Homework 4 for STA 602 – Fall 2022

Due at 12:00pm on September 30, 2022 (Show your work!)

1. PH 3.9

2. This problem continues on Problem 5 in HW3. Recall that for the customer service call example we have seen in class, we model the number of calls received on n different days $X_1, X_2, \ldots, X_n$ as independent Poisson(θ) random variables given the unknown mean θ. Suppose that data were collected on 10 different days, during which a total of 200 calls were received. Also suppose that you adopt a Gamma(α, β) prior with $\alpha = 10$ and $\beta = 1$.

   (a) What are the bias and variance of the Bayes estimator, which you found in HW3, under squared error loss? What is its MSE?

   (b) Make a plot of the above MSE and along with that for the maximum likelihood estimator $\bar{X}$. Comment on for what ranges of θ one is better than the other and give the reason behind that.

3. In this problem, recall the four estimators for θ in the political poll example we saw in the "Lecture 6" notes:

$$\delta_1(X) = \frac{X}{n}, \quad \delta_2(X) = \frac{1}{2}, \quad \delta_3(X) = \frac{X+12}{n+24}, \quad \text{and} \quad \delta_4(X) = \frac{X+1}{n+2}.$$

   Also consider a fifth estimator

$$\delta_5(X) = \frac{X + \sqrt{n}/2}{n + \sqrt{n}}$$

   (a) Calculate the bias, variance, and MSE for δ_5. (δ_5 is the so-called "minimax" estimator under squared error loss, in that one can show that the maximum value of its MSE is smallest among all estimators though we are not going to show why that is the case.)

   (b) Plot the MSE for the five estimators when $n = 5$. (The first four are already computed in the notes so you just need to plot those lines and add the line for δ_5. You don't need to use the same colors as in the notes. You can use `col=1,2,3,4,5` in R for the 5 estimators respectively.) How does the MSE for δ_5 compare with the others? Repeat this for $n = 100$.

   (c) Write δ_5 as a weighted average between δ_1 and δ_2. How does the weight compare when we have small n (e.g., $n = 5$) versus large n (e.g., $n = 100$)? In terms of strength of Bayesian shrinkage, what does that imply?

(d) Use Monte Carlo to estimate the mean absolute error (MAE), i.e., the risk function under absolute error loss, of all five estimators on a grid of θ values $\theta = \{0, 0.01, 0.02, \ldots, 1\}$ and plot them. Do this for $n = 5$ and $n = 100$. (You can use a fairly large Monte Carlo sample size, e.g., $S = 100,000$ to get precise estimates.) Comment on how the MAEs compare with the corresponding MSEs. Is δ_5 still "minimax"?

Remark: For plots (b) and (d), make sure you choose a good range for the y-axis so that the lines can be clearly visualized and compared. Refer to "Lecture 6" notes for examples.