22S:138, Bayesian Statistics
Fall 2005, Homework 2

Due: Mon. 9/12

Suppose you are the coach of a college baseball team. A new player is joining your team. As a high school player, he had 40 hits in 120 at-bats. You wish to estimate θ, his underlying true probability of getting a hit in any at-bat as a college-level player.

1. Specify a Beta prior that seems appropriate to capture your knowledge or uncertainty about θ before he plays in any college-level games. Use any information you have that seems important – his high-school record, what you know about college- level ball players, etc. You may use R/Splus functions if you wish. Explain in a few sentences how you chose the values of α and β. There is no one right answer here – I want to see how you think about this and what procedure you use.

2. Specify a Beta prior that you think might reflect the player's father's beliefs about θ.

3. Suppose the player now plays eight college-level games, has thirty at-bats, and gets 5 hits. Thus the data are:

$$y = 5, \quad n = 30$$

   We will use a binomial likelihood for these data. This requires the assumption that, conditional on θ, each at-bat is an independent Bernoulli trial with success probability θ. There are several reasons why independence might actually *not* be a reasonable assumption in this problem. Give one.

   Note: For our present purposes, we'll use a binomial likelihood anyway. We'll come up with a better model when we talk about hierarchical models later in the semester.

4. Obtain the following characteristics of the posterior distribution $p(\theta|y)$

   (a) name of posterior distribution and its parameter values

   (b) posterior density plot (you may use either the `plot.beta` or the `plot.betaupdate` R function if you wish)

   (c) posterior mean and mode

   (d) 95% posterior interval for θ

   (e) posterior probability that $\theta > .25$

   based on the data from the college-level games under each of three priors:

   (a) your prior from part 1.

   (b) the father's prior

   (c) a uniform prior

5. Comment on whether or not the conclusions in part 4 are robust to different prior specifications.