1 Overview

An important skill that a working statistician or mathematician should have is the ability to independently investigate an advanced topic and then communicate the findings to a wider audience. This final project is designed as an opportunity for you to demonstrate semester-long cumulative knowledge of the theory of mathematical statistics, and to investigate the framework and application of a particular statistical method or procedure. As part of your investigation, you will apply the method to a particular data set to answer a research question of interest.

You will work either individually or in small groups of up to 3 people to research a topic in mathematical statistics beyond the scope of what we've covered in class, and then share your results in a 5 - 10 page research paper. This is an opportunity to investigate an area of interest that isn't covered in class this semester. Alternatively, you may dive deeper into a topic that we only briefly covered in class this semester.

2 Objectives

Your final research paper should:

1. Discuss any necessary background information needed to understand your topic.
2. Provide context, motivation, and history indicating how your topic fits into the field of mathematical statistics.
3. Explain how your method relates to other topics studied in our course.
4. Give an explicit and rigorous statement of the subject of your research, as it appears in the literature.
5. Present your own informal interpretation of what the topic means or represents.
6. Apply your method or procedure to at least one data set (either simulated or real), in order to answer a research question of interest.
7. Discuss and provide proof or derivation of one significant theorem or result related to your topic.
8. Incorporate significant material from one or more references beyond our class textbook.
9. Suggest related areas for further research on the topic.

3 Project components

3.1 Proposal

A 1-page proposal for your research project is due by 12:00pm on Saturday, April 27. This proposal should outline the scope and intent of your project. Although changes to the direction of the project are possible after this point, this timeline should give both you and I

the opportunity to ensure the project is appropriate for this assignment. Only one proposal needs to be submitted per group.

Requirements
Your proposal should contain the following:

- A list of all group members.
- The topic your group intends to research.
- A 1-2 paragraph description of the topic you wish to explore, which should describe briefly how this topic fits with other topics we've explore in this class, and also address why someone might be interested in your investigation.
- A description of several possible research questions involving data that might be answerable using your method or procedure, along with a brief indication of why you believe your method can be used to answer these questions.
- A list of at least two potential sources (think textbooks, journal articles) other than our course textbook that can be used in your investigation. Full citation is not necessary, but you should include enough information about the source so that it can be found online or in the library without too much difficulty.

3.2 Research paper

The research paper is your opportunity to provide a technical and in-depth treatment of your research topic. The target audience for your paper is your classmates, who you can assume are familiar with the topics discussed throughout MATH/STAT 310-311, have significant level of statistical and mathematical maturity, and can follow detailed technical arguments comparable to those presented in the textbook. However, you should not assume that your audience has any in-depth knowledge of your particular research topic, or knows the precise statements of definitions, theorems and examples relevant to your topic.

A rough draft of the paper is due by 12:00pm on Saturday, May 11 for peer review. The final draft of your paper is due by 11:59pm on Monday, May 20. Only one research paper needs to be completed and submitted per group.

Requirements
In addition to accomplishing the objectives listed in the Overview section above, your paper must:

- Be typed single-spaced using 10 - 12 pt font with 0.75 - 1.25 inch margins.
- Be *at least* 5 pages long, with 2 additional pages per group member beyond the first (e.g. the minimal length of a group of a single member is 5 pages, whereas the minimal length of a group of three people is 9 pages). Going beyond the minimal length is fine.

- Include a self-created graph, diagram, chart, or some other visual aid, which can be either be computer generated or hand-drawn.
- Make-use of legible mathematical typesetting where appropriate (it is not necessary to use LaTeX, but any equations used should be displayed in a readable and unambiguous manner).
- Include citations to at least one relevant source beyond our textbook, per group member (i.e. a group of three must include at least three references other than our course textbook).
- Include an Appendix at the end of your paper with the code you used for your Application. This can be done easily by knitting your .Rmd file to PDF and using some software to combine your final paper with the Appendix. When discussing results for your application (Objective 6), you may simply state results and reference/point the reader to the code in the Appendix!
- Be written using an appropriate style for a professional academic publication, with clearly defined and consistent statistical notation, and correct grammar, spelling and punctuation.

3.3 Project presentations

Each group will give a brief presentation on their project during our scheduled exam period, which is 9:00am-12:00pm on Saturday, May 18. The duration of each presentation will depend on the number of groups, but tentatively plan for presentations between 10-12 minutes long. The presentations should be a distilled version of the research paper, but delivered in a way that is engaging to the audience.

Requirements
The presentation and presenters should:

- Address objectives 1-6 through a combination of oral and written (i.e. chalkboard and/or slides) delivery.
- Be practiced so the delivery is smooth and coherent.
- Be prepared to answer questions from other students in the course.

3.4 Reflection

A 1-page reflection on the final project is due by 11:59pm on Monday, May 20. Each person should submit an individual reflection. More details about what should be addressed in the reflection will be shared during the final week of class.

4 Possible topics

You have considerable latitude in choosing your topic, and what you choose to research will depend on your background and interests. The topic you choose should be something relatively new to you (although it's fine if you've seen some parts of it before, as long as the core component of the project leads you to discover something you didn't know before).

A list of potential topics can be found below. For those topics appearing in our *Probability and Statistics* textbook, the relevant section is listed. Topics have been loosely arranged by theme, although some topics may easily fit within more than one listed theme. Please don't pick a topic just because there is a section from the textbook listed; pick a topic that actually sounds interesting to you!

You may also choose a topic not listed below, although I'd advice you to consult with me before delving too deeply into the topic to make sure it is of appropriate scope.

Bayesian statistics

- Posterior predictive distributions
- Invariance principle for Bayesian estimates, the Jeffreys prior

Properties of Statistics and Estimators

- Sufficient statistics (7.7)
- The delta method for asymptotic distributions (6.3)

Tests for categorical data and Non-parametric tests

- χ^2 test for goodness-of-fit (10.1)
- Fisher's Exact Test for goodness-of-fit
- Goodness-of-fit tests for composite hypotheses (10.2)
- χ^2 test for independence and homogeneity (10.3, 10.4)
- Komogorov-Smirnov Test for normality (10.6)
- Robust estimator (10.7)
- 1-sample nonparametric tests
- 2-sample nonparametric tests

Simulation and Computational Methods

- Bias-corrected bootstrap confidence intervals
- Permutation-based hypothesis tests