HOMEWORK 10, STAT 251

Remember to submit your R code along with the completed homework.

(1) A study was done to see if taking a fish oil supplement lowered blood pressure. For the seven subjects the change in blood pressure was: 8, 12, 10, 14, 2, 0, 0. The study organizers felt it was reasonable to assume the data were conditionally *iid* and normally distributed. For the prior distribution of μ, we have chosen a Normal distribution with a mean of 6 and a variance of 32. For the prior distribution for σ^2, we have chosen an inverse gamma with shape=3 and rate=50.

  (a) What is the posterior probability that μ is greater than 0?
  (b) What is the posterior probability that $\sigma^2 > 30$?
  (c) Plot the prior and the posterior for μ on the same graph.
  (d) Plot the prior and posterior for σ^2 on the same graph.

(2) Next we will consider problems 3-5 from the previous homework, but relaxing the assumption that the mean is known. As before, here is the total serum cholesterol for 9 urban residents of Guatemala: 197, 199, 214, 217, 222, 223, 227, 228, 234. It is assumed that $Y_{urban,i}|\sigma^2_{urban} \overset{iid}{\sim} N(\mu_{urban}, \sigma^2_{urban})$, where $Y_{urban,i}$ denotes the total serum cholesterol for the i^{th} individual in the sample of Guatemalan urban residents. The prior belief about σ^2_{urban}, the variance of the serum cholesterol of urban residents of Guatemala, is that it has an inverse gamma distribution with shape=2.1 and rate=480. The prior belief about μ_{urban}, the mean serum cholesterol of urban residents of Guatemala, is that it has a normal distribution with a mean of 180 and a variance of $10^2 = 100$.

  (a) Get draws from the posterior distribution, and plot the posterior distribution for the mean on one graph and the posterior distribution for the variance on another graph.
  (b) What is the posterior expected value of σ^2_{urban}?
  (c) What is the posterior expected value of μ_{urban}?
  (d) Now, assume the prior belief about the mean is that it is normally distributed with a mean of 180 and a variance of $20^2 = 400$. (The prior for σ^2_{urban} is still the same.) Again get posterior draws for the mean and variance. This time plot the two posterior distributions for the mean, and include them on the same graph (the posterior for μ_{urban} from part (a) in gray, and the posterior for μ_{urban} from this part in black).[1]

(3) Here is the total serum cholesterol for 10 rural residents of Guatemala: 139, 142, 143, 144, 145, 148, 155, 162, 171, 181. The response variable (total serum cholesterol) is assumed to be conditionally *iid* and normally distributed. The prior belief about the serum cholesterol of rural residents of Guatemala is that the population mean should have a mean of 150 with a variance of 500. The prior for the variance is an inverse gamma with shape=3 and rate=180. Plot the posterior for the mean.

(4) Now we consider comparing μ_{urban} to μ_{rural} and σ^2_{urban} to σ^2_{rural}. Note the the former is akin to the independent two-sample t-test to comparing means from two populations which is regularly taught in stat 121. The former is not taught in stat 121 because the sampling distribution is beyond the scope of that course. In Bayes, inference for comparing σ^2_{urban} to σ^2_{rural} follows exactly as before!

  (a) Plot the posterior distribution of the difference in the means of the two groups: $\mu_{urban} - \mu_{rural}$. What is the 95% (posterior) credible interval for this difference? What can you conclude?
  (b) Plot the posterior distribution of the ratio of variances (i.e., of $\frac{\sigma^2_{urban}}{\sigma^2_{rural}}$). What is the 95% posterior credible interval for this ratio? What can you conclude?

(5) Plot the posterior predictive distribution of Y_{urban} and the posterior predictive distribution of Y_{rural} on the same graph. (This of course requires a legend be included in your plot to distinguish one from the other.). Obtain an interval for difference in prediction of the urban group relative to the rural. How does this interval compare to that one found in part 4 (a)?

[1] This allows us to gauge the sensitivity of our posterior inference on μ to the variance of the prior beliefs of μ_{urban}.