Notes on Spark Project

Paulinha

Random networks

Under a Boolean approach, is it possible to find cost and benefit functions that would result in networks resembling the structure of ecological networks? We have a model for understanding transition phases in mutualisms based on increasing costs of interactions, parameterized using empirical networks and sampled cost/benefit functions. One of our goals is to compare the results of the empirical data with the expectations under a null model. Additionally, we ask whether the random structure of a networks can be "pruned" given the structure of costs and benefits generating/resulting in networks that resemble empirical ones.

Cost and benefit functions

Costs and benefits can be describe as functions with a giving underlying distribution.

Boundary conditions for costs and benefits and the persistence/extinction of species

What are the boundary conditions for having all species present ($benefits > costs$) and for having all species extinct of the system (i.e., when $benefits < costs$)? In other words, we want to know what is the probability of having the benefits surpass the costs, resulting in all species present in the network, and alternatively, what is the probability of having all species removed from the network if the costs of interaction surpass their benefits.

Assuming normally and independently distributed costs $C \sim N(\mu_c, \sigma_c^2)$ and benefits $B \sim N(\mu_b, \sigma_b^2)$, we want first to know $P(B > C)$.

$$
\begin{aligned}
P(B > C) &= P(B - C > 0) \\
&= 1 - P(B - C \leq 0)
\end{aligned}
$$

If B and C are independent and normally distributed so is their difference, with the expectation of their mean as $\mathbb{E}(B - C) = \mathbb{E}(B) - \mathbb{E}(C) = \mu_b - \mu_c = \mu$ and of their variance as $Var(B - C) = \sigma_b^2 + \sigma_c^2 = \sigma$. Now we subtract the mean and divide by the variance so that our difference is normally distributed with $\mu = 0$ and $\sigma^2 = 1$) such that $\frac{B-C-\mu}{\sigma} \sim N(0, 1)$. And so:

$$
P(B > C) = 1 - \Phi\left(\frac{-\mu}{\sigma}\right) \tag{1}
$$

where Φ is the cumulative distribution function of $N(0, 1)$.

Now let's consider an example: if species benefits have the distribution $B \sim N(1.2, 2)$ and costs are distributed as $C \sim N(2, 1.5)$, the distribution describing these values is:

Using the derived expression, the resulting probability of having all species present in this system, i.e., the $P(B > C)$, meaning benefits are greater than costs is: 0.3344646. Alternatively, the probability that all species goes extinct in this example is: 0.6655354.

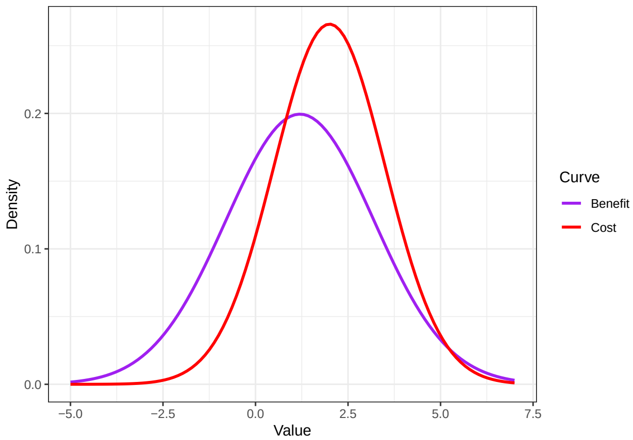

Figure 1: Example of cost-benefit functions following a normal distribution

In more general terms, we can explore how the difference between benefits and costs influences the probability of all species coexisting. Assuming a standard deviation of $\sigma^2 = 1$ we explore the effect of the mean difference between costs and benefits in species' probability of persistence, deriving the probability according to equation 1.

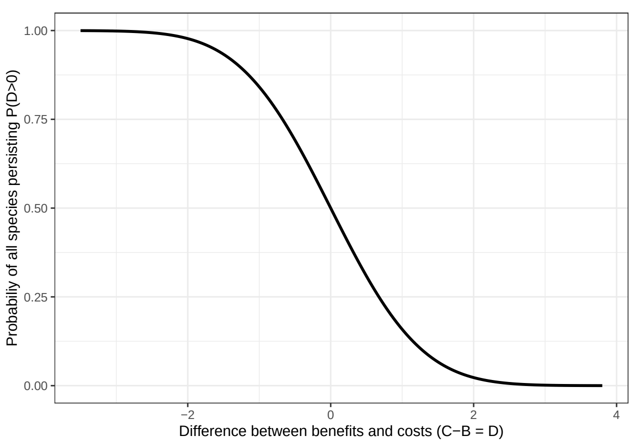

Figure 2: Probability of species' persistence given the difference in benefits and costs

Lognormal distribution

As costs and benefits should, in principle, have the same sign, it is more appropriate to use a distribution defined for positive values. The derivation for the lognormal distribution is not as straight forward as the one for the normal distribution, given we are interested in the difference between two i.i.d. random variables drawn from a lognormal distribution. One solution to find the probability that all species are present is through integration and another solution is proposed by ADD REF

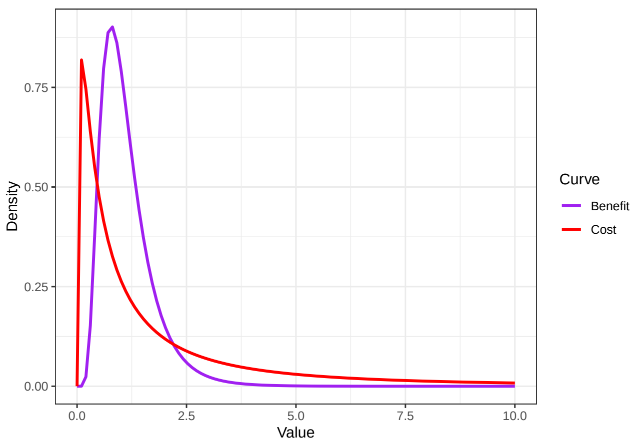

Figure 3: Costs and benefits drawn from a lognormal distribution