Reading for Oct. 30 section discussion

October 23, 2017

The goal of this problem is to think carefully about the design and interpretation of simulation studies. In particular, we'll work with Lo et al. (2001), a recent article in JASA (the Journal of the American Statistical Association), which is a leading statistics journal. The article is available as *lo_etal_2001.pdf* under the *ps* directory on Github. Read the first three pages and Section 3 of the article. You don't need to understand their algorithm for testing the null hypothesis [i.e., you can treat it as some black box algorithm] or the theoretical development, though it may help to skim through some of the material on the algorithm for context.

For problem 1 of PS6, briefly answer the following questions. Before section on Monday Oct. 30, please put your answers to these questions in a text file named *ps6_question1.txt* and push to the *ps6* directory of your Github repository.

- What are the goals of their simulation study and what are the metrics that they consider in assessing their method?
- What choices did the authors have to make in designing their simulation study? What are the key aspects of the data generating mechanism that likely affect the statistical power of the test? Are there data-generating scenarios that the authors did not consider that would be useful to consider?
- Do their tables do a good job of presenting the simulation results and do you have any alternative suggestions for how to do this?
- Interpret their tables on power (Tables 2 and 4) - do the results make sense in terms of how the power varies as a function of the data generating mechanism?
- How do you think the authors decided to use 1000 simulations. Would 10 simulations be enough? How might we decide if 1000 simulations is enough?