Choose the Right Probability Distribution (Ecology, Evolution & Behavioral Biology)

Dr. F.J. Rodenburg, L. Outhuis & L. Bikker

2022-08-08

These exercises assume you are familiar with the basics of ▶ **probability distributions**. That means that you should know when the following distributions are good approximations or not:

- The normal distribution
- The Poisson distribution
- The binomial distribution

For these questions, you have to **identify the outcome** from a short description of a study, and **choose the appropriate probability distributions** that could be used in the analysis.

Q1 — Tree Circumference

Trees come in a large variety of shapes and sizes, depending on the climate and many other factors. To estimate the differences in tree trunk width, researchers measured the circumference of trees from 192 different tree species. What is the outcome here, and with what underlying probability distribution could it be approximated?

Answer:

Q2 — Species Richness

A simple measure for biodiversity is the species richness (the number of different species in a given area).

1. If you want to compare the species richness of urban and rural areas, what probability distribution could you assume in the analysis?
2. Can you think of a downside of using species richness as a measure of biodiversity?

Answer:

Q3 — Millipedes

During the study in **Q2**, you notice millipedes are more common in urban areas. To find out if they also differ physically from millipedes in rural areas, you decide to measure their lengths. To avoid bias, you make sure to capture all millipedes there are, and not just the slow ones. What would be the outcome in this follow-up experiment, and what distribution do you expect these values to take on?

Answer:

Q4 — Shannon Index

The Shannon index is a quantitative measure often used in ecology to describe biodiversity in a specific community. For k different species in a community, it is calculated as follows:

$$\text{Shannon diversity index} = -\sum_{i=1}^{k} p_i \cdot \log(p_i) \tag{1}$$

Here, p_i is the proportion of animals (or plants) belonging to species i.

Suppose you want to compare the Shannon index of 20 urban and 20 rural areas, what probability distribution can you use to approximate the index?

Answer:

HINT: Try running/changing the code below. What values can the index take on?
(Make sure the proportions add up to 1.)

```
# Example of low biodiversity (two species, one underrepresented):
p <- c(0.1, 0.9)
-sum(p * log(p))

# Example of high biodiversity (10 species, all equally represented):
p <- rep(0.1, 10)
-sum(p * log(p))
```

Q5 — Bird Song Frequency

It is common knowledge that in nature, it is usually the male bird that sings. In a study about bird song researchers recorded 150 different songs. Of these, only 10% turned out to be female. If the researchers wanted to demonstrate this underrepresentation of female songs, what probability distribution could be associated with it?

Answer:

Q6 — *In Vivo* Drosophila Evolution

Drosophila melanogaster is a species that is often used to study evolution in the lab. One possible evolutionary experiment, is to count the number of flies that have a particular type of mutation after exposure to hypoxia for 200 generations. Which probability distribution do you think your outcome will take on?

Answer: