Homework 1 (due date: Sep.19th before the class)

Note: this homework should be typed in Latex or any other similar tools that generate pdf file. Please turn in your homework with R code file to the TA at `yli193@jhu.edu`. You could print it out and turn in the paper version before the class if you would like.

Problem 1

Suppose that 6 observations are taken at random from a uniform distribution on the interval $(\theta - 1/2, \theta + 1/2)$, with θ unknown, and that their values are 11.0, 11.5, 11.7, 11.1, 11.4 and 10.9. Suppose that the prior distribution of θ is a uniform distribution on the interval (10,20). Determine the posterior distribution of θ.

Problem 2

Suppose that the time in minutes required to serve a customer at a certain facility has an exponential distribution for which the value of the parameter θ is unknown and that the prior distribution of θ is a gamma distribution for which the mean is 0.2 and the standard deviation is 1. If the average time required to serve a random sample of 20 customers is observed to be 3.8 minutes, what is the posterior distribution of θ?

Problem 3

Show that the family of beta distributions is a conjugate family of prior distributions for samples from a negative binomial distribution with a known value of the parameter r and an unknown value of the parameter p, with $0 < p < 1$.

Problem 4

Sample survey: Suppose we are going to sample 100 individuals from a county (of size much larger than 100) and ask each sampled person whether they support policy Z or not. Let $Y_i = 1$ if person i in the sample supports the policy, and $Y_i = 0$ otherwise.

1. Assume $Y_1, \ldots, Y_{100}$ are, conditional on θ, i.i.d. binary random variables with expectation θ. Write down the joint distribution of $\Pr(Y_1 = y_1, \ldots, Y_{100} = y_{100} \mid \theta)$ in a compact form. Also write down the form of $\Pr(\sum Y_i = y \mid \theta)$.

2. For the moment, suppose you believed that $\theta \in \{0.0, 0.1, ..., 0.9, 1.0\}$. Given that the results of the survey were $\sum_{i=1}^{100} Y_i = 57$, compute $\Pr(\sum_{i=1}^{100} Y_i = 57)$ for each of these 11 values of θ and plot these probabilities as a function of θ.

3. Now suppose you originally had no prior information to believe one of these θ-values over another, and so $\Pr(\theta = 0.0) = Pr(\theta = 0.1) = \ldots = Pr(\theta = 0.9) = Pr(\theta = 1.0)$. Use Bayes' rule to compute $p(\theta \mid \sum_{i=1}^{100} Y_i = 57)$ for each θ-value. Make a plot of this posterior distribution as a function of θ.

4. Now suppose you allow θ to be any value in the interval [0, 1]. Using the uniform prior density for θ, so that $p(\theta) = 1$, plot the posterior density $p(\theta) \times \Pr(\sum_{i=1}^{100} Y_i = 57 \mid \theta)$ as a function of θ.

5. As discussed in this chapter, the posterior distribution of is beta(1+ 57, 1 + 100-57). Plot the posterior density as a function of θ. Discuss the relationships among all of the plots you have made for this exercise.

Problem 5

Sensitivity analysis: It is sometimes useful to express the parameters a and b in a beta distribution in terms of $\theta_0 = a/(a+b)$ and $n_0 = a+b$, so that $a = \theta_0 n_0$ and $b = (1-\theta_0)n_0$. Reconsidering the sample survey data in Problem 4, for each combination of $\theta_0 \in \{0.1, 0.2, \ldots, 0.9\}$ and $n_0 \in \{1, 2, 8, 16, 32\}$ find the corresponding a, b values and compute $\Pr(\theta > 0.5 \mid \sum Y_i = 57)$ using a beta(a, b) prior distribution for θ. Display the results with a contour plot, and discuss how the plot could be used to explain to someone whether or not they should believe that $\theta > 0.5$, based on the data that $\sum_{i=1}^{100} Y_i = 57$.