STAT 345-01: Nonparametric Statistics

Problem Set 4

Assigned 2018 September 20
Due 2018 September 27

Show your work on all problems! Be sure to give credit to any collaborators, or outside sources used in solving the problems. Note that if using an outside source to do a calculation, you should use it as a reference for the method, and actually carry out the calculation yourself; it's not sufficient to quote the results of a calculation contained in an outside source.

Please hand in parts one and two separately.

1 Part One

1.1 Conover Problems on Signed Rank Tests

Exercise 5.7.4

Exercise 5.7.8

Problem 5.7.4

1.2 Conover Problems on Rank Sum Tests

Exercise 5.1.6

Problem 5.1.2

2 Part Two

2.1 Power Curve of Wilcoxon Signed Rank Test

Download the ipython/jupyter notebook
`http://ccrg.rit.edu/~whelan/courses/2018_3fa_STAT_345/data/ps04.ipynb`
using the username and password given in class, and carry out the exercises, evaluating the cells and adding commands as necessary to complete the problem according to the instructions. Submit your completed notebook as a hardcopy, or via email (either the final `.ipynb` file or a pdf, which can be created using nbconvert).

2.2 Mann-Whitney Test on Gravitational Wave Search Results

Please turn in some sort of transcript of your python session, along with answers to the questions posed. If you want to submit electronically, please send either a pdf (with a white background so it can be printed) or a plain ASCII file. No Word documents!

The paper "Search for gravitational-wave inspiral signals associated with short Gamma-Ray Bursts during LIGO's fifth and Virgo's first science run" (*The Astrophysical Journal* **715**, 1453 (2010), available on campus via `http://stacks.iop.org/ApJ/715/1453` and off campus via `http://stacks.iop.org.ezproxy.rit.edu/ApJ/715/1453`) reported on a search for gravitational waves in 22 time windows, each corresponding to a gamma-ray burst. None of these results was individually significant, i.e., the highest detection statistic was consistent with noise given the "trials factor" of 22. To test whether there was an overabundance of signals too weak to be individually significant, the authors also performed a Mann-Whitney U test of these 22 "on-source" results against a set of 6801 "off-source" searches at times with no GRB, and reported, "Applying the U-test, we find that the two distributions are consistent with each other; if the on-source and off-source significances were drawn from the same distribution, they would yield a U-statistic greater than what we observed 53% of the time." Download the results from
`http://ccrg.rit.edu/~whelan/courses/2018_3fa_STAT_345/data/ps04_onsource.dat`
`http://ccrg.rit.edu/~whelan/courses/2018_3fa_STAT_345/data/ps04_offsource.dat`
using the username and password given in class, and perform the Mann-Whitney U test or the equivalent Wilcoxon rank-sum test. Convert the statistic to an equivalent z score via the normal approximation using the expected mean and standard deviation (even if your software does this for you, check using the raw numbers; note that since both lists have multiple 0 values, you will need to use the corrected formula for the variance considering these ties). Check that the stated p-value of .53 is correct.