STAT 431 — Applied Bayesian Analysis — Fall 2022

Homework 2

Due: September 23, 11:59 PM (US Central)

In your submission, please include computer code and output that you used to answer the following problems. All numerical values should be given to at least three significant digits, unless otherwise stated.

1. A telemarketing salesperson makes a series of telephone calls, some of which (successful calls) result in a sale, while others (unsuccessful calls) do not. The salesperson must fulfill a daily quota of 5 successful calls. On a given day, the number Y of **un**successful calls made before this quota is fulfilled follows this particular negative binomial distribution:

$$Y \mid \theta \quad \sim \quad \text{NegBinomial}(\theta, m = 5)$$

in the notation of BSM, Appendix A.1. Here, $0 < \theta < 1$, but θ is otherwise unknown.

   (a) [1 pt] Considering this scenario, describe briefly in words what θ represents, in practical terms.

   (b) [1 pt] If it were true that $\theta = 0.2$, what would be the **expected total** number of calls the salesperson would need to make (including both the successful and unsuccessful calls) to fulfill the daily quota?

   (c) [2 pts] Yesterday, the salesperson made $y = 20$ unsuccessful calls before fulfilling the quota. Give an expression for the likelihood function of θ (up to proportionality) based on y.

   (d) [2 pts] Assume a uniform ("flat") prior for θ. Using the likelihood of the previous part, derive an expression for the posterior density (up to proportionality).

   (e) [2 pts] Identify the posterior distribution corresponding to the density you derived in the previous part. (Name it and specify the values of its constants.)

2. Consider the same scenario as in the previous problem. Based on the same model for the $y = 20$ unsuccessful calls observed yesterday, a Bayesian data analyst determines the following posterior for θ:

$$\theta \mid y = 20 \quad \sim \quad \text{Beta}(\alpha = 6, \beta = 25)$$

   (a) [1 pt] What specific prior did the analyst use? (Specify either an expression proportional to its density, or a distribution's name and constants.)

   (b) [2 pts] Based on this posterior, approximate the posterior mean and posterior standard deviation of θ. (Retain at least three significant digits.)

   (c) [2 pts] Based on this posterior, approximate a 95% equal-tailed (Bayesian posterior) credible interval for θ. (Retain at least three significant digits.)

   (d) [1 pt] Based on this posterior, approximate the posterior probability that $\theta \leq 0.2$. (Retain at least three significant digits.)

(e) [1 pt] Based on this posterior, approximate the posterior **predictive** probability that, compared to yesterday, the salesperson will need to make **strictly fewer** calls today to fulfill today's quota of 5 successful calls. (You may assume that the numbers of unsuccessful calls today and yesterday are conditionally independent. Retain at least three significant digits in your numerical answer.)

3. GRADUATE SECTION ONLY

Consider data y from a binomial distribution with known n and unknown success probability θ. Assume a **beta** prior distribution on θ, with constants $\alpha > 0$ and $\beta > 0$. (Distributions are as defined in BSM, Appendix A.1.)

In this problem, you will show that the posterior standard deviation becomes equivalent (in a frequentist sense) to the frequentist standard error as sample size n increases:

(a) [2 pts] Find a formula for the posterior *standard deviation* $\sqrt{\text{Var}(\theta \mid y)}$ (in terms of y, n, α, and β).

(b) [3 pts] Now adopt the frequentist perspective in which θ is fixed ($0 < \theta < 1$) and the data Y is random.

Recall that the usual (frequentist) standard error estimator for $\hat{\theta} = Y/n$ is

$$\text{SE}(\hat{\theta}) \;=\; \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$

Using the fact that $\hat{\theta} \to \theta$ as $n \to \infty$ (w.p. 1, by the strong law of large numbers), show that

$$\frac{\sqrt{\text{Var}(\theta \mid Y)}}{\text{SE}(\hat{\theta})} \;\longrightarrow\; 1 \qquad \text{as } n \to \infty$$

for any $0 < \theta < 1$. (Remember, α and β remain fixed.)