HW6

Neal Marquez

February 24th, 2019

In order to investigate presidential preferences in Seattle before the Democratic caucuses in the 2016 Presidential election, a pollster interviewed likely caucus goers in each of several neighborhoods, chosen at random. The two candidates in contention were Clinton and Sanders. On Capitol Hill, 19 of 91 respondents favored Clinton, while 72 favored Sanders.

Question 1

Find the posterior distribution of the true proportion of likely caucus goers actually favoring Sanders analytically (i.e. not by simulation), using a uniform prior. Find the posterior mode, posterior mean, posterior median, and posterior 95% confidence interval. Plot the posterior density. Compare the results with the results from a standard frequentist analysis.

We have uniform prior distribution across all probability of a caucus goers favoring Sanders by using a $\text{Beta}(1,1)$ prior. The benefit of using a Beta distribution as our prior is that it leads to a Beta distribution as the prior. We can show this in the generic case of a prior with values $\text{Beta}(\alpha, \beta)$. Recall that the likelihood for the Bernoulli distribution, our family used for the Bernoulli trails of selecting Sanders over Clinton, is as follows

$$p(\boldsymbol{x}|\theta) = \theta^k (1-\theta)^{n-k}$$

where k is the number of trials that are in favor of Sanders and n is the total number of trials. A prior of $\theta \sim \text{Beta}(\alpha, \beta)$ will then give us.

$$p(\theta) = B(\alpha, \beta)^{-1} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

Using bayes rule we know $p(\theta|x) \propto p(\boldsymbol{x}|\theta)p(\theta)$ which leaves us with the following

$$p(\theta|\boldsymbol{x}) \propto B(\alpha, \beta)^{-1} \theta^{k+\alpha-1} (1-\theta)^{n-k+\beta-1}$$

This value is proportional to the following

$$p(\theta|\boldsymbol{x}) \propto \text{Beta}(\alpha+k, \beta+n-k)$$

Again using a $\text{Beta}(1,1)$ prior we find that the posterior then is

$$p(\theta|\boldsymbol{x}) \propto \text{Beta}(73, 20)$$

The mean, median, and mode for the Beta distribution is well defined and is as follows

$$\text{Mean} = \frac{\alpha}{\alpha+\beta}$$

$$\text{Median} = \frac{\alpha - \frac{1}{3}}{\alpha + \beta - \frac{2}{3}}$$

$$\text{Mode} = \frac{\alpha - 1}{\alpha + \beta - 2}$$

This leaves us with the following estimates.

| Statistic	Value
Mean	0.785
Median	0.787
Mode	0.791

We can show hos this estimate compares to the frequentist approach by calculating the maximum likelihood estimate. We plot the posterior and the MLE, $\frac{k}{n}$, in the plot below and show that the estimates are nearly identical. Indeed, when we use a uniform prior the posterior shape is strongly influenced by the likelihood.

Posterior Distribution (MLE at dotted line).
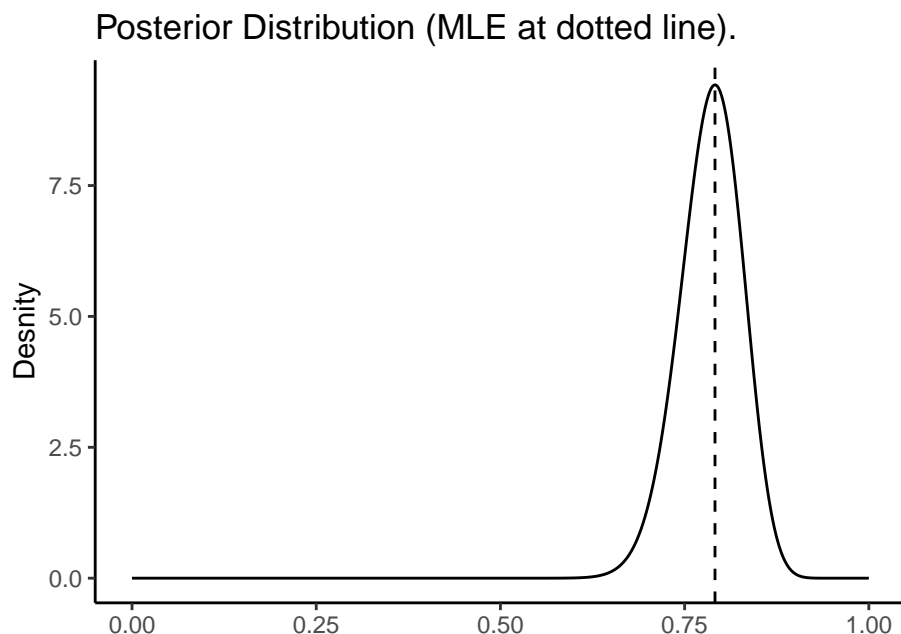
Question 2

*Do the same calculations by Monte Carlo and compare the results.*

We can create a set of Monte Carlo results by simulating from the posterior distribution. This exercise is helpful as it shows that even when we do not know the closed form of the distribution of the posterior we may simulate from it and get accurate results here we take 10000 Monte Carlo simulations from the posterior and show how close it resemble the true posterior from question 1. The plot below shows the estimated density from samples of the Monte Carlo simulation compared against the true posterior density.

Posterior Analytical vs Simulation Computation
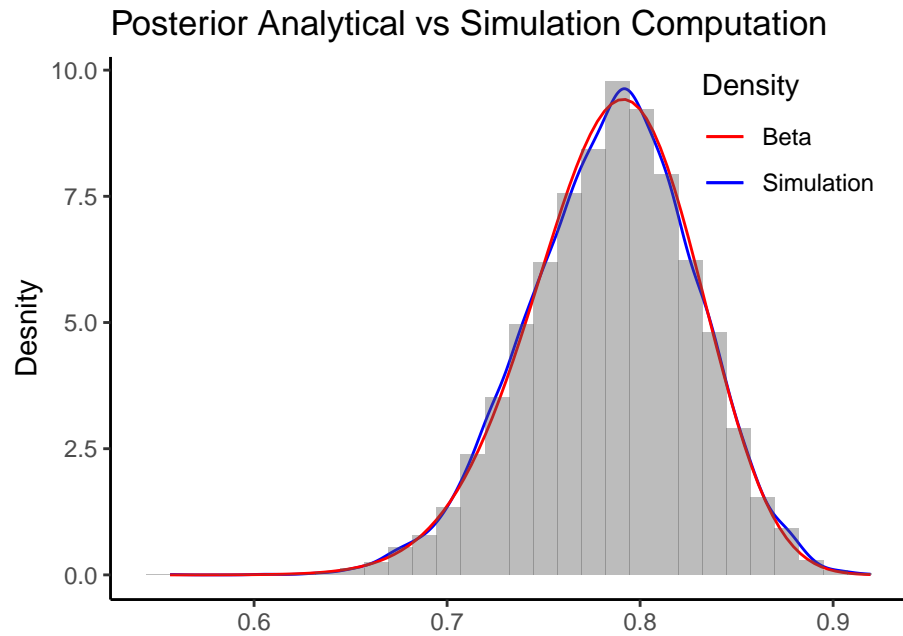
Question 3

Find and plot the posterior distribution of the logit of the proportion favoring Sanders.

Another useful property of estimating the posterior distribution by simulation is that functions of the parameter of interest are easy to calculate with their uncertainty simply by applying the function to each of the simulations. Below we plot the estimated density of the logit of the probability of favoring sanders which supports a range from $(-\infty, \infty)$ rather than $(0, 1)$.

## Posterior Logit Probability of Sanders Support
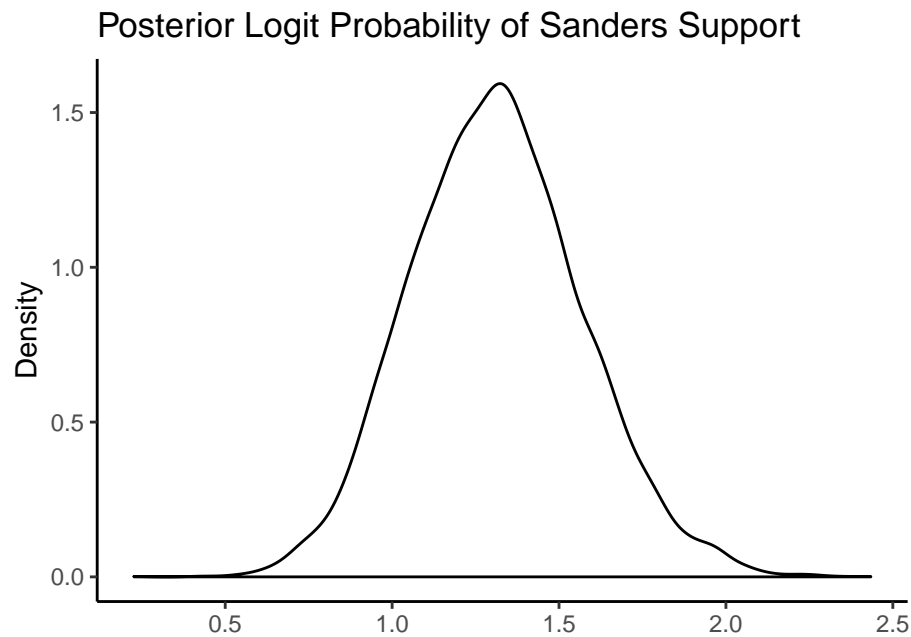
Question 4

Suppose it is known that it is rare for any candidate to get more than 80% of the votes in a two-person caucus. Specify a prior distribution to represent this information. Redo question 1 with this prior. Comment on the differences between the results with the two prior distributions.

We can specify a prior where the 95% confidence interval covers a range from .2 to .8 and is centered at .5. Finding this distribution is trivial because in order for the mean of any Beta distribution to be .5 α must equal β. This is convenient as it also gives us a symmetrical distribution as skewness is 0 when α equals β. Finding the correct value of the parameters is simply a matter of solving a single equation of the cumulative density function (CDF) where the $F(x)$ where $x = .2$ is .025.

The CDF for a beta distribution is slightly complicated, however, we can use R to help us find the solution to this problem. The code below and the following plot shows how we calculate the correct parameters for the prior and the CDF of the resulting distribution.

```
minFunc <- function(x) ((.025 - pbeta(.2, x, x))^2)^.5
optPar <- nlminb(0, minFunc, lower=.0001)
priorPar <- optPar$par
```

Prior Distribution as Derived from R optimizer
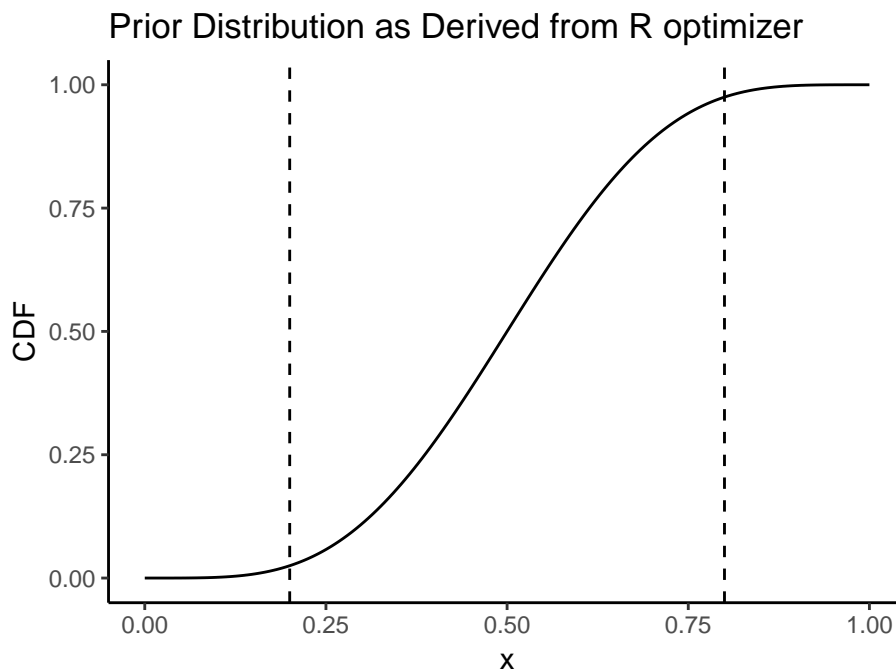
We may now recalculate the posterior using this prior with $\alpha = \beta = 4.54$. The resulting posterior distribution is then a Beta distriobution with α of 76.54 and β of 23.54. Below we plot the results of the two priors which we label uninformative and informative respectively. Notice that despite the difference in the priors the resulting posterior distribution has a very similar shape and placement. This is because of the amount of data that we have informing our estimate.

Distribution Comparison