

Final Project Guidelines

Presentations on Wednesday, May 2, 9:45-11:45 a.m.

This is the date/time designated by the University for our final exam period.

General Description: The final project involves preparing an oral presentation (@15 minutes + 3-5 minutes for questions). The oral presentation will be held during our scheduled final exam period. Also, you may work in groups of no more than 2 students.

The project presentations should involve three main components outlined below.

1. Find a research manuscript on some *resampling method* of *your choice*, typically involving a problem with *dependent data*.
2. Provide a concise summary of the method (e.g., what is the purpose, what are the mechanics, what are the assumptions/limitations?).
3. Include a “creative component” in the project. This creative component involves implementing or examining the resampling method in a manner of your choice. For example, this could be a data application (e.g., numerical illustration) or a simulation investigation or a software development or a limited theoretical study of the method.

More Guidance and Details regarding each point above:

1. Find a research manuscript on some *resampling method* (bootstrap, subsampling, empirical likelihood) of *your choice*, typically involving a problem with *dependent data*.
 - The resampling method could involve bootstrap, subsampling, or empirical likelihood (EL), but it has to address dependent data (time series or perhaps even spatial data).
 - You can find articles using Google but also try the Current Index of Statistics (CSI) (the department has a link for this on its website).
 - Your selected manuscript may or may not be published, but try to select an article written in the last 5-6 years.
 - There are some other resampling methods that we have not discussed in class: wild bootstraps for dependent data, time domain bootstrap samples from resampling in the frequency domain, generalized versions of empirical likelihood, resampling methods for high dimensional data, even spatial or spatio-temporal developments, etc. These are merely suggestions, but I’m sure that more possibilities exist.
 - Most importantly, pick something that interests you.
2. Provide a clear and concise summary of the method.
 - Briefly summarize the nature of the method. (What is the method for? Why it is relevant or how is the method useful?)
 - For what kind of dependent data is the method intended? What are the assumptions on the data generating process? Are these assumptions stringent or reasonable? How widely applicable is the method?
 - Describe the mechanics of the method in some reasonable amount of detail (but don’t get bogged down).

- What are the advantages or disadvantages (limitations) of the method? Again, generally where is the method going to prove most useful in statistical inference and why is this inference important?

The presentation should highlight the main points above and be based on your judgement. In other words, this summary and critique of a method should be *your own words and thoughts*. You are **not** to copy anything, at any point, from the manuscript word-for-word (use your own words). Most research manuscripts, if they are reasonably written, will address some of points mentioned above but you should develop your own thoughts on the potential merits or disvalues of a method.

3. You will probably have to develop computer code (in R or otherwise) for implementing or illustrating the method in some respect.
 - Try to avoid directly using other people's source code or R-packages for the implementing the method (assuming that an R-package has been developed for the method), or at least reasonably verify that source code is working properly.
4. Include a “creative component” in the project. This creative component involves investigating or examining the resampling method in a manner of your choice (apart from any numerical work presented in the manuscript), such as
 - implementing the method to conduct a real data analysis. This would involve having some real data and using the resampling method to perform inference with meaningful interpretation.
 - conducting a limited simulation study. The study could examine the performance of method (based on some criterion like MSE or coverage accuracy), but the context should differ from anything directly presented in the original manuscript (e.g., use a different inference problem or different processes). Or the simulation study could be comparative, investigating and comparing different resampling methods based on some performance criterion.
 - develop some theoretical results, even in a context which may be limited in scope. This option may be the least accessible, but it may be possible to develop or extend some theoretical results on the performance of the method if the theoretical problem is focused and narrowly framed (e.g., assume the process is m -dependent with all moments). In other words, it's ok to make strong assumptions if it would become easier to develop any theory at all.
 - try applying the method in an inference context where it wasn't originally intended. Sometimes this can help in understanding the limitations of a method or how it compares to other resampling methods in the same problem.

This “creative component” is *intended to be limited*, meaning that you should be able to complete this within a week or so after having the other steps in place.

At the time of your presentation, *I will request a copy of your slides*. If appropriate or helpful, you may also include a few “extra” slides with supporting details (beyond slides directly intended for the oral presentation itself), but this is unnecessary.

If you have questions or concerns, please come and talk with me.