, *Brownian motion* and the *Ornstein-Uhlenbeck process*, are especially useful in this case. For Brownian motion, the diffusion over $-\infty < x < \infty$ is given by

$$m(x) = a \qquad v(x) = b \tag{A5.16}$$

where $b > 0$. The general solution under Brownian motion starting at x_0 is that x_t is normally distributed, with mean $x_0 + at$ and variance $\sigma_t^2 = bt$. There is no equilibrium solution, as the process converges to a normal with infinite variance (and infinite mean if $a \neq 0$).

Example 5. Lande (1976) used Brownian motion to approximate the change in the phenotypic mean for a neutral character with constant additive genetic variance. 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 23), 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 such that a random population has 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 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 from their initial value.

The careful reader will recall from Chapter 23 that drift also changes σ_A^2, with the assumption of a constant σ_A^2 being reasonable only for $t < N$. Alternatively, we could assume that the population has been at its current size sufficiently long enough so that additive variance is at its mutation-drift equilibrium value (Equation 23.19b), $\sigma_A^2 = 2N_e\sigma_m^2$. 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 time until 50% of the means exceed K standard deviations is obtained from $K\sigma_z/\sqrt{t\,2\sigma_m^2} = 0.675$. Since $\sigma_z^2 = \sigma_A^2 + \sigma_E^2$,

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

where $h_m^2 = \sigma_m^2/\sigma_E^2$, the mutational heritability. From Table 1 in Chapter 9, h_m^2 has an approximate average value of 0.006. Taking this value, and repeating the calculation above (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 relative to the fixed variance example above is that additive variance is much smaller due to the small population size. Contrast this when $N_e = 100$, where $t = 3300$, while (assuming $h^2 = 0.5$) the constant variance assumption yields $t = 3600$.

The Ornstein-Uhlenbeck process is essentially Brownian motion with a linear restoring force that tends to bring the mean back to 0. The resulting diffusion for $-\infty < x < \infty$ is

$$m(x) = -ax \qquad v(x) = b \tag{A5.17}$$

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

$$\mu_t = x_0e^{-at} \qquad \sigma_t^2 = \frac{b}{2a}\,(1 - e^{-2at}) \tag{A5.18}$$

Thus the stationary distribution is normal with mean zero and variance $b/(2a)$

Example 7. Lande (1976) examined the distribution of phenotypic means under drift and stabilizing selection. Consider 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. Under nor-optimal selection, if phenotypes before selection are normally distributed with mean μ_t and phenotypic variance σ^2, they remain normal after selection, with new mean $\mu_t + s$ and variance σ_z^2 (assuming weak selection, $\omega >> \sigma_z^2$), where

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

Let x_t be the mean in generation t of a randomly drawn replicate population. Assuming $R = h^2 s$, 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 $\Delta\mu = h^2 s$ upon substituting for s. Hence, the distribution of phenotypic means in generation t is normal, with mean

$$\mu_t = \mu_0 \exp\left[- t \frac{\sigma_A^2}{\sigma_z^2 + \omega} \right]$$

and variance

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

Example 8. Diffusions can provide a more general solution for the equilibrium distribution of means under selection and drift. If phenotypes are normally distributed, then from Equation 25.12b the change in mean is

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

Hence, an approximating diffusion under selection and drift is to let x = current mean and set

$$m(x) = \sigma_A^2 \frac{d \ln(\overline{W})}{dx} \quad \text{and} \quad v(x) = \frac{\sigma_A^2}{N_e} \tag{A5.19}$$

Solving for the equilibrium distribution of means,

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

thus

$$\varphi(x) = \frac{C}{v(x)G(x)} = \frac{CN_e}{\sigma_A^2} \overline{W}(x)^{2N_e} \propto \overline{W}(x)^{2N_e} \tag{A5.20}$$

where we have written $\overline{W}(x)$ to remind the reader that mean population fitness is a function of the phenotypic mean (this result is due to Lande 1976). Thus, as 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 A5.19-20 also have the assumption that the phenotypic and additive genetic variance remains constant as the mean changes, as well as the additional assumption that the phenotypic distribution remains normal.