Final Project

EPsy 8264

Fall 2018

This is an independent project, meaning that you need to complete it individually. You need to do the data analysis and complete the write-up of the results on your own. This being said, you are welcome to discuss the project with others to garner opinions and suggestions for improvement. Think of it like you are the solo analyst and author on a manuscript—while the work is independent, it does not need to be done in isolation. You can also consult with the instructor for additional help and clarification. Additionally, you may use your class notes, books, or any other related material you like (including the internet).

There are two primary parts to this project:

- A proposal, which is due on November 06, 2018.
- The final project report, which is due on December 12, 2018 (by 12:00pm).

Early submissions on both parts are encouraged.

Goal and Outline

The goal of this project is for you to complete and write-up either (1) an empirical analysis using advanced regression models and methods (e.g., not just conventional linear multiple regression), or (2) a methodological paper exploring the properties of particular advanced regression models and methods.

Empirical Analysis

Projects carrying out an empirical analysis of data, should mimic an applied research article including detail from that begins by laying out the scientific background and motivation, methods and analysis, and resulting conclusions, as well as a discussion of scientific implications. Although the substantive background is imprtant for motivating the analysis, your grade will be primarily based on how well you communicate and justify the statistical modeling and analysis employed and how well you explain and interpret the results in a way that readers not familiar with advanced regression models and methods can understand. In particular, a strongly suggested outline for an empirical project is:

- **Introduction.** What are the objectives of, or motivation for, the work? What substantive background/literature is needed for the audience to understand the work? Why specialized statistical models and methods are required for analysis of the data (so why methods familiar to the researchers are not appropriate). State clearly the scientific objectives and why they are of interest. Are there specific hypotheses to be addressed?
- **Methods.** What are the variables of interest and how were the data collected? What was the sampled population? What statistical model have you chosen, including its interpretation and how researchers' questions may be cast formally in terms of the model. What technique(s) did you use to analyze the data? Exploratory/descriptive? Modeling? Diagnostics? Summarize the data (this is often most effective when done graphically with few descriptive statistics). Provide a rationale for choosing the model, any assumptions you have made, and why these assumptions are reasonable for the researchers' situation. What method did you use to fit the model?
- **Results.** What are the main results of your analyses? Describe these using the language and in context of the substantive area of interest. How do these results relate to your objectives and research questions?

- **Discussion.** Summarize the main points of your work. Discuss any scientific or statistical problems with the data, the analyses, or the interpretation. Mention any future research directions.
- **References.** In a format suggested by flagship journals in your research domain. References are formatted differently in *Psychological Methods* than they might be in *Journal of Public Health*.
- **Appendix.** Any syntax that produces results cited in your report should be included in the appendix. This syntax should be commented. Do not include the output; only include syntax here. Also do not include any syntax outside of what you used to produce results referenced in the report. The appendix is for my information only; your report should not refer to the appendix (i.e., the report should not ask readers to go to pages of code or output. Any results that the audience needs to see should be provided in the body of your report.

Methodological Paper

Methodological-focused projects should address the study and improvement of the statistical methods themselves. While these projects may also use data, the data are just a tool for studying the methods. (Often the data in methodological studies are simulated.) An outline for a methodologically-focused project may be:

- **Introduction.** What are the objectives of, or motivation for, the work? What are the strengths and weaknesses of the chosen model or methods? State clearly the scientific objectives and why they are of interest. Are there specific methodological hypotheses to be addressed?
- **Methods.** How will you study the method? If you are simulating data, you need to describe (in detail) how you will simulate the data (e.g., what will be assumed). What are the methods and criteria you will use to evaluate the methodological question? Provide a rationale for choosing theses methods.
- **Results.** What are the main results of your analyses? Describe these using the language and in context of the methodological area of interest. How do these results relate to your objectives and research questions?
- **Discussion.** Summarize the main points of your work. Discuss the analyses, assumptions, and provide interpretation. Mention shortcomings/limitations and any future research directions.
- **References.** In a format suggested by flagship journals in your research domain. References are formatted differently in *Psychological Methods* than they might be in *Journal of Public Health*.
- **Appendix.** Any syntax that produces results cited in your report should be included in the appendix. This syntax should be commented. Do not include the output; only include syntax here. Also do not include any syntax outside of what you used to produce results referenced in the report. The appendix is for my information only; your report should not refer to the appendix (i.e., the report should not ask readers to go to pages of code or output. Any results that the audience needs to see should be provided in the body of your report.

If you are looking for inspiration, you may want to look through a recent issue of *Psychological Methods* or *Educational and Psychological Measurement*. Two other resources include:

- Feinberg, R. A. & Rubright, J. D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice, 35*, 36–49. doi:10.1111/emip.12111
- Harwell, M., Kohli, N., & Peralta, Y. (2017). Experimental design and data analysis in computer simulation studies in the behavioral sciences. *Journal of Modern Applied Statistical Methods, 16*(2), 3–28. doi: 10.22237/jmasm/1509494520

Audience

When you write-up the project, you need to write for an appropriate audience. For example, methodological projects will have a much different audience (assume mathematical and statistical sophistication) than empirical project (better to assume a less statistically sophisticated audience).

Project Proposal

You will need to submit a project proposal by **November 06, 2018**. This proposal should be no more than two pages in length. In this proposal, you will include the research question(s) that you propose to answer in your project, a brief rationale for examining those questions (why are they important, etc.), and information about the dataset(s) or methodology you plan to utilize to answer those questions. For empirical projects, identify the outcome variable(s) you plan to use in the analysis, along with any predictors and covariates. If you have thoughts or ideas about initial models you may fit to answer the research questions, include those as well. For methodological studies, include information about what properties of the method you will be studying, as well as how you plan to study those properties.

Details in this proposal may change as you explore the data and carry out your analyses, so you aren't locked-in to the proposal. Rather, it is to give me a more systematic idea of what you are thinking about for the project. It will also give you a basis for writing the *Introduction* and *Methods* sections of the project.

Grading

The project report is worth a total of 40 points (and 40% of your final course grade). Note that the proposal is not graded. The following guidelines will be used to grade your project:

- **Technical Detail (15 points).** Is the analysis problem or research focus of the report formulated clearly using appropriate terminology for the audience? Are the appropriate statistical tools used properly? Are the assumptions checked when possible? Are the research or analysis conclusions justified? Are the analysis interpretations and implications appropriately described for the audience?
- **Thoroughness (10 points).** Is there enough substantive or methodological background provided so that a fellow researcher can understand the general ideas of your study, and the implications of your study for your population of interest? Was anything obvious overlooked, for example, a too simple procedure performed as a substitute for a more appropriate procedure, an assumption not checked, conclusions not stated for each objective? Is the work of others appropriately cited, and the references for those cited materials included?
- **Clarity (10 points).** Is the report easy to understand? Does it follow a logical structure, both within and across sections? Is it concise and clearly written?
- **Innovation (5 points).** Were the data or was the science particularly challenging or complex? Was any extra information provided which added to the quality of the report (EDA, diagnostics)? Was the report exceptionally well done, smoothly flowing with easily understandable presentation of information? *Note: These are points to be explicitly earned; the default score for this category is 0.*