

Session 4 - Applications of Probabilistic Programming - Exercises

Consider the `ecology.ipynb` notebook.

An ecologist is doing a survey to estimate the proportion of a population of bats that come from species A and the proportion that comes from species B. The two bat species are identical in appearance apart from their size and weight. Species A is typically smaller, with a mean weight of just 10 grams. The weight of individuals within the species differ, with a standard deviation of 5 grams. Species B is slightly larger and heavier, with a mean weight of 15 grams, and a standard deviation of 2.5 grams.

Over a night time survey, the ecologist traps and weighs 20 bats.

1. Write a probabilistic program to describe this setting.

Hint: Use a Bernoulli random variable that represents whether each individual bat that was caught is from Species A or Species B. The probability that controls this Bernoulli random variable will represent the proportion of bats from Species A and Species B. Then use a likelihood function where whether the bat is from Species A or Species B is used to select the mean and standard deviation of the Normal distribution that describes the measurement process.

2. Use the probabilistic program to estimate the proportion of bats in the overall population that come from Species A.
3. Plot the posterior probability distribution for the proportion of bats in the overall population that come from Species A.
4. The posterior samples of the Bernoulli random variable represents whether each individual bat that was caught was actually from Species A or Species B? Use these samples to calculate the probability that the first bat caught was from Species A.
5. Plot a bar graph representing the probability that each bat was from Species A.