

Introduction to Bayesian Data Analysis

Tutorial 9

- (1) Problem 9.1 (Hoff) Extrapolation: The file `swim.dat` contains data on the amount of time, in seconds, it takes each of four high school swimmers to swim 50 yards. Each swimmer has six times, taken on a biweekly basis.
 - (a) Perform the following data analysis for each swimmer separately:
 - (i) Fit a linear regression model of swimming time as the response and week as the explanatory variable. To formulate your prior, use the information that competitive times for this age group generally range from 22 to 24 seconds.
 - (ii) For each swimmer j , obtain a posterior predictive distribution for Y_j^* , their time if they were to swim two weeks from the last recorded time.
 - (b) The coach of the team has to decide which of the four swimmers will compete in a swimming meet in two weeks. Using your posterior predictive distributions for Y_j^* , make a recommendation to the coach on which swimmer to send to the meet.
- (2) Problem 9.2 (Hoff) Model selection: the file `azdiabetes.dat` contains data on health-related variables of a population of 532 women. In this exercise we will be modelling the conditional distribution of glucose level (`glu`) as a linear combination of the other variables, excluding the variable `diabetes`.
 - (a) Fit a regression model using the g -prior with $g = n$, $\nu_0 = 2$ and $\sigma_0^2 = 1$. Obtain posterior confidence intervals for all of the parameters.
 - (b) Perform the model selection and averaging procedure. Obtain $Pr(\beta_j \neq 0 | \mathbf{y})$, as well as posterior confidence intervals for all of the parameters. Compare to the results in part (a).