

ST440/550 – Assignment 2 – Due 1/29

In this assignment you will write an *R* function to perform a Bayesian analysis of count data. The response is the number of successes in N independent trials, $Y \in \{0, 1, 2, \dots, N\}$, and the unknown parameter is the success probability, $\theta \in [0, 1]$. Assume the Bayesian model with likelihood $Y|\theta \sim \text{Binomial}(N, \theta)$, prior $\theta \sim \text{Beta}(a, b)$, and thus posterior $\theta|Y \sim \text{Beta}(Y + a, N - Y + b)$.

- (1) Write an *R* function that takes Y , N , a , and b as inputs. The function should produce a plot (clearly labeled!) that overlays the prior and posterior density functions (both using the `dbeta` function), and it should return a list with the posterior mean and posterior standard deviation.
- (2) What values of a and b would make good default values to represent a prior that carries little information about θ ? Make these the default values in your function.
- (3) What values of a and b give prior mean 0.7 and prior standard deviation 0.2?
- (4) Now we observe $Y = 20$ events in $N = 30$ trials. Use your code from (1) to conduct a Bayesian analysis of these data. Perform the analysis twice, once with the uninformative prior from (2) and once with the informative prior in (3).
- (5) Summarize the results. In particular, how does this analysis compare to a frequentist analysis and how much are the results affected by the prior?

You should turn in your responses to these questions in 1-2 pages (i.e., one piece of paper with text on both sides). Be sure all plots are labeled and code is commented!