12/8/2022

Some remarks on Bayesian statistics
All the statistics we've done so far in this course are sometimes referred to as "frequentist" statistics. The basic idea is that we collect data, and then we make an inference on that data without making prior assumptions about the conclusion. It has been the most popular form of statistics for quite some time, but there are critiques of the process. There is a lot of discussion about P-values, for instance, and the consequences of publishing only affirmative results that can be misleading. It's worth considering these controversies when analyzing studies.

But there is another approach to statistics that comes at some of these issues from a completely different perspective. Bayesian statistics is founded in the idea of conditional probability and Bayes' Rule (which we've covered in this class). Recall:

$$P(A \mid B) = \frac{P(A \,\&\, B)}{P(B)}.$$

In Bayesian statistics, prior information is taken into account as we take in new data. Typically, some prior distribution is assumed, and the new data is taken in to modify the distribution, and this process continues each time new data is assimilated. This can be useful if you have reasons to believe something about the result (say for physical reasons, from differential equations or other areas of mathematics and science), then that information can be incorporated into the analysis. New data can modify these prior assumptions, and lots of data can modify them a lot. Many Bayesians argue that this process is more similar to the way that people actually reason. The reference [4] linked below is an entire online text about Bayesian statistics and they have a more in-depth discussion of the difference between Bayesian and frequentist statistics.

It is worth noting, however, that while the philosophy of Bayesian statistics and frequentist approaches are quite different, and lead to different computational methods, nonetheless, in the presence of sufficient data, the results of the analysis converge to the same numerical solutions.

Next semester, we'll continue with the frequentist approach, but we may touch on a couple of techniques that were born from Bayesian approaches.

Review for final exam
Review material from Exam #1 and Exam #2 for the comprehensive portion of the Final.

For the recent material, focus on the following:

- Non-parametric statistics
    - Wilcoxon test for sign test/rank-sum test
    - Non-parametric ANOVA (Kruskal-Wallis, Friedman's ANOVA)
    - Permutation tests
    - Bootstrapping
- χ^2 tests

- Goodness-of-fit tests
- Test of homogeneity
- Test of independence
- Fisher Exact Test

The final exam will have the same format as the two previous exams.

Don't forget to submit the final draft of your final project as well. Be sure to include suggested changes from your rough draft into your final version.

References:

1. https://assets.openstax.org/oscms-prodcms/media/documents/IntroductoryStatistics-OP_i6tAI7e.pdf
2. https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineering_and_the_sciences.pdf
3. http://www.scholarpedia.org/article/Bayesian_statistics
4. https://statswithr.github.io/book/the-basics-of-bayesian-statistics.html