Exercise session 11 – Prob. & Stat. 2 — Dec 12, 2023

Nonparametric tests

Possible choices:

- for 2-sample tests: Permutation test (scipy.stat.permutation_test), Mann-Whitney U-test, also known as Wilcoxon sum test (scipy.stats.mannwhitneyu) (we didn't cover the latter one)

- for 1-sample & paired tests: Sign test (easiest using scipy.stats.binomtest), Wilcoxon signed rank test (scipy.stats.wilcoxon)

1. Warm-up: apply the Wilcoxon signed rank test on data $1, 3, 4, -2$. (H_0: median is 0 (≤ 0, ≥ 0). Do it by hand and/or by computer. Think about what it means.

Use the sign test on the same data.

2. It is claimed that a student can increase his/her score at a test by at least 20 point, if he/she is provided with sample problems in advance. To test this, 20 students were divided in ten pairs (of comparable abilities), one student of each pair got sample problems a week ahead of the test. Students did not cooperate.

Results of students with sample problems: $591, 621, 683, 579, 451, 680, 691, 769, 563, 575$.

Results of students without sample problems: $509, 540, 688, 502, 424, 683, 568, 748, 530, 524$.

(a) Test the null hypothesis, at the 0.05 level of significance, that sample problems increase score by at lest 20 points versus the alternative, that the increase is less than 20 points. Use the sign test and the Wilcoxon signed rank test.

(b) Meditate about why the tests give different answers. How to decide, which one is better? Which one should you use? When should you decide?

3. Students coming for an oral exam had to wait the following number of minutes before being examined: $17, 15, 20, 20, 32, 28, 12, 26, 25, 25, 35, 24$. Test at the 0.05 level of significance the professor's claim that the median waiting time is at most 20 minutes. Use the sign test and the Wilcoxon signed rank test.

4. Another warm-up: use the Mann-Whitney U-test on data $1, 3$ in one group and $2, 4$ in second group. Also use the permutation test.

5. Two brands of cell-phones are compared in terms of how many years they last before failure.

A: 2.1, 4.0, 6.3, 5.4, 4.8, 3.7, 6.1, 3.3

B: 4.1, 0.6, 3.1, 2.5, 4.0, 6.2, 1.6, 2.2, 1.9, 5.424

Test the hypothesis that the durability of both brands is the same (against the alternative that they are unequal). Use Mann-Whitney U-test and/or permutation test.

6. In table on `http://www.biostathandbook.com/kruskalwallis.html`, 27 dogs are ranked by their dominance. Test the hypothesis that male dogs are more dominant. Explain, why in this case we cannot use a parametric test.