STATS 551 Homework 1

Introduction & Single Parameter Models

Responsible GSI: Wayne Wang

Practice of Bayes Formula (3×10 points). Suppose that if $\theta = 1$, then y has a normal distribution with mean 1 and standard deviation σ, and if $\theta = 2$, then y has a normal distribution with mean 2 and standard deviation σ. Also, suppose $\Pr(\theta = 1) = 0.5$ and $\Pr(\theta = 2) = 0.5$.

1. For $\sigma = 2$, derive the formula for the marginal probability density for y, p(y), and sketch/visualize it in R.
2. What is $Pr(\theta = 1|y = 1)$, again supposing $\sigma = 2$?
3. Describe how the posterior density of θ changes in shape as σ is increased and as it is decreased.

Guideline for Submission: submit a hard copy (handwritten or printed).

Normal distribution with unknown mean (3×10 points). A random sample of n students is drawn from a large population, and their weights are measured. The average weight of the n sampled students is $\bar{y} = 150$ pounds. Assume the weights in the population are normally distributed with unknown mean θ and known standard deviation 20 pounds. Suppose your prior distribution for θ is normal with mean 180 and standard deviation 40.

1. Give your posterior distribution for θ. (Your answer will be a function of n.)
2. A new student is sampled at random from the same population and has a weight of $\tilde{y}$ pounds. Give a posterior predictive distribution for $\tilde{y}$. (Your answer will still be a function of n.)

3. For $n = 10$, give a 95% posterior interval for θ and a 95% posterior predictive interval for $\tilde{y}$. Do the same for $n = 100$.

Guideline for Submission: submit a hard copy (handwritten or printed).

Nonconjugate single parameter model (2×20 points). Suppose you observe $y = 285$ from the model Binomial$(500, \theta)$, where θ is an unknown parameter. Assume the prior on θ has the following form

$$p(\theta) = \begin{cases} 8\theta, & 0 \le \theta < 0.25 \\ \frac{8}{3} - \frac{8}{3}\theta, & 0.25 \le \theta \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

1. Compute the unnormalized posterior density function on a grid of m points for some large integer m. Using the grid approximation, compute and plot the normalized posterior density function $p(\theta|y)$, as a function of θ.

2. Sample 10000 draws of θ from the posterior density and plot a histogram of the draws.

Guideline for Submission: submit R markdown (or jupyter notebook) with annotated code followed by results. Discussions about the results should follow the results.

Optional Reading. Read one of the following papers and post your summary and thoughts on Canvas.

1. Efron, B. (2005). Bayesians, frequentists, and scientists. Journal of the American Statistical Association, 100(469), 1-5.

2. Biostatistics and Bayes, Norman Breslow, Statist. Sci., Volume 5, Number 3 (1990), 269-284.

3. Bayesian Methods in Practice: Experiences in the Pharmaceutical Industry, A. Racine, A. P. Grieve, H. Fluhler and A. F. M. Smith, Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 35, No. 2 (1986), pp. 93-150.