MATH 538: Bayesian Data Analysis (Fall 2024)
Assignment #1

Due date: Sunday September 15, 2024

Please upload your solutions in the form of one pdf file onto Canvas before midnight on the due date. Upload your written report and R codes by the deadline. Although you upload your R codes, all codes along with outputs should be included in your pdf solutions file as well.

1. **Book Problems**
   (a) Chapter 1: 1, 2, 3
   (b) Chapter 2: 5, 17

2. **Probability** X_1 and X_2 have joint PMF:

X_1	X_2	$P(X_1 = x_1, X_2 = x_2)$
0	0	0.10
1	0	0.20
2	0	0.15
0	1	0.15
1	1	0.10
2	1	0.30

   (a) Compute the marginal distribution of X_1.
   (b) Compute the marginal distribution of X_2.
   (c) Compute the conditional distribution of $X_1|X_2$.
   (d) Compute the conditional distribution of $X_2|X_1$.
   (e) Are X_1 and X_2 independent? Justify your answer.

3. **Likelihood** Suppose we observe 7 artichoke plants that have the following heights (in meters):

   1.8 1.7 1.4 1.6 1.9 1.5 1.2

   Let y_i be the height of artichoke plant i, and suppose we want to model the y_i as independent realizations from a Gamma distribution with parameters α and β (i.e., with mean α/β and pdf $f(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$)

(a) Suppose that α is known to be 4. Find the maximum likelihood estimator of β.

(b) Now suppose that α is unknown, but β is known to be 2.

i. Try to find the maximum likelihood estimator of α and see that you cannot find it in closed form.

ii. Plot the likelihood as a function of α

iii. Use the plot to find an approximate maximum likelihood estimator of α (For this problem, it is sufficient to just eyeball the plot; optionally, you can do the maximization numerically.)

4. **MAP** Consider the scenario in Problem 3 where α is known to be 4 and β is unknown. Now, let's assume the following conjugate prior for β is a Gamma with $\alpha_0 = 1$ and $\beta_0 = 0.8$, which reflects the prior belief that β has mean of 1.25 with variance 1.56.

   (a) Write out the posterior distribution of β given the data y, $p(\beta|data)$.

   (b) Plot the posterior (as a function of β).

   (c) Use the plot, or numerical evaluation to obtain the value of β that maximizes the posterior distribution (MAP).

   (d) Simulate 5000 values from the posterior and use the quantile function in R to obtain the 2.5^{th}, 50^{th}, and 97.5^{th} quantiles for β. Also report the mean of your posterior samples.