STA402L: HOMEWORK 2

DUE: 11:59 PM ON FRIDAY, JANUARY 30

Instructions. Solutions must be submitted to Gradescope as a single PDF. Programming exercises must be completed in R, should be clearly presented, and include all R code. Lab questions are restated here for convenience, but you should refer to the lab itself for details.

Total points. Book exercises: 16; Lab exercises 14; Overall: 30.

BOOK EXERCISES

B1. (3 points) Show that the posterior variance of the beta–binomial model can be written as

$$\mathrm{Var}(\theta \mid y) = \frac{\mathbb{E}(\theta \mid y)\,\mathbb{E}(1-\theta \mid y)}{a+b+n+1}\,.$$

B2. Hoff 3.1.

(a) (1 point)

(b) (1 point)

(c) (1 point)

(d) (1 point)

(e) (2 points)

B3. Hoff 3.3.

(a) (3 points)

(b) (3 points)

(c) (1 point)

LAB EXERCISES

L1. (2 points) Plot a histogram of θ from the `rstan` object called `pool_output`. Describe the distribution.

L2. (2 points) Visualize the posterior distributions of the θ_i with boxplots. In the plot, there should be one box and whiskers object for each θ_i.

L3. (2 points) Take a few minutes to look at the contents of the two files `lab-02-pool.stan` and `lab-02-nopool.stan`. How are they different?

L4. (2 points) What observable quantity does the parameter a represent about our prior beliefs with respect to these data? What does b represent?

L5. (2 points) What do we actually observe in the rat tumor data with respect to these quantities?

L6. (2 points) How well do our different prior beliefs—the ones represented by the different parameter settings above—match up with the data?

L7. (2 points) Why might we have observed such a difference between the two approaches when using the prior $\text{Beta}(1, 1)$?