

ST440/540 – Exam 1 - Due February 19

THIS IS AN EXAM - DO NOT DISCUSS THE PROBLEM WITH ANYONE (INCLUDING OTHER STUDENTS OR THE TA)! If you have questions, please email me.

1. The first task is to gather a dataset. The data must consist of three variables
 - X := Binary indicator ($X \in \{0, 1\}$) of the two groups of interest
 - Y := Outcome variable
 - Z := Some other variable that may be related to X and/or Y

Describe your dataset and motivation for the study. In your description, be sure to include the sample size and data source for each variable. The variable X will only be used for last two questions and Z will only be used in the final question.

2. Ignoring X and Z , select and justify a likelihood for Y and conjugate priors for the parameter(s) in the chosen model. Derive (i.e., a few lines of math) the equation (e.g., “Beta($2Y, a + b$)”) of the posterior distribution for the model parameter(s).
3. Apply the method in (2) to your data (use Y only). Plot the posterior distribution of the model parameter(s) and study sensitivity to the prior.
4. Verify that the chosen likelihood is appropriate for the data. Detail your approach and graphically display the results.
5. Conduct (still ignoring Z) a Bayesian hypothesis test of whether the distribution of Y is different for the two levels of X . State your hypotheses in math and words, describe the methods you use to conduct your test, summarize the results and give the consequences of the results in the context of your problem.
6. Use Z to divide the data into subgroups. Repeat the analysis in (5) for each subgroup. Use a Bayesian analysis to determine if the difference between the distribution of Y by X varies across subgroups.

Your paper should be written as a professional document with full sentences, clearly labeled figures and tables and few spelling/grammar errors. Organize your report with subsections corresponding to the questions above. Summarize your analysis in **no more than two pages** (12 font, single space, one-inch margins). Include your (commented) code after the two-page summary and submit a single, 2+ page PDF document. Only the first two pages of the PDF will be graded. **In-class students should turn in the exam in class on Monday, Feb 19. Online students should submit the exam on moodle.**

HAVE FUN!