Homework 6

Due: 2/17 Before Class

Welcome to homework 6!

For this homework, we want you to use the VO2.dat dataset to predict MaxVO2ML (maximum squared milliliters of air in lungs) based on the other covariates in the dataset. The idea is simple here. Find the "best" model.

Finding the "best" model is obviously subjective. We define the "best" model here to excel in both information criterion and interpretability. In identifying this model, use one of the probabilistic programming languages (PPLs), such as JAGS or Stan, that we have been using in class to build the model.

When you have decided on a model that you feel is justified, report all useful convergence diagnostics and information criterion. After deciding on a model, REPLICATE this model in both JAGS and Stan. It is expected that you will have attempted multiple different models, but only report statistics for a single model (using both languages).

For the grading of this assignment, we expect you to explain your reasoning and provide interpretation for the model you choose.

There is added incentive to this assignment. The top 2 students with lowest reported WAIC plus a complexity penalty will each be awarded 2 bonus points to their assignment. The catch is that the WAIC will be penalized for complexity (number of terms). For the penalization, each individual covariate in the model will receive 1 point of penalty, and interactions will receive 2 penalty points. A model has to be simple and effective.