Assignment 9 (2 pages)

1. Repeat Assignment 8, using Stan instead of JAGS.

2. In Section 3.7 (starting on page 74) and in class, we discussed the analysis of a bioassay experiment. We simulated from the posterior distribution using a uniform prior via the grid method (this code is in Chapter3.R).

   (a) Use JAGS or stan to fit this model to the data, using the following prior distributions (which effectively match the one used in the grid method):

   $$\alpha \sim Uniform(-5, 10)$$

   $$\beta \sim Uniform(-10, 40)$$

   Don't forget to check the model diagnostics and re-simulate if necessary. Obtain at least 10,000 simulations per chain of α and β, summarize each posterior distribution (mean, variance/sd, 95% credible interval) and overlay the simulated values on the contour plot from Chapter3.R to verify that your simulated values seem to be from the 'correct' posterior distribution.

   (b) Add a variable to your model that takes the value 1 if $\beta > 0$ and 0 otherwise. What fraction of your simulated values of β are greater than 0?

   (c) Add the variable $LD50$ to your model, where $LD50$ is defined on page 77. Summarize the posterior distribution of $LD50$. Draw a histogram of the simulated posterior values of $LD50$ and compare it to the histogram in page 77 (they should be very similar).

   (d) Suppose we now wish to give 20 animals a dose level of -0.2. Simulate the posterior predictive distribution of the number deaths in the group of 20 animals. Draw a histogram of this distribution and give a 95% credible interval.

3. In Chapter 5 and in class, we discussed a hierarchical model for rat tumors (the data is on blackboard). We simulated from the posterior distribution using the grid method (the code is in Chapter5.R).

   On page 110, the prior distribution was specified to be uniform on $(\alpha/\alpha+\beta, (\alpha+\beta)^{-1/2})$, and on page 111, we computed using a grid $(\ln(\frac{\alpha}{\beta}), \ln(\alpha+\beta)) \in [-2.3, -1.3] \times [1, 5]$

   Now, if we are only considering values of $\ln(\alpha+\beta)$ in the range $[1, 5]$, then that means we are only considering values of $\alpha+\beta$ in the range $[e^1, e^5]$, and thus we are only considering values of $(\alpha+\beta)^{-1/2}$ in the range $[e^{-5/2}, e^{-1/2}] \approx [0.08, 0.61]$

   Likewise, if we are only considering values of $\ln\left(\frac{\alpha}{\beta}\right) = logit\left(\frac{\alpha}{\alpha+\beta}\right)$ in the range $[-2.3, -1.3]$ then we are only considering values of $\frac{\alpha}{\alpha+\beta}$ in the range $\left[\frac{e^{-2.3}}{1+e^{-2.3}}, \frac{e^{-1.3}}{1+e^{-1.3}}\right] \approx [0.09, 0.22]$

   Therefore we can use the following prior distributions in JAGS or stan to emulate what we did via the grid method:

$$\frac{\alpha}{\alpha+\beta} = U \sim Uniform(0.09, 0.22)$$

$$(\alpha+\beta)^{-1/2} = V \sim Uniform(0.08, 0.61)$$

   (a) Fit the model using JAGS or Stan. Include the variables $lnx = \ln(\alpha/\beta)$ and $lny = \ln(\alpha+\beta)$ in your model so that values of these variables are simulated. Don't forget to check the model diagnostics and re-simulate if necessary. Obtain at least 10000 simulations per chain of α, β, lnx, and lny. How does your answer compare with the answer from the grid method described on page 112? (If you have done it correctly, your answer should be similar.)

   (b) Suppose we plan to conduct a 72nd experiment. Simulate a posterior distribution for θ_{72}. Summarize this distribution and draw a histogram.

   (c) Now suppose that the 72nd experiment will contain 30 rats. Simulate the posterior predictive distribution of the number of rats that will develop tumors, and give a 95% credible interval.