Gov 2001: Problem Set 3
Assessment Problem

Due Wednesday, February 24 by 6pm

Instructions

You should submit your answers and R code to the problems below using the Quizzes section on Canvas.
Remember that you should treat this assessment problem as you would treat a final exam. You are not allowed to discuss the problem with classmates, the teaching staff, or any other people. You also may not post questions about the assessment problem to the Canvas discussion boards.
You may consult any readings, notes, or R code from the class, and you can also use the internet as a resource, but remember that all answers and code must be entirely your own.
Also don't forget that you only have one opportunity to submit your answers to the assessment problem on Canvas. Be sure you check your work before you click the submit button.
Remember that the purpose of these assessment problems is for you to get an honest sense of how well you understand the content from the class. If you struggle with an assessment problem one week, you should spend the following week being sure that you catch up on the topics you didn't understand.

Problem 1

You are analyzing the way that Senator Byrd (D-WV) voted on education in 1999. The way this works is as follows – Senator Byrd either voted *Yea* on a bill or he voted *Nay* on a bill. There are no other possible outcomes. Your data (reported by the National Education Association) show that out of 12 bills pertaining to education, Senator Byrd voted *Yea* on 9 of them. That is, $y = 9$.

For the following parts, let θ represent the *Yea* rate (or 'success' rate) of the 12 votes in your dataset.

Likelihood Inference

1.A) What is the likelihood of θ given y and n, $L(\theta|y, n)$?

1.B) Plot the likelihood using R. On the x-axis you should have θ and on the y-axis you should have the likelihood.

1.C) Using your plot from part B, provide a rough estimate of the maximum likelihood estimate. (Note: There is no need to use calculus here; an eye-ball estimate is fine.)

1.D) Now, assume that your estimate from Part C is correct (that is, the true θ equals your estimate). What is the probability of seeing 9 "Yea" votes out of 12 total votes if this is correct?

Bayesian Inference

Assume for the rest of this problem that you want to introduce some prior information. You believe that the prior distribution of θ is proportional to:

$$p(\theta) \propto (1-\theta)^{\beta-1}$$

Assume β is a known constant.

1.E Find $p(\theta|y, \beta, n)$, the posterior distribution of θ, up to some multiplicative proportionality constant.

1.G Write a function in R for this posterior distribution. Your function should take as its inputs θ, β, y and n. Your function can give outputs proportional to the posterior distribution (i.e. without a normalizing constant). Using the function, the data you have collected, and assuming that $\beta = .5$, plot the posterior density. On the x-axis you should have θ and on the y-axis you should have the posterior density.

1.H What is the probability that θ is between 0.6 and 0.65 given the data, the prior distribution, and $\beta = .5$? Keep in mind that the posterior function you wrote probably is proportional (not equal) to $P(\theta|y, \beta, n)$, so to get the probability you will need to divide by the normalizing constant, that is, the integral of the function across the whole support of θ.

R code

Please submit all your code for this assessment problem as a .R file. Your code should be clean, commented, and executable without error.