Natural Selection and the Price Equation

The role of natural selection in the evolution of organisms adapted to their environment was first described by Charles Darwin and Alfred Russel Wallace in a joint presentation to the Linnean Society of London in 1858. Working independently of one another, they reasoned that selection will alter the distribution of a trait expressed by the members of a population whenever the following conditions are satisfied:

- Phenotypic variation: the trait is variable in the population.
- Fitness differences: the trait influences reproductive success.
- Heritability: offspring tend to resemble their parents in this trait.

Under these conditions, the population will change over time so that traits with higher fitness will increase in frequency at the expense of traits with lower fitness. In particular, in stable environments, the population may evolve to become better adapted to the environment. Traits can affect fitness in different ways and may involve differences in survival, mating probability or fecundity (number of offspring). In particular, although natural selection has often been equated with the phrase 'the survival of the fittest', selection can also decrease survival if this is compensated by an increase in the number of offspring that are left by an individual. It is also important to note that selection is not the only process that influences how populations evolve. Mutation and recombination are both important evolutionary processes because they help to generate the genetic variation on which selection acts, while random variation in reproductive success can lead to a loss of variation or even allow less fit mutations to spread through a population.

Our goal in these notes is to describe some simple mathematical models of natural selection. We begin with an important conceptual result known as the Price equation, which was first derived by George R. Price. We first introduce some notation. We will consider a single real-valued trait (such as height or weight) that currently assumes $k$ possible values, say $z_1, \ldots, z_k$, in a population containing $n$ individuals. Let $n_i$ be the number of individuals in the current generation that have trait value $z_i$ and notice that the population mean of the trait in this generation is

$$\bar{z} = \frac{1}{n} \sum_{i=1}^{k} n_i z_i = \sum_{i=1}^{k} p_i z_i,$$

where $p_i = n_i/n$ is the frequency of trait $z_i$ in the present generation.

Assume that the population has non-overlapping generations and let $n_i'$ be the number of offspring born to the $n_i$ individuals with trait $z_i$ and write $n' = n'_1 + \cdots + n'_k$ for the size of the population in this next generation, which may differ from $n$. If we let $w_i = n'_i/n_i$, then we can interpret $w_i$ as the fitness of trait $z_i$, i.e., $w_i$ is the mean number of offspring per capita born to individuals with trait $z_i$. Furthermore, the mean fitness of the population in the current generation is

$$\bar{w} = \frac{1}{n} \sum_{i=1}^{k} n_i' z_i = \sum_{i=1}^{k} p_i' z_i,$$

where $p_i' = n_i'/n$. However, it is more convenient to compute this mean fitness using the frequency of traits in the next generation:

$$\bar{w} = \frac{1}{n'} \sum_{i=1}^{k} n_i' z_i.$$
\[
\bar{w} = \frac{1}{n} \sum_{i=1}^{k} n_i w_i = \frac{1}{n} \sum_{i=1}^{k} n'_i \frac{n'}{n},
\]

i.e., \(n' = n\bar{w}\). To describe how the population mean of the trait changes over one generation, we need to specify a relationship between the parental and the offspring values of the trait. To this end, let \(z_i'\) denote the mean value of the trait in the \(n'_i\) offspring born to the individuals expressing trait \(z_i\) and let \(\Delta z_i = z_i' - z_i\) denote the change in these means. Notice that \(\Delta z_i\) measures the transmission bias in the trait as expressed in the \(i\)'th group. For example, if the trait is partially controlled by environmental factors, then a change in the environment could bias the mean value of the trait to either higher or lower values, e.g., an improvement in early childhood nutrition will tend to increase the average heights of offspring over those of their parents. With these conventions, we see that the mean value of the trait amongst the offspring of the current generation is

\[
\bar{z}' = \frac{1}{n'} \sum_{i=1}^{k} n'_i z_i' = \frac{1}{n\bar{w}} \sum_{i=1}^{k} n_i w_i (z_i + \Delta z_i)
\]

where we have used the fact that \(z_i' = z_i + \Delta z_i\) to obtain the final expression.

To find the change in the population mean of the trait, we calculate \(\Delta \bar{z} = \bar{z}' - \bar{z}\):

\[
\Delta \bar{z} = \frac{1}{n\bar{w}} \sum_{i=1}^{k} n_i w_i (z_i + \Delta z_i) - \frac{1}{n} \sum_{i=1}^{k} n_i z_i
\]

\[
= \frac{1}{n\bar{w}} \sum_{i=1}^{k} n_i (w_i - \bar{w}) z_i + \frac{1}{n\bar{w}} \sum_{i=1}^{k} n_i w_i \Delta z_i
\]

\[
= \frac{1}{\bar{w}} \sum_{i=1}^{k} p_i (w_i - \bar{w}) z_i + \frac{1}{\bar{w}} \sum_{i=1}^{k} p_i w_i \Delta z_i.
\]

Both sums in this final line can be interpreted as population means. The second term is easier to interpret and is just \(\mathbb{E}[w_i \Delta z_i]\), i.e., this is the population mean of the product of the fitness and the transmission bias. To interpret the first term, observe that

\[
\sum_{i=1}^{k} p_i (w_i - \bar{w}) \bar{z} = \bar{z} \sum_{i=1}^{k} p_i w_i - \bar{z} \bar{w} \sum_{i=1}^{k} p_i
\]

\[
= \bar{z} \bar{w} - \bar{z} \bar{w} = 0,
\]

where we have used the fact that the population frequencies sum to 1, i.e., \(p_1 + \cdots + p_k = 1\). It follows that

\[
\sum_{i=1}^{k} p_i (w_i - \bar{w}) z_i = \sum_{i=1}^{k} p_i (w_i - \bar{w}) (z_i - \bar{z}) = \text{Cov}(w_i, z_i),
\]

i.e., the first term is proportional to the covariance between the fitness and the trait value. Combining these results, we arrive at our main result:
**Theorem 1** (Price Equation). *The change in the mean value of the trait is given by*

\[
\Delta \bar{z} = \frac{1}{\bar{w}} \text{Cov}(w_i, z_i) + \frac{1}{\bar{w}} \mathbb{E}[w_i \Delta z_i].
\]

The Price equation decomposes the change in the mean value of the trait into the sum of two contributions. The first term quantifies the change in mean due to selection. For example, if fitness increases with the value of the trait, then the covariance \( \text{Cov}(w_i, z_i) \) will be positive, meaning that individuals with larger values of the trait will tend to leave more offspring, and so selection will tend to increase the mean value of the trait in the population. On the other hand, if the trait has no effect on fitness, then all of the \( w_i \)'s will be equal to the same constant value \( \bar{w} \). In this case, the covariance will be equal to 0, i.e.,

\[
\text{Cov}(w_i, z_i) = \sum_{i=1}^{k} p_i(\bar{w} - \bar{w})(z_i - \bar{z}) = 0,
\]

and so selection will have no influence on the mean value of the trait. In contrast, the second term quantifies the average effect of the transmission bias, weighted by the fitness of the trait. If transmission is unbiased, e.g., if each offspring inherits its parent’s trait without any change, then \( \Delta z_i = 0 \) for all \( i = 1, \cdots, k \) and this term will vanish.

By treating fitness as a trait, i.e., setting \( z_i = w_i \), we obtain an important special case of the Price equation. Since \( \text{Cov}(w_i, w_i) = \text{Var}(w_i) \), we have

\[
\Delta \bar{w} = \frac{1}{\bar{w}} \text{Var}(w_i) + \frac{1}{\bar{w}} \mathbb{E}[w_i \Delta w_i].
\]

In particular, if fitness is inherited without any transmission bias, then the second term vanishes and the Price equation reduces to the following identity

\[
\Delta \bar{w} = \frac{1}{\bar{w}} \text{Var}(w_i),
\]

which is a version of Fisher’s **Fundamental Theorem of Natural Selection**. Since the variance of a non-degenerate random variable is always positive, the fundamental theorem asserts that in the absence of transmission bias, the population mean fitness will increase from one generation to the next provided that there is variance in fitness in the current generation. This result was first derived by R. A. Fisher and has been the subject of much debate ever since.

The controversy lies not in the identity itself, which is mathematically correct, but rather in the validity of the condition that fitness is inherited in an unbiased fashion. There are many reasons why this need not be true. For instance, if offspring traits differ from parental traits because of mutation or environmental noise, then in general we might expect these random perturbations to reduce the average offspring fitness, especially in a population which is near a fitness optimum. In this case, the second term will be negative and may even dominate the first term, so that the overall change in mean fitness is negative. Likewise, if fitness is context dependent, then environmental changes may alter the fitness of each trait value so that traits that are beneficial under current conditions become disadvantageous in subsequent generations. For example, while a larger body size may be advantageous in a cooler environment, the opposite will be true if the
environment begins to warm. Here too the second term will tend to be negative, meaning that the mean fitness may decrease in some generations.