Population Genetics Course: Selection Notes 2

1. Dynamics of Selection on a Favourable Allele (Haploid Population)

Assume that the genotypic fitnesses are as follows:

A_1	A_2
1	$1 + s$

If the frequency of A_1 in a given generation is p and that of A_2 is $q = 1 - p$, a similar argument to that used in **Selection Notes 1** shows that the values in the next generation are:

$$p' = p/(1 + sq)$$

Similarly, the new allele frequency q' of A_2 is given by

$$q' = q(1 + s)/(1 + sq)$$

We can get rid of the annoying term $1 + sq$ in the denominators of these equations by taking the ratios of the two equations:

$$(q'/p') = (q/p)\ (1 + s) \qquad (1)$$

Writing u for the ratio q/p, we then find that the values of u in successive generations are connected by the simple relation

$$u' = u(1 + s)$$

If this is repeated over many generations, we can write the value of u in a given generation t in terms of its value in an initial generation 0 as

$$u_t = u_0(1 + s)^t \qquad (2)$$

Taking natural logarithms, this gives

$$\ln(u_t / u_0) = t \ln (1 + s)$$

If s is small, $t \ln(1 + s)$ is approximately ts (see your basic mathematics notes), so that the time to change allele frequencies by a specified amount can be written as

$$t \approx (1/s) \ln\{q_t p_0 / q_0 p_t\} \qquad (3)$$

This brings out the very important point that there is only a *logarithmic* dependence of the time needed to change the frequency of a favourable allele on its initial value, if the allele affects the fitness of its heterozygous carriers.

2. Survival Probability of a Favourable Mutation (Haploid Population)

We assume that individuals of a given genotype are characterized by a characteristic distribution of completed family size (the distribution of the number of adult individuals contributed by an adult to the next generation), such that P_i is the probability that an individual contributes i offspring ($i = 0, 1, 2, ...$). Let E be the probability of ultimate extinction of the mutant allele, i.e. the probability that no descendants are ultimately left by the initial copy. This can happen if it leaves no offspring in the first generation, if it leaves one offspring in the first generation but its descendants ultimately go extinct (probability E), and so on.

In general, if i offspring are contributed to the first generation, the chance that they all ultimately fail to leave descendants is E^i. The net chance of extinction thus satisfies the relation

$$E = P_0 + P_1 E + P_2 E^2 + P_3 E^3 + ... + P_i E^i + .. \qquad (1)$$

For any given assumed distribution of offspring number, this equation can always be solved numerically. One general result that can be derived from this is that E is less than one only if the mean of the offspring distribution exceeds one i.e. a mutant is doomed to stochastic extinction if its *absolute* fitness is less or equal to one, but will have a finite chance of survival if its absolute fitness is greater than one, regardless of its performance relative to the rest of the population.

Haldane's approximation:

In the case of weak selection and a constant population size, and assuming a haploid asexual population initially fixed for allele A_1, we can obtain a very useful approximate result for the chance of survival of a new favourable mutation A_2, which is due to Haldane (1927).

If the population size of the A_1 population is constant, it has a mean number of offspring per individual of 1. (If the mutant is neutral, it follows that its mean offspring number is also one, so that it is doomed to extinction, using the result stated above.) If mutants A are favoured by selection, we can write their mean offspring number as $1+s$, where s is assumed to be small.

Assume further that the offspring distribution is Poisson, which is reasonable if there is random variation in reproductive success for a given genotype. This means that the offspring distribution is given by:

$$P_i = (1+s)^i \exp\{-(1+s)\} / i ! \qquad (2)$$

(where we use exp x to denote e^x; see your basic maths notes).

Substituting this into equation (1), we get

$$E = \{\exp -(1+s)\}\ \{1 + (1+s)E + (1+s)^2E^2/2 + (1+s)^3 E^3/3! \ldots\}$$
$$= \{\exp -(1+s)\}\ \{\exp E\,(1+s)\}$$

$$= \exp\{-(1-E)(1+s)\} \qquad (3a)$$

Let the chance of survival of the mutation be $Q = 1-E$; we get:

$$1 - Q = \exp\{- Q(1+s)\} \qquad (3b)$$

By using the formula for expanding exp $-x$ as a series in x, the right hand-side of (3b) can be written as

$$1 - Q(1+s) + Q^2(1+s)^2/2, \qquad \text{plus higher-order terms in } U \text{ and } s$$

Since Q and s must be similar in size, this can be approximated further by

$$1 - Q(1+s) + Q^2/2$$

Equating this to the left-hand side of (3b), we get the approximation

$$Q \approx 2s \qquad (4)$$