STAT 6390 Course Project

Description

A project team consists of up to four students (including yourself). Students should decide their own team members. Each team of students will carry out a project involving development and/or application of Bayesian methods to problems of their own choosing. Some examples are:

- Carry out a complete Bayesian analysis of a real dataset. This might involve:
  - description of the research question and dataset
  - specifying an appropriate Bayesian model
  - determining appropriate values for prior parameters
  - fitting the model
  - checking convergence
  - model checking & diagnostics
  - analyzing the output
  - reporting and interpreting the results
- Carry out a Bayesian analysis of a dataset for which a classical analysis has been reported in a journal. Compare and contrast the results obtained by the two approaches.
- Fit a Bayesian model to a dataset using several different choices of prior (hyperparameters and/or functional form). Discuss the meaning of the different results, and the robustness of the model to prior specifications.
- Fit several different plausible Bayesian models to the same dataset. Carry out a check of model adequacy and model fit. Discuss the results.
- Compare different methods of fitting the same model to the same dataset
  - normal approximations
  - MCMC
  - other simulation methods (such as those based an empirical Bayes approach)
  - analytical computation (if feasible)

Note that a team should avoid using simple models directly from Chapter 2 or 3 since they are mostly for illustrative purposes. They could be components for a more complicated model. The difficulty level of the project will play an important role in grading. Utilization of advanced computing techniques is desirable.

Timeline

Projects will have two check points:

- **Project proposal (due 11/11)**

  Groups should hand in one project proposal (with all group members' names on it) by the proposal due date given above. The project proposal is not graded. It exists primarily for you to get feedback on your project idea. The instructor will return the proposals with comments. The proposal should comprise one or two pages addressing the following questions:

  - What is the topic of the project?
  - What are the main issues or problems to be addressed?

- What data will be used, including the variables measured?
- What are your initial thoughts on appropriate models/distributions?
- What questions and/or concerns do you have about the project?

- **Project report (due 12/12)**

  Each team should prepare ONE formal report, of which the main body is no longer than 7 pages (double-spaced or not—you decide, and an up to 10-page appendix is allowed). Please submit the report in one single PDF file and computer code through Canvas Assignments, all by 11am, 12/12.

  The report should address the following points whenever applicable:

  1. Statement of the problem: Describe the questions you address and any key background issues surrounding the questions.
  2. Data: Describe the data used in the project.
  3. Analyses: Describe the analyses you did, including models and prior distributions. Be ready to explain why you believe these methods are justified.
  4. Results: Present relevant descriptive statistics and summaries of posterior inferences. Include tables or graphs that support your analyses (be judicious here--too many tables and graphs hurts the clarity of your message).
  5. Conclusions: Answer your question of interest.
  6. Discussion: What implications do your results have for the population you sampled from? What could be done to improve the study if it was done again? What types of biases might exist? Other potential applications that you can use your analysis strategy?
  7. References: list any publications (e.g., books, papers) you cite, as well as websites you use or download data from. If an R package plays an essential role in your analysis, you should list it, too. This part is not included in the seven pages for the main body.

Grading guidelines

Your instructor will grade your project. We will be looking for the following characteristics:

- **Consistency:** Did you answer your question of interest?
- **Clarity:** Is it easy for your readers to understand what you did and the arguments you made?
- **Relevancy:** Did you use Bayesian statistical techniques wisely to address your question?
- **Interest:** Did you tackle a challenging, interesting question (☺), or did you just collect descriptive statistics (☹)?
- **Techniques:** How well do you apply the techniques you've learned during the semester? and how good is your understanding about the Bayesian theory? If you are using techniques that we did not cover in class, you should succinctly explain them.
- **English:** Check grammar and typos before you turn in your report. English errors should be minimized.