

## 5 Diffusion approximation

We already know that the probability of fixation of a neutral allele in a finite population of size  $N$  is  $1/N$ . Here, we derive the probability of fixation of a selected allele.

### 5.1 Derivation of the fixation probability (Diffusion approximation)

Let  $u(p)$  be the ultimate probability of fixation of allele  $\mathbf{A}$  given that its initial frequency is  $p$ . Then, using Markovian nature of the process of change in  $p$ , one can write

$$u(p) = \sum_{\Delta p} \text{Prob}(\Delta p) u(p + \Delta p), \quad (46)$$

where  $\text{Prob}(\Delta p)$  is the probability of a particular change in  $p$  in one generation, and  $u(p + \Delta p)$  is the probability of fixation given that this particular change has happened. Equation (46) can also be thought of as a variant of the generalized partition theorem in the probability theory:  $\text{Prob}(A) = \sum_i \text{Prob}(A|B_i)\text{Prob}(B_i)$ , where  $\text{Prob}(A)$  is the probability of event  $A$  and  $\text{Prob}(A|B_i)$  is the conditional probability of  $A$  given  $B_i$ . The expression in the right-hand side of (46) can also be interpreted as the mathematical expectation of  $u(p + \Delta p)$  with respect to all possible changes  $\Delta p$ , allowing one to write

$$u(p) = E_{\Delta p}(u(p + \Delta p)).$$

Expanding  $u(p + \Delta p)$  in a Taylor series

$$u(p + \Delta p) \approx u(p) + u'(p)\Delta p + \frac{1}{2}u''(p)(\Delta p)^2,$$

one finds that

$$u(p) \approx u(p) + u'(p)E[\Delta p] + \frac{1}{2}u''(p)E[(\Delta p)^2].$$

We assume that  $E\{(\Delta p)^2\} = \text{var}\{\Delta p\} + (E\{\Delta p\})^2 \approx \text{var}\{\Delta p\}$  that is that the expected change in allele frequency is small. Let  $m(p)$  and  $v(p)$  be the expected change in  $p$  and the expected variance of the change in  $p$ . Then,  $u(p)$  satisfies to a linear homogeneous second order ODE

$$m(p)u'(p) + \frac{1}{2}v(p)u''(p) = 0 \quad (47)$$

with boundary conditions  $u(0) = 0, u(1) = 1$ . The solution can be found explicitly:

$$u(p) = \frac{\int_0^p G(x)dx}{\int_0^1 G(x)dx}, \quad (48)$$

where

$$G(x) = e^{-\int \frac{2m(x)}{v(x)} dx}.$$

To find this solution we can use *Maple*. We also can do in an old-fashioned way: first, convert equation (47) to a first order ODE  $f' + (2m/v)f = 0$  by changing the dependent variable to  $f(p) = u'(p)$ , second, compute the integrating factor  $H = \exp(\int \frac{2m}{v} dp)$ , third, solve the ordinary differential equation  $d(fH)/dp = 0$  for  $f$ , and, finally, integrate the answer to find  $p$  subject to the boundary conditions.

## 5.2 Example: the probability of fixation of an advantageous allele

We consider a one-locus two-allele haploid populations of finite size  $N$ . Let the fitnesses of alleles **a** and **A** be 1 and  $1 + s$ , respectively. The mean and variance of the change in the frequency  $p$  of allele **A** in one generation are

$$m(p) = spq, \quad v(p) = pq/N,$$

where  $q = 1 - p$ . Substituting these values into the general expression (48) one finds that the probability of fixation is

$$u(p) = \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}.$$

Assuming that there is a single copy of the mutant allele initially ( $p = 1/N$ ),

$$u(1/N) = \frac{1 - e^{-2s}}{1 - e^{-2Ns}}. \quad (49)$$

If  $s$  is small (weak selection) but  $Ns$  is large (large population),

$$u \approx 2s,$$

that is the fixation probability is approximately twice the selective advantage. In large populations, selection will overwhelm drift once the advantageous allele is at all common. But only a very small proportion of advantageous mutations have a chance to become common!

## 5.3 Example: the probability of fixation of a deleterious allele

For a deleterious allele ( $s < 0$ )

$$u \approx \frac{e^{2|s|} - 1}{e^{2N|s|} - 1},$$

which can be large enough if  $Ns$  is small. Thus, in small populations, the fixation of deleterious alleles can occur with a non-negligible probability! If  $Ns$  is large, the probability of fixation  $u \approx 2s/e^{2Ns}$  and is very small.

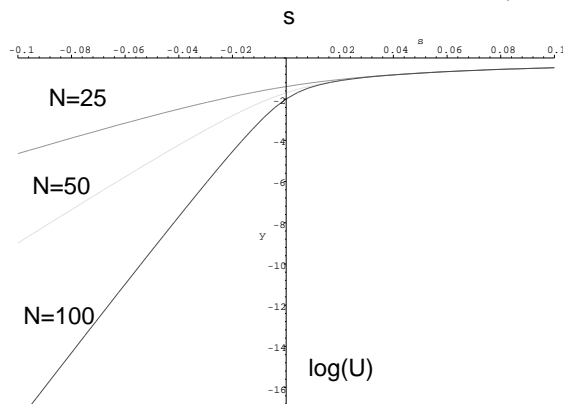


Figure 4: Fixation probability (49) as function of the allele effect  $s$  and the population size  $N$ .

---

Homework. Consider a one-locus two-allele model with frequency-dependent fitnesses

$$\begin{aligned}w_A &= 1 + sp, \\w_a &= 1 + sq,\end{aligned}$$

describing selection against rare alleles. In the deterministic approximation, the corresponding dynamic system has two locally stable monomorphic equilibria (corresponding to fixation of alleles). Finite populations subject to mutation can “jump” between these equilibria. These “jumps” will happen if an allele that has a low frequency initially gets fixed. Try to show that the probability of fixation of allele **A** that is represented by a single copy in a population of size  $N$  can be approximated (if  $N \gg 1$  and  $Ns > 4$ ) as

$$U = e^{-Ns/2} \sqrt{2s/N} \frac{1}{\sqrt{\pi}}.$$

Numerically compare  $U$  with the probability of fixation of a neutral allele for a set of parameter values (for example, with  $N = 10, 100, 1000$  and  $s = 0.1, 0.01, 0.001$ ).

---

## 5.4 Major points

In a finite population

- the probability of fixation of an advantageous allele is approximately twice its selective advantage;
- the fixation of slightly deleterious alleles can occur with a non-negligible probability;
- “jumps” between the corresponding locally stable deterministic equilibria are possible.

## 5.5 Additions: Time till absorption

Assume that both  $x = 0$  and  $x = 1$  are absorbing boundaries (e.g., corresponding to a loss or fixation of an allele). Let  $\phi(t, p)$  be the density function of the time until absorption occurs given the initial value of  $x$  is  $p$ . Under diffusion approximation,  $\phi(t, p)$  satisfies the equation

$$\frac{\partial \phi(t, p)}{\partial t} = a(p) \frac{\partial \phi(t, p)}{\partial p} + \frac{1}{2} b(p) \frac{\partial^2 \phi(t, p)}{\partial p^2}.$$

Let

$$\bar{t}(p) = \int_0^\infty t \phi(t; p) dt$$

be the average time till absorption. Note that  $\bar{t}(0) = \bar{t}(1) = 0$  and that the first and second derivatives of  $\bar{t}(p)$  are

$$\frac{d\bar{t}(p)}{dp} = \int_0^\infty t \frac{\partial \phi(t; p)}{\partial p} dt$$

and

$$\frac{d^2 \bar{t}(p)}{dp^2} = \int_0^\infty t \frac{\partial^2 \phi(t; p)}{\partial p^2} dt.$$

Following Ewens (1979,2004) and using integration by parts,

$$\begin{aligned} -1 &= - \int_0^\infty \phi(t; p) dt = \\ &= -[t\phi(t, p)]_0^\infty + \int_0^\infty t \frac{\partial \phi}{\partial t} dt \\ &= 0 + \int_0^\infty t \left[ a(p) \frac{\partial \phi}{\partial p} + \frac{1}{2} b(p) \frac{\partial^2 \phi}{\partial p^2} \right] dt, \end{aligned}$$

so that

$$1 + a(p) \frac{d\bar{t}(p)}{dp} + \frac{1}{2} b(p) \frac{d^2 \bar{t}(p)}{dp^2} = 0. \tag{50a}$$

Thus, the average time till absorption can be found by solving the second order ordinary differential equation (1a) subject to boundary conditions

$$\bar{t}(0) = \bar{t}(1) = 0. \tag{50b}$$