

STATISTICAL RETHINKING 2022

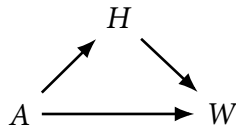
WEEK 2

These solutions are due on Friday 21 January 2022. The optional challenge problem is on the second page.

1. Construct a linear regression of weight as predicted by height, using the adults (age 18 or greater) from the Howell1 dataset. The heights listed below were recorded in the !Kung census, but weights were not recorded for these individuals. Provide predicted weights and 89% compatibility intervals for each of these individuals. That is, fill in the table below, using model-based predictions.

Individual	height	expected weight	89% interval
1	140		
2	160		
3	175		

2. From the Howell1 dataset, consider only the people younger than 13 years old. Estimate the causal association between age and weight. Assume that age influences weight through two paths. First, age influences height, and height influences weight. Second, age directly influences weight through age-related changes in muscle growth and body proportions. All of this implies this causal model (DAG):



Use a linear regression to estimate the **total** (not just direct) causal effect of each year of growth on weight. Be sure to carefully consider the priors. Try using prior predictive simulation to assess what they imply.

3. Now suppose the causal association between age and weight might be different for boys and girls. Use a single linear regression, with a categorical variable for sex, to estimate the total causal effect of age on weight separately for boys and girls. How do girls and boys differ? Provide one or more posterior contrasts as a summary.

4 - OPTIONAL CHALLENGE. The data in `data(Oxboys)` (`rethinking` package) are growth records for 26 boys measured over 9 periods. I want you to model their growth. Specifically, model the increments in growth from one period (`Occasion` in the data table) to the next. Each increment is simply the difference between height in one occasion and height in the previous occasion. Since none of these boys shrunk during the study, all of the growth increments are greater than zero. Estimate the posterior distribution of these increments. Constrain the distribution so it is always positive—it should not be possible for the model to think that boys can shrink from year to year. Finally compute the posterior distribution of the total growth over all 9 occasions.