Project:
Applying an Advanced Topic in Applied Statistics

Important reminder This project is in lieu of a test, and as such you are not allowed to work with anyone, including other students. Any questions that you have about it should be directed to me only.

Assignment

Choose a real-world problem and address it with a method or topic in applied statistics that was not covered in Math 260 or in this course. Write a paper reporting on your study, and present the results of your study to the subject in class.

Purpose

In order of priority, the goals of this assignment are:

1. To learn how to apply statistical methods that you have taught yourself in order to address a particular real-world problem.

2. To learn how to report on the results of an advanced statistical study, both in a paper and in a presentation.

3. To learn how to use an advanced statistical method that is new to you.

While learning how to use the specific actual statistical method is a goal, it is not the most important one.

Due dates

This project has two due dates:

- Monday, April 23 to Wednesday, May 2 in-class presentation (exact date TBA)
- Wednesday, May 2 paper due

Before the presentations begin, we will decide in class on your exact in-class presentation date.

You should submit all your files to me by email. Your R Markdown file should be named:

yourLoginName-paper2.Rmd

If you have data files, you should give them file names that begin with yourLoginName- but after the hyphen you may choose the name.

If you have data files more than 5 MB, please discuss with me how to submit those. (Don’t submit them by email.)
Grading

Only the paper is graded; the in-class presentation is not graded separately but is an important part of the learning goals. The only way that the in-class presentation would contribute to your project grade is if you did not prepare adequately for it. If that is the case, then that will decrease your overall grade. Otherwise, the presentation is solely for the experience and is not graded.

I will grade your papers separately in two broad areas: content and writing. Content refers to the mathematical aspects of the paper and how well you have shown that you understand them. Writing refers to clarity of exposition, which includes structure, grammar, spelling, typos, layout, and so forth.

There is a trade-off between your study’s level of sophistication and the depth of understanding that I expect. If you choose to conduct a relatively simple study, you will have to explain it extremely well and demonstrate that you have a solid understanding of it in order to earn a high grade. If you choose to conduct a more advanced statistical study, your explanation and understanding of the material still need to be good, but I will not hold them to quite so high a standard as with a simpler study.

Assignment details

In this project, you should report the results of your statistical study to a reader who has a background similar to that of students in Math 360. The study must be approved by me in advance of your presentation.

The main things to keep in mind when writing a report on your study are:

1. The study should not just report your results. Rather, it should illustrate and explain the process of how to obtain those results. (Most of the paper should be devoted to the explanation, with the results occupying only a small part of it.)

2. Include everything relevant for readers to understand how your study was conducted, and possibly even to conduct a version of your study themselves.

3. Be explicit about any assumptions that you are making in your study.

4. Be precise (and correct) about what conclusions you can and cannot draw from your study.

5. Include a frank and open discussion of not only the strengths but also the possible weaknesses of your study. This is important for establishing yourself as a credible researcher.

The overall structure should be similar to papers that you wrote in Math 260 (with clearly labeled sections, an introduction describing what the paper is about and how it is structured, a section containing your conclusions and results, and a final section in which you discuss the strengths and weaknesses of your study). Your paper should be written in R Markdown (submit the actual R Markdown file, not the word document), and you should send me all relevant data files, appropriately named. Be sure to include a bibliography in which you cite not only your data sources and the sources you used to determine how to use your method, but also R and any R Packages that you used in the analysis.

As usual, your paper should not contain unformatted R text output. (If it must, please consult me beforehand.) It of course may contain graphics generated in R (or rather R Markdown).
There is no required length for your paper. It should be long enough to clearly explain what you are trying to explain, and no longer. I am guessing that this will make most students’ papers about 12-18 pages long (length if it were double-spaced), but you should not strive specifically for this range. (I would like the distribution of page lengths in the class to be approximately normally distributed with mean 15 and standard deviation 2.5.)

Your in-class presentation should be 12-15 minutes, and it should be both interesting and informative to your fellow classmates. You shouldn’t necessarily present everything that is in your paper, since people can follow highly technical arguments in a paper much better than they can absorb them in a talk. A short talk such as this calls for especially careful consideration of which details to include and, more importantly, which to leave out.

You may use any type of technology that you like in your presentation: R Markdown (Slidy), Powerpoint, chalk, none of the above, etc.

Also, I offer the following two suggestions to help your presentation go smoothly:

1. *Rehearse* your presentation, preferably in front of at least one other person. Rehearsing the presentation is the best (if not the only) way to get the timing right. Rehearsing also helps catch typos, errors, stumbling points, and other subtleties that can detract from or even derail a presentation.

2. *Check the technology.* If you are going to use any technology in your presentation, try it out on the equipment that you will be using. If you will use the classroom computer, check that your files display correctly on it. If you will hook your laptop into the classroom projector, make sure beforehand that you can indeed connect it and that when you do, your files display correctly.

Following these suggestions won’t necessarily eliminate all the errors and foibles from your talk, but it will at least help to minimize them.