Categorical traits and the strength of selection

Stephen Francis Mann | stephenfmann@gmail.com

LOGOS Research Group | **University of Barcelona**

Department of Linguistic and Cultural Evolution | **Max Planck Institute for Evolutionary Anthropology**

1. Introduction

Selection acts on categorical traits. The Price equation is usually expressed in terms of change in the average value of a trait that takes numeric values (such as radius; Box 3). However, selection can occur on categorical traits too (such as colour; Box 5).

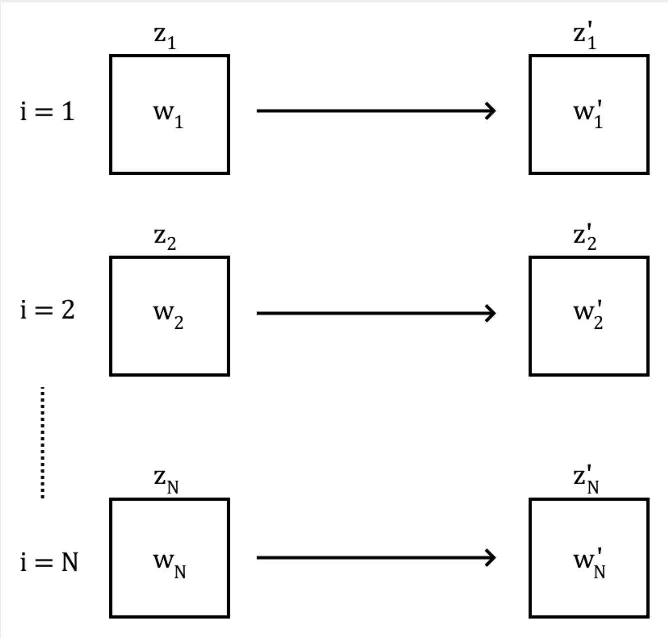
Formal descriptions should reflect the breath of selection. For both numeric and categorical traits, different trait values are associated with differential changes in population share. A general account of selection should encompass selection on categorical variables.

This poster advocates:

1. Representing categorical traits with **one-hot vectors**, such that a general form of the Price equation holds true for them (Boxes 3, 4 & 5);
2. Measuring the strength of selection in terms of **relative entropy**, which has sense for both numeric and categorical traits (Box 6).

2. Framework & Terminology

George Price's framework models populations undergoing selection. The population is divided into N types, indexed by i.
The **population share** of the ith type is w_i before selection. The population share of its direct descendants is w_i'.
The **trait value** of the ith type is z_i. The trait value associated with its direct descendants is z_i'.
The **selection coefficient** (fitness) of type i is $c_i = \frac{w_i'}{w_i}$.
Selection occurs when the population share of at least one type differs from its descendants, i.e. $c_i \neq 1$.



4. Representing categorical traits as one-hot vectors

Vectors represent categorical traits. Let trait values z_i be vectors. Elements $_j z_i$ give the proportion of category j belonging to type i.

Example: with three colours, blue, red and green:

$$\langle 1,0,0\rangle \implies \text{all blue}$$
$$\left\langle 0,\frac{1}{2},\frac{1}{2}\right\rangle \implies \text{half red, half green}$$

One-hot vectors represent distinct types. Exactly one entry takes the value 1, the rest are 0.

$$z_1 = \langle 1,0,0\rangle \implies \text{blue}$$
$$z_2 = \langle 0,1,0\rangle \implies \text{red}$$
$$z_3 = \langle 0,0,1\rangle \implies \text{green}$$

The mutation vector z_i' says what proportion of a type's descendants are of a given type. Examples:

$$z_1' = \langle 1,0,0\rangle \implies \text{blue items remain blue}$$
$$z_2' = \left\langle 0,\frac{3}{4},\frac{1}{4}\right\rangle \implies \text{quarter of red items become green}$$
$$z_3' = \left\langle \frac{1}{3},0,\frac{2}{3}\right\rangle \implies \text{third of green items become blue}$$

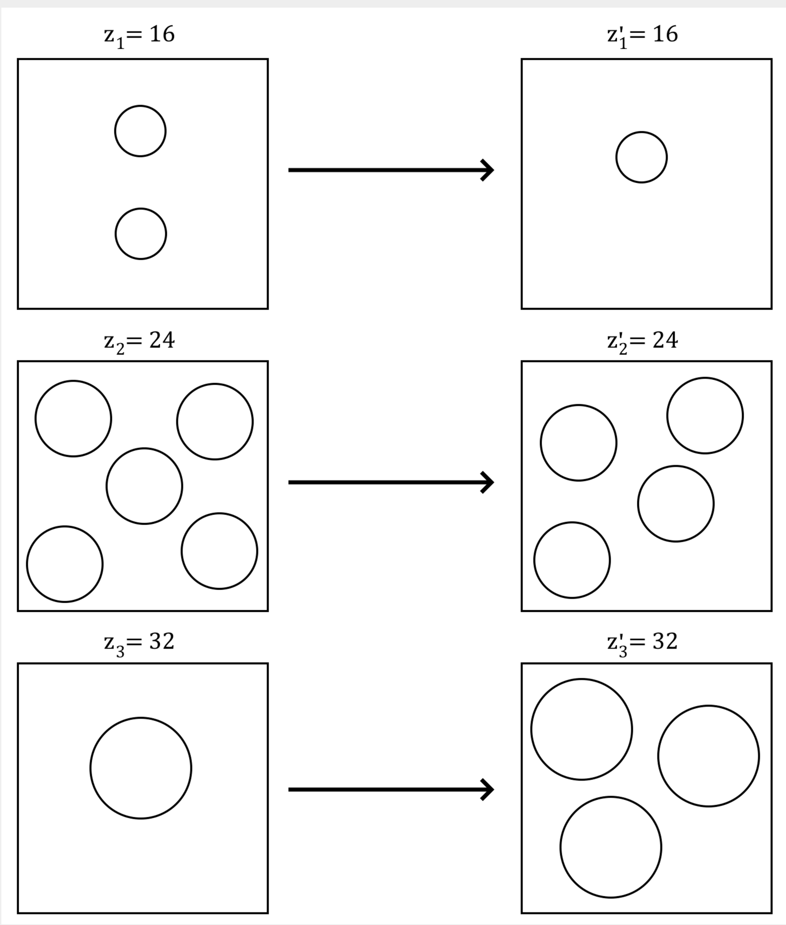
3. The Price equation for numeric variables

Weighted sums are population **averages**

The Price equation describes change as a consequence of selection plus other causes. It is usually written:

$$\underbrace{\Delta\bar{z}}_{\text{change in average trait value}} = \underbrace{\Delta_s\bar{z}}_{\text{change due to selection}} + \underbrace{\Delta_t\bar{z}}_{\text{change due to transmission}} = \mathrm{cov}_w(c,z) + E_{w'}(\Delta z) \qquad (1)$$

Example: w_i is the population share of type i $\left(\frac{2}{8},\frac{5}{8},\frac{1}{8}\right)$, and z_i is the radius of items of type i $(16, 24, 32)$:



Covariance measures change due to selection. The population average radius increases, and the covariance is:

$$\begin{aligned}\mathrm{cov}_w(c,z) &= \mathbb{E}_w(cz) - \mathbb{E}_w(c)\mathbb{E}_w(z)\\ &= 2+12+12 - (1)(23)\\ &= 26 - 23\\ &= \mathbf{3}\end{aligned}$$

So the average radius increase due to selection is 3 size units. There is no mutation in this example, so selection accounts for all of the change in average radius.

Selection without numbers? The population share of different types can change, without those types being defined in terms of numeric traits like radius. How can the Price equation describe change in this case?

5. The Price equation for categorical variables

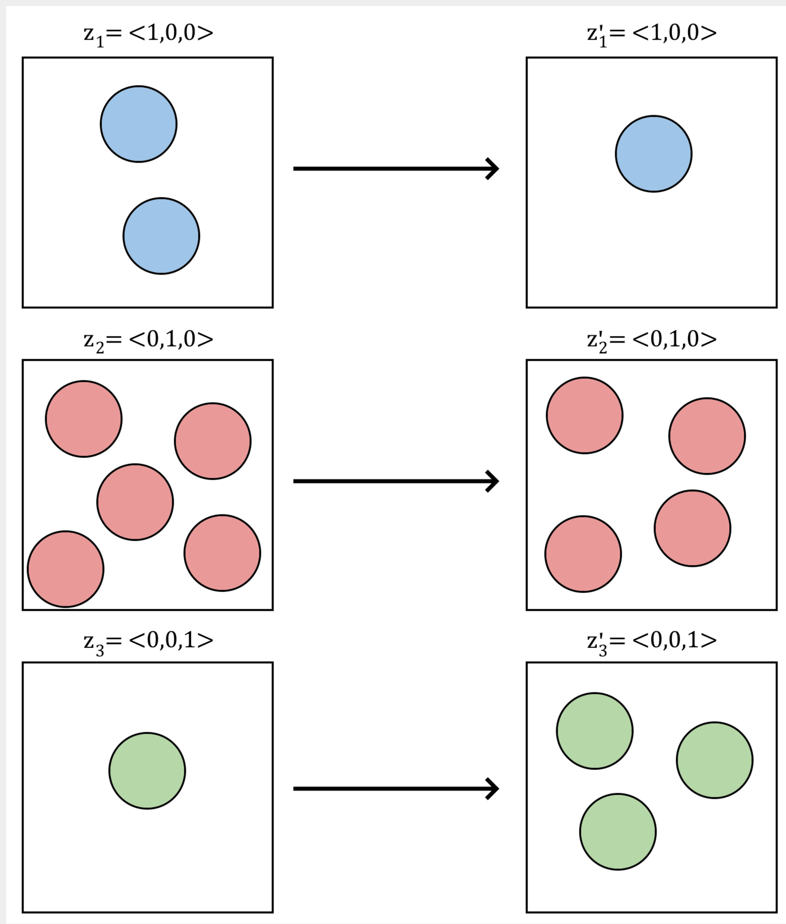
Weighted sums are population **proportions**

A Price equation for categorical variables results from generalising equation (1):

$$\underbrace{\sum_i w_i' z_i' - \sum_i w_i z_i}_{\text{change in weighted sum of trait values}} = \underbrace{\sum_i z_i(w_i' - w_i)}_{\text{change due to selection}} + \underbrace{\sum_i w_i'(z_i' - z_i)}_{\text{change due to transmission}} \qquad (2)$$

Equation (2) becomes (1) when z is a numeric trait, but equation (2) also holds when z is not numeric.

Example: w_i is the population share of type i $\left(\frac{2}{8},\frac{5}{8},\frac{1}{8}\right)$, and z_i is the colour of items of type i:



Vectors capture the change in population proportions. The selection term of equation (2) is:

$$\begin{aligned}\sum_i z_i(w_i' - w_i) &= \langle 1,0,0\rangle\left(\frac{1}{8}-\frac{2}{8}\right)\\ &+ \langle 0,1,0\rangle\left(\frac{4}{8}-\frac{5}{8}\right)\\ &+ \langle 0,0,1\rangle\left(\frac{3}{8}-\frac{1}{8}\right)\\ &= \left\langle -\frac{1}{8}, -\frac{1}{8}, \frac{2}{8}\right\rangle\end{aligned}$$

The vector $\left\langle -\frac{1}{8}, -\frac{1}{8}, \frac{2}{8}\right\rangle$ says that the first two types lose $\frac{1}{8}$ of a share each, while the third type gains $\frac{2}{8}$. There is no mutation in this example, so selection accounts for all of the change in population proportions.

How much selection? Without covariance, is there a single number that captures the magnitude of selection?

6. Relative entropy as selection strength

Covariance captures selection strength for numeric traits. The larger the covariance between fitness and radius, the stronger the selection in favour of larger radius. Negative covariance indicates selection against larger radius; i.e. in favour of smaller entities.

Covariance cannot be a general measure of selection strength. There can be selection when the average value of a trait doesn't change (e.g. selection for intermediate radius; only the variance decreases). And there can be selection on traits that are not numeric (e.g. colour).

A general measure of selection strength. With traits z as one-hot vectors, let the population distribution before selection be $Z_1 = \sum_i w_i z_i$ and the population distribution implied by the selection term be $Z_2 = \sum_i w_i' z_i$.

I propose the following measure, defined as a difference of relative entropies measured from Z_1 and Z_2 to a reference distribution Z^*:

$$S_*(Z_1, Z_2) = D(Z^*||Z_1) - D(Z^*||Z_2) \qquad (3)$$

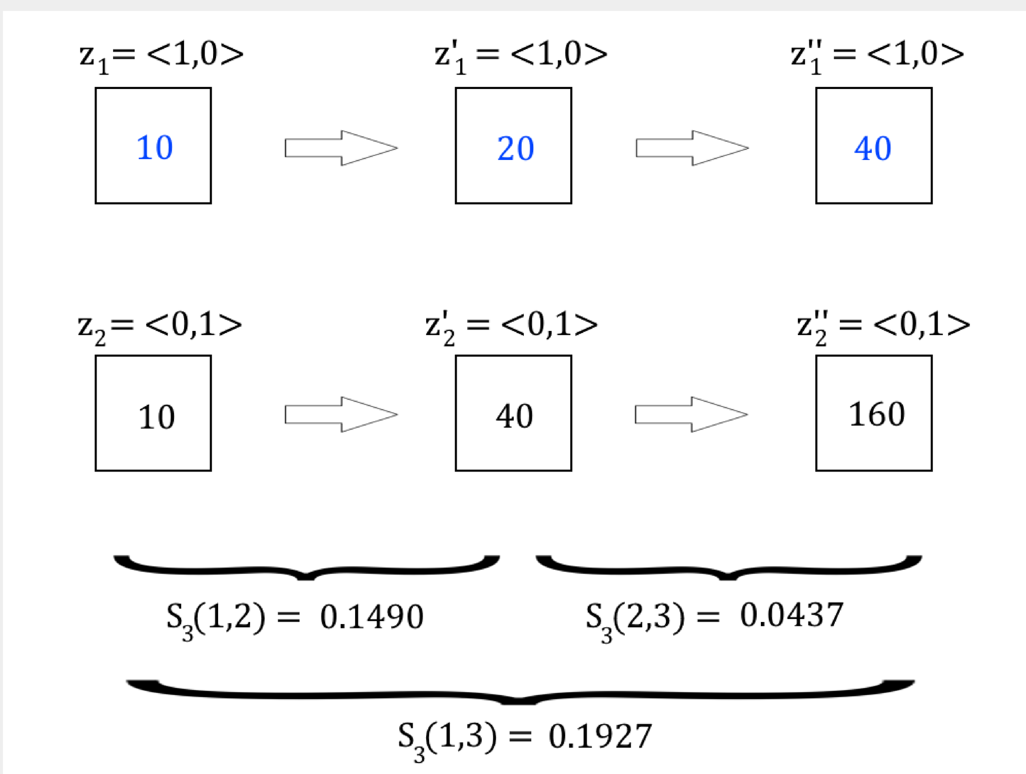
Where relative entropy is defined as $D(Q||P) = \sum_i q_i \log \frac{q_i}{p_i}$.

Useful properties of definition (3) include the following:

- **Additivity** when no mutation: $S_3(Z_1,Z_2) + S_3(Z_2,Z_3) = S_3(Z_1,Z_3)$
- **Simplifies** to $D(Z_2||Z_1)$ when $Z^* = Z_2$
- Applies to **numeric** and **categorical** traits
- Derived from **existing tools** in probability theory
- Related to the notion of **substitution load** in evolutionary biology.

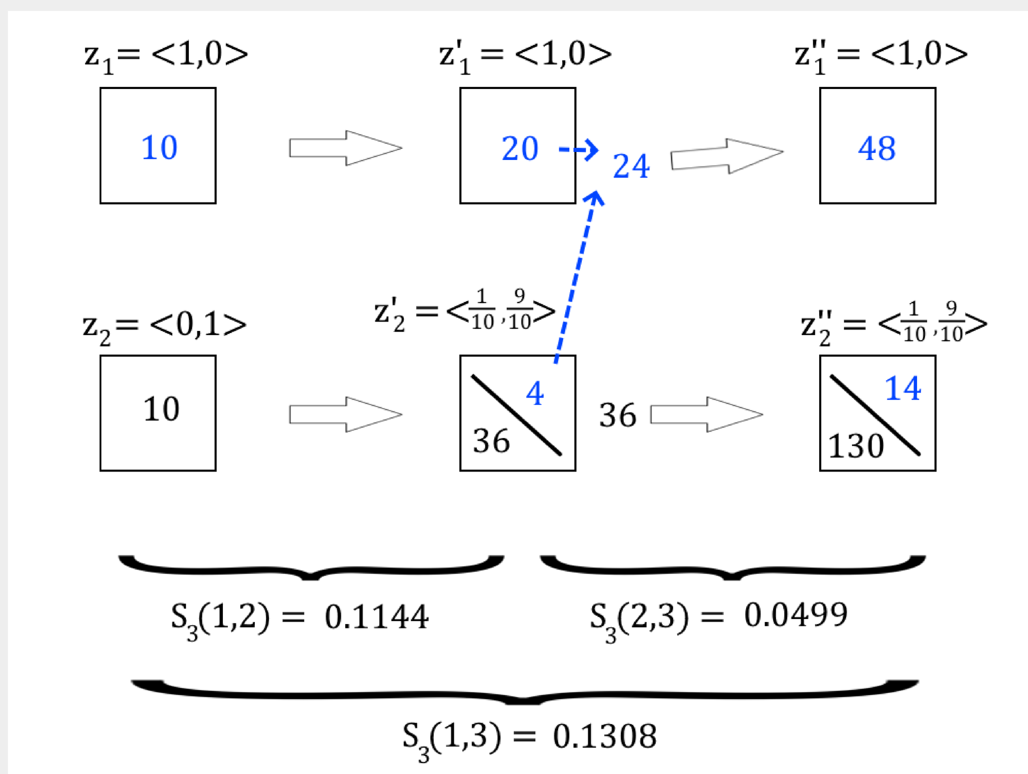
Example: no mutation. Consider a population of two equally distributed types. One doubles in size at every generation, while the other quadruples.

With zero mutation, the overall selection strength is the sum of the strengths at each intermediate step:



Example with mutation. When the more successful type mutates slightly, the overall selection strength is **lower** than the sum of the intermediate strengths.

Selection is being 'held back' by the loss of more fit types:



When selection exactly counterbalances mutation. With enough mutation, $z_2' = \left\langle \frac{1}{4}, \frac{3}{4}\right\rangle$, the two types will remain equally distributed, and the selection strength is the same at every timescale (not pictured).
This captures the fact that selection is occurring, despite the fact that the relative distribution of types remains constant.