

STATISTICAL RETHINKING 2023

WEEK 6

The theme of this homework is tadpoles. You must keep them alive.

1. Conduct a prior predictive simulation for the Reedfrog model. By this I mean to simulate the prior distribution of tank survival probabilities α_j . Start by using this prior:

$$\begin{aligned}\alpha_j &\sim \text{Normal}(\bar{\alpha}, \sigma) \\ \bar{\alpha} &\sim \text{Normal}(0, 1) \\ \sigma &\sim \text{Exponential}(1)\end{aligned}$$

Be sure to transform the α_j values to the probability scale for plotting and summary. How does increasing the width of the prior on σ change the prior distribution of α_j ? You might try `Exponential(10)` and `Exponential(0.1)` for example.

2. Revisit the Reedfrog survival data, `data(reedfrogs)`. Start with the varying effects model from the book and lecture. Then modify it to estimate the causal effects of the treatment variables `pred` and `size`, including how `size` might modify the effect of predation. An easy approach is to estimate an effect for each combination of `pred` and `size`. Justify your model with a DAG of this experiment.

3. Now estimate the causal effect of `density` on survival. Consider whether `pred` modifies the effect of `density`. There are several good ways to include `density` in your Binomial GLM. You could treat it as a continuous regression variable (possibly standardized). Or you could convert it to an ordered category (with three levels). Compare the σ (tank standard deviation) posterior distribution to σ from your model in Problem 2. How are they different? Why?

4-OPTIONAL CHALLENGE. Return to the Trolley data, `data(Trolley)`, from Chapter 12. Define and fit a varying intercepts model for these data. By this I mean to add an intercept parameter for the individual participants to the linear model. Cluster the varying intercepts on individual participants, as indicated by the unique values in the `id` variable. Include `action`, `intention`, and `contact` as treatment effects of interest. Compare the varying intercepts model and a model that ignores individuals. What is the impact of individual variation in these data?