

The Price equation

Elliot Aguilar and Erol Akçay

March 20, 2019

In this lecture, we will get to know one of the most general and powerful frameworks to study evolutionary change, the Price equation, so named because it was proposed in its most general form by George Price, who was one of the most interesting characters in evolutionary theory in the second half of the 20th century. He came into evolutionary biology at age 45, after first having a couple of different careers as a chemical physicist, writer, and computer scientist. Between 1967 and 1975, when he died, he made two major contributions to evolutionary biology: he invented evolutionary game theory and the Price equation.

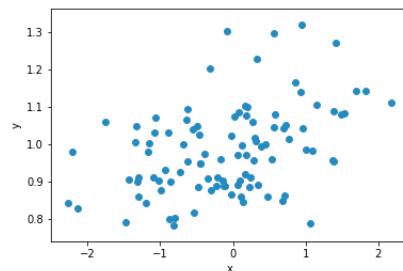
see Frank (1995) for more on Price's contributions to evolutionary theory.

The Price equation is a general description of how a mean of a phenotype (any measurable characteristic) in a population changes from one time point to another. As such, it has a different focus than what we have considered so far: the notion of genes are not an essential part of the Price equation (though the Price equation can be used to do population genetics). In fact, neither is biology an essential ingredient of the Price equation. The equation is an identity that describes change in the mean of a measurable characteristic between any two sets of entities (trees, squirrels, songs, cars, countries, etc.), entities in one group can be labeled as descendants of entities in the other group.

Let's first look at the mathematics underlying the Price equation, which is somewhat tedious (this should be familiar by now) but at heart, quite simple. We first need a refresher on a mathematical object that you've probably seen in your stats class.

Covariance and Regression

First we define two important concepts for understanding the Price equation. Consider a sample from pairs of two random variables, X and Y , as in the figure below. We want to express the relation between the values of the variables in each pair, *in this sample of points*.



There are a couple of different ways to express this relationship. We begin with **the covariance**:

$$\text{cov}(x, y) = E[(x - \bar{x})(y - \bar{y})] \quad (1)$$

This captures the idea that if we were to plot a set of points in the 2D plane whose coordinates were the successive realizations of the random variables X and Y , then the expected product of the difference of each random variable and its mean would give us an idea of the extent to which the two random variables covary. We can expand the above definition to find a more useful expression for the covariance: Let's define the covariance:

$$\begin{aligned} \text{cov}(x, y) &= E[xy - x\bar{y} - y\bar{x} + \bar{x}\bar{y}] \\ &= E(xy) - E(x\bar{y}) - E(y\bar{x}) + E(\bar{x}\bar{y}) \\ &= E(xy) - \bar{y}\bar{x} - \bar{x}\bar{y} + \bar{x}\bar{y} \\ &= E(xy) - E(x)E(y) \end{aligned} \quad (2)$$

Now, covariance has a problem—if we change the units of X and Y (say, inches vs. centimeters) then our covariance will change as well. You can see this because what we're considering are “distances,” the numerical value of which will vary depending on our measuring stick. In some cases, we may want another measure of the relationship between two random variables that gives us how much, say, y changes on average, if we change x by a standard amount, regardless of the scale we use to measure x (but not regardless of the scale used to measure y). This measure is called **the linear regression coefficient**, which we'll consider below.

Again, return to the plot of points whose coordinates were the realizations of two random variables X and Y . We can try to fit a linear function to these points, (of the form $y = mx + b$, where m is the slope and b is the y-intercept) so see how well X ‘predicts’ Y . We say that $E(y) = \beta_0 + \beta_{y,x}x$, or, in other words, we expect Y to be a linear function of X , though there is some random noise that may make any particular y deviate from this line. Here, β_0 is now our y-intercept and $\beta_{y,x}$ is the slope of the line, known as the regression coefficient or linear correlation. We read $\beta_{y,x}$ as the “regression of y on x .” It can be shown that,

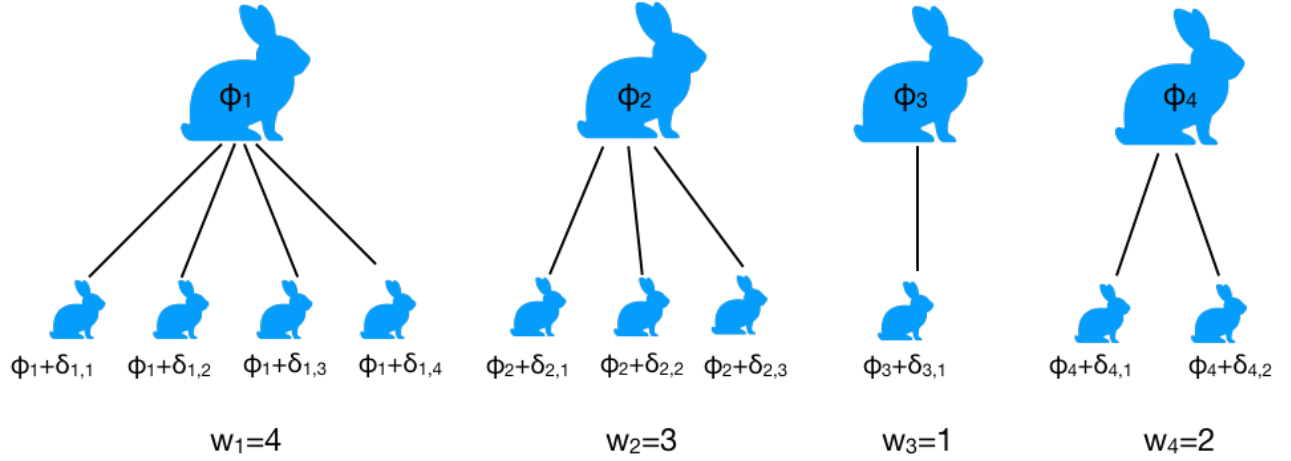
$$\beta_{y,x} = \frac{\text{cov}(x, y)}{\text{var}(x)} \quad (3)$$

a fact that we will use later.

The Price Equation

We're now ready to derive the Price Equation (our derivation will follow Rice, 2004, chapter 6; see readings). We begin by noting that an organism interacts with its environment (including its social environment) through its phenotype,

which in general will determine its fitness. The Price equation is about how change in the mean phenotype can be expressed in terms of the relationship between phenotypes and fitness. Below we will assume that an individual's phenotype can be expressed as a real number value. Consider a situation as in the figure below, which depicts a small population of rabbits and their descendants:



Here's a bit of a notation dump that will help us construct the Price equation:

1. N population size in the ancestor generation.
2. ϕ_i phenotype of individual i in the ancestor generation.
3. $\bar{\phi}$ the mean phenotype of individuals in the ancestor generation.
4. $\delta_{i,j}$ the difference between the phenotype of i and her j th descendant.
5. δ_i the mean difference in phenotype between i and her descendants.
6. w_i the fitness measured in number of descendants of individual i .
7. \bar{w} the mean fitness of individuals in the ancestor generation.

Now, let $\bar{\phi}'$ be the mean phenotype in the descendant generation. We can write:

$$\bar{\phi}' = \frac{\sum_{i=1}^N \sum_{j=1}^{w_i} (\phi_i + \delta_{i,j})}{\sum_{i=1}^N w_i} \quad (4)$$

If equation (4) looks daunting, just remember the simple fact that the average phenotype of descendants is the sum of all descendant phenotypes divided by the total number of descendants. Now remember from the definitions given above that the phenotype of descendant j of ancestor i can be written as the phenotype of the ancestor plus the difference between the two, or $\phi_i + \delta_{i,j}$. The first sum in the numerator is just summing the phenotypes of all the descendants of i (who has w_i descendants), while the outer sum then adds over all the N ancestors. The denominator is the total number of descendants. Below we will

make use of the following identities:

$$\sum_{i=1}^N w_i = N\bar{w} \quad (5)$$

$$\sum_{j=1}^{w_i} \phi_i = \phi_i w_i \quad (6)$$

$$\sum_{j=1}^{w_i} \delta_{i,j} = w_i \delta_i \quad (7)$$

Consult the definitions given above and check to see why each of these identities is true. Next, we can use them to rewrite (4) as,

$$\bar{\phi}' = \frac{\sum_{i=1}^N \phi_i w_i + \sum_{i=1}^N \delta_i w_i}{N\bar{w}} \quad (8)$$

or,

$$\bar{\phi}' = \frac{1}{\bar{w}} [E(\phi_i w_i) + E(\delta_i w_i)] \quad (9)$$

Where the above expectations are with respect to the ancestors that range 1 through N (e.g. $\frac{1}{N} \sum_{i=1}^N \phi_i w_i = E(\phi_i w_i)$). Now using (2), we can make the substitution $E(xy) = \text{cov}(x, y) + E(x)E(y)$, and we have,

$$\bar{\phi}' = \frac{1}{\bar{w}} [\text{cov}(\phi_i, w_i) + \bar{\phi}\bar{w} + E(\delta_i w_i)] \quad (10)$$

$$\bar{\phi}' - \bar{\phi} = \frac{1}{\bar{w}} [\text{cov}(\phi_i, w_i) + E(\delta_i w_i)] \quad (11)$$

For clarity we'll rewrite $\bar{\phi}' - \bar{\phi}$ as $\Delta\bar{\phi}$, which gives us the Price Equation (sometimes also called Price's Theorem) for evolutionary change,

$$\Delta\bar{\phi} = \frac{1}{\bar{w}} [\text{cov}(\phi_i, w_i) + E(\delta_i w_i)] \quad (12)$$

The above equation tells us that the change in the mean phenotype in the population is the sum of two terms: 1) $\frac{1}{\bar{w}} \text{cov}(\phi_i, w_i)$ is the covariance between fitness and phenotype, which includes all directional forces in evolution, such as selection and drift, which act on the replicating units, and 2) $\frac{1}{\bar{w}} E(\delta_i w_i)$ which includes nondirectional forces, like mutation, regression to mean phenotype, or even selection acting at lower levels than the unit of replication we're considering (for instance, we could write the price eq. down for groups, as long as we define the phenotypes of a group, and the second term on the right hand side of 2.9 will contain selection acting on individuals within those groups. see Rice, 2004 for an example of multilevel selection). There were very few assumptions involved in deriving equation (12), but to be clear let's list them:

1. **real-valued phenotypes:** Phenotypes were assigned real number values.
2. **ancestor-descendant relationships:** Descendants could all be attributed to ancestors. Think of this as being able to draw lines connecting and descendant to its ancestor(s), as we did for the rabbits above.

These are really innocuous assumptions. Even the second one, which sounds a bit more restrictive, is actually a lot less so than it appears. For example, it is possible to draw lines to more than one ancestor for each descendant (as would happen in sexual species). As long as everyone has the same number of ancestors (again, as in most organisms), this is almost trivial, but it can even work if some descendants have one ancestors while others have 27 (to make a wild number). It is also possible that some of the “descendants” are really the same individuals as “ancestors,” e.g., because they simply survived into the second time point we censused the population. The Price equation is still valid. It also does not really matter how far apart the ancestor and descendant populations are from each other in time, as long as one can draw the connections from ancestors to descendants. Of course in all of these cases, the interpretation of the quantities w and δ , and the covariances change. But the identity stays the same.

See Kerr and Godfrey-Smith (2009) for a generalization for this case.

The Price equation is thus a general statement of evolutionary change. There’s essentially no biology involved in its derivation, and thus it holds for any biological (or otherwise) system we could imagine, as long as the two assumptions above are met.

Haploid model

We’ll use an example from previous lectures to demonstrate the Price equation, the haploid single-locus model from the previous lecture. There are two alleles 1 and 2 at this locus, with fitness w_1 and w_2 , respectively. The phenotype here under consideration is the genotype, which is transmitted perfectly, and $E[\delta_i w_i] = 0$. The mean phenotype in the population is equal to the frequency of allele 1,

$$\Delta \bar{\phi} = \Delta p = \frac{1}{\bar{w}} \text{cov}(p_i, w_i) = \frac{1}{\bar{w}} \frac{1}{N} \sum_{i=1}^N p_i w_i - p = \frac{1}{\bar{w}} [w_1 p - \bar{w}] \quad (13)$$

$$= \frac{1}{\bar{w}} p(1-p)[w_1 - w_2] = \frac{\text{var}(p_i)}{\bar{w}} [w_1 - w_2] \quad (14)$$

Evolutionary Game theory

As mentioned above, in addition to deriving the Price equation, George Price also invented evolutionary game theory. Let’s look at the Hawk-Dove game we have seen before, but using the Price equation now. To refresh your memory, the Hawk-Dove game runs as follows: imagine a population of organisms that can play one of two strategies: Hawk (fight) or Dove (retreat). The organisms meet in antagonistic encounters over a resource which produces a fitness benefit. The payoff matrix below describes the payoff to a player (given by the row) when up against an opponent’s strategy (given by the column):

	Hawk	Dove
Hawk	$1/2(V-C)$	V
Dove	0	$1/2V$

When a Hawk encounters a Dove, he gets the entire benefit, V . When a Hawk encounters a Hawk, he pays a cost C for fighting and can expect to be the winner half the time. When a Dove encounters another Dove, he can again expect to be the winner half the time. We can now write the expected fitnesses for hawks and doves:

$$w_H = w^* + \frac{1}{2}(V - C)p - (1 - p)V \quad (15)$$

$$w_D = w^* + \frac{1}{2}(1 - p)V \quad (16)$$

We can now use the Price equation to compute the change in mean phenotype (the frequency of Hawks). Inheritance is again perfect, so $E[\delta_i w_i] = 0$. Thus,

$$\Delta \bar{\phi} = \frac{1}{\bar{w}} \text{cov}(p_i, w_i) = \frac{1}{\bar{w}} [p w_H - \bar{w} p] = \frac{p(1 - p)}{\bar{w}} (w_H - w_D). \quad (17)$$

Using the fitness functions above, we can calculate $w_H - w_D = \frac{1}{2}[V - C]$, and,

$$\Delta p = \frac{p(1 - p)}{2\bar{w}} (V - Cp). \quad (18)$$

Fisher's Fundamental Theorem

The Price equation also allows us to get at one of the deepest (and most puzzling, in some ways) results in evolutionary genetics. What if we consider fitness itself as the phenotype of interest, which we can do (because it is a real valued attribute of individuals)? But fitness is a tricky quantity, since in order to know the actual fitness of the descendants, we have to actually know their descendant population (the grand-descendants) as well. So for writing the change in mean fitness, we need to know three generations instead of two. But suppose we do know the three generations involved, and therefore know the fitnesses of everyone in the descendant (middle) generation. We can then write:

$$\Delta \bar{w} = \frac{1}{\bar{w}} [\text{cov}(w_i, w_i) + E(\delta_i w_i)] = \frac{1}{\bar{w}} [\text{var}(w) + E(\delta_i w_i)] . \quad (19)$$

Now, suppose fitness is inherited in an unbiased way, i.e., the descendants of an individual, on average can expect to have the same fitness as their parents. In that case, the second term will vanish, and we will be left with

$$\Delta \bar{w} = \frac{1}{\bar{w}} \text{var}(w) ,$$

which is always positive. This is the simplest derivation and expression of what R.A. Fisher called “the Fundamental Theorem of Natural Selection.” It says

that the change in the mean fitness of a population is proportional to the variance in fitness. Since that quantity is always positive, it appears that mean fitness would always go up. But not so fast.

We made an assumption in deriving (), that the fitness of offspring of a given parent, on average, is the same as their parents' fitness. But that turns out to be a very non-innocuous assumption. In particular, consider the Hawk-Dove game above, with the population composed mostly of Doves. We know that the frequency of Hawks is then predicted to increase. But that means that the offspring of both Hawks and Doves will encounter more Hawks than their parents did, which reduces the fitness of both. Therefore, on average, the offspring of both Hawks and Doves, will have lower fitness than their parents, and the second term will not vanish (indeed will be negative). Thus, the total change in mean fitness can still go down.

In fact, Fisher was not claiming that the total evolutionary change in mean fitness was the variance in w , but only that part of the change attributable to selection. By selection Fisher meant the additive effects of genes; all other effects, including genetic effects like dominance, Fisher considered to be environmental. Specifically, he referred to the other factors influencing the change in mean fitness as the “deterioration due to the environment.” In our example above, this deterioration would be the proliferation of Hawks, which are a nuisance to everyone. This is because Fisher thought that the change in mean fitness must be near zero, or else a population would quickly reach extinction or overpopulation. Since the effect of selection must be positive, the so-called effect of the environment must be negative or a “deterioration.”

In sum, the Price equation is a very general statement that can be used to obtain many classical and non-classical results. Yet its interpretation has always been somewhat fraught, and people have debated (and continue to do so) whether it is useful at all. See (Frank, 1997) for an exposition of the different interpretation of the Price equation (coming down on the side of being useful; you can guess what side I'm on from the fact that you experienced this lecture).

References

- Frank, S. A. 1995. George price's contributions to evolutionary genetics. *Journal of theoretical biology* 175:373–388.
- . 1997. The price equation, fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* pages 1712–1729.
- Kerr, B., and P. Godfrey-Smith. 2009. Generalization of the price equation for evolutionary change. *Evolution: International Journal of Organic Evolution* 63:531–536.

Rice, S. H. 2004. Evolutionary theory: mathematical and conceptual foundations. Sinauer Associates Sunderland.