22S:138, Bayesian Statistics
Fall 2008, Homework 1

Due: Mon., 9/8 in class

1. Refer to section 1.4 (pp. 9-11) in Gelman *et al.* The question of interest is whether or not the woman is a carrier of the hemophilia gene.

   (a) Use the information given in the section to determine

   - prior probabilities for the 2 possible "models": $\theta = 1$ (carrier) and $\theta = 0$ (not carrier)
   - two different sets of likelihood probabilities

     Using the notation from the section, for the ith son born to the woman, $y_i = 0$ indicates the son is not affected by hemophilia; $y_i = 1$ indicates that he is affected. For each model, you will need:

     $$Pr(y_i = 0|model)$$

     and

     $$Pr(y_i = 1|model)$$

   (b) Suppose the woman has 3 sons. None of them are identical twins, and we will consider their hemophilia outcomes to be independent conditional on her carrier status. The 3 outcomes are (this is different from in the textbook example):

   $$y_1 = 0$$
   $$y_2 = 1$$
   $$y_3 = 0$$

   Do a sequential Bayesian analysis in which you compute the posterior probability that the woman is a carrier using the data from each son one-at-a-time. Either use the "bayes" function that we used in lab, or write out the required calculations by hand. Make a table with columns for model, prior, likelihood, product, and posterior. Turn in printouts of your R output, or handwritten tables.

   (c) Also answer the following questions:

   - What was the posterior probability that the woman was a carrier after the first son's status became known?
   - Did the posterior probability change based on the data from the second son? Why or why not?
   - Did the posterior probability change based on the data from the third son? Why or why not?

2. (Question taken from Berry.) Suppose the following statements are true for Iowa City:

   (a) The probability that it rains both today and tomorrow is 0.2.
   (b) The probability that it rains today but not tomorrow is 0.1.
   (c) The probability that it rains tomorrow but not today is 0.1.
   (d) The probability that it rains neither today nor tomorrow is 0.6.

   Find the probability that:

   (a) It will rain today.
   (b) It will rain at some time during the time spanning today and tomorrow.
   (c) It will rain tomorrow.
   (d) It rains tomorrow, given that it rains today.

3. Do question 6 on p. 31 of Gelman *et al.* Note: if you do not know what "fraternal" and "identical" twins are, look it up or ask someone.

4. Consider event B: The current population of the state of South Carolina is greater than the current population of Monaco. Draft someone and use the "balls in an urn" calibration experiment (as was done in class on 8/27) to help that person assess his or her $P(B)$. Report this probability to the nearest $\frac{1}{8}$, for example, by concluding that $\frac{3}{8} \leq P(B) \leq \frac{1}{2}$. Report who the person is. Give all your questions and his or her answers. If there are extenuating circumstances that affect your answer, then please describe them. For example, the person might have left the room during the process and — you suspect — consulted an encyclopedia.

   Note: If your partner is another student in the class, then one of you should use event B as above and the other instead should consider event C: The average weight of adult male emperor penguins is greater than 60 pounds.