22S:138, Bayesian Statistics
Fall 2006, Homework 3

Due: Mon. 9/18

1. Bolstad textbook: problems 9.2 and 9.5, p. 164-165
2. Comment on whether your results in part 4 of last week's homework are robust to prior specifications.
3. (problem from Johnson and Albert) In 1986, the St. Louis Post Dispatch was interested in measuring public support for the construction of an indoor stadium. The newspaper conducted a survey in which they interviewed 301 registered voters. In ths sample, 135 voters opposed the construction of a new stadium. Let p denote the proportion of all registered voters in the St. Louis voting district opposed to the stadium.

   Suppose a city councilman wants to increase the city sales tax to pay for the construction of the new stadium. She claims that under half of the registered voters are opposed to stadium construction. She would like to use the sample survey data of the newspaper to test the two hypotheses:

$$H_0: \; p \; \geq \; .5$$
$$H_a: \; p \; < \; .5$$

   (a) A frequentist method of testing these hypotheses is based on the p-value. The p-value is the probability of observing the sample result obtained, or something more extreme, if indeed exactly half of the registered voters in St. Louis were opposed to construction; that is,

$$p - value = Pr(y \leq 135 | p = .5)$$

   where y is a binomial random variable with sample size $n = 301$ and success probability $p = 0.5$. Compute the p-value for this example (use of an R function will make this easy). If this probability is small, then one concludes that there is significant evidence in support of hypothesis $H_a: \; p \; < \; .5$.

   (b) Now consider a Bayesian approach to testing these hypotheses. Suppose that a uniform prior is assigned to p. Find the posterior distribution of p and use it to compute the posterior probabilities of H_0 and H_1.

4. Referring to the previous problem, again suppose that a uniform prior is placed on the proportion p, and that from a random sample of 301 voters, 135 oppose construction of a new stadium. Also suppose that the newspaper plans on taking a new survey of 20 voters. Let y^* denote the number in this new sample who oppose construction.

   (a) Find the posterior predictive probability that $y^* = 8$. (Use one of the R functions that I wrote and that were introduced in lab.)

   (b) Find the 90% prediction interval for y^*. Do this by finding the predictive probabilities for each of the possible values of y^* and ordering them from largest probability to smallest. Then keep adding the most probable values y^* into the probability set until the total probability exceeds .90 for the first time.

5. (REQUIRED for stats and biostats grad students; OPTIONAL for undergrads and students in other majors.) Referring to the previous problem, suppose more generally that the newspaper plans on taking a new survey of n^* voters. Find the formula for the posterior predictive probability of y^* successes in a sample of size n^* (still conditioning on the uniform prior and the results of the first survey). [Hint: You will need to compute the integral of the kernel of a beta density. Look at the expression for the normalized beta density, and recall that it has to integrate to 1. The predictive probability that you need can be expressed as a ratio involving some functions.]