# Lab 3: Wilcoxon Rank-Sum Test

This lab focuses on a new two-sample test called the Wilcoxon rank-sum test. In this lab, you will first work through the test by hand (on a small data set). Then you will do a larger test using `R`.

## Sleep deprivation

"Researchers have established that sleep deprivation has a harmful effect on visual learning. In a recent study, Stickgold, James, and Hobson (2000) investigated whether subjects could "make up" for sleep deprivation by getting a full night's sleep in subsequent nights. This study involved randomly assigning 21 subjects (volunteers between the ages of 18 and 25) to one of two groups: One group was deprived of sleep on the night following training with a visual discrimination task, and the other group was permitted unrestricted sleep on the first night. Both groups were allowed unrestricted sleep on the following two nights and then were retested on the third day. Subjects' performance on the test was recorded as the minimum time (in milliseconds) between stimuli appearing on a computer screen for which they could accurately report what they had seen on the screen. Previous studies had shown that subjects deprived of sleep had performed significantly worse the following day, but it was not clear how long these negative effects would last." (Rossman and Chance, 2001)

The data for this experiment are given as the performance score immediately after training minus the performance score on the third day. So negative values indicate a decrease in performance.

## Wilcoxon rank-sum test

The Wilcoxon rank-sum test is a permutation test that uses the sum of the ranks as a test statistic.

### What are ranks?

Let $X_1, X_2, \ldots, X_n$ be your data set of $n$ observations. The rank of $X_i$ among the $n$ observations is given by

$$rank(X_i) = \text{the number of } X_j's \leq X_i.$$

For example, the table below shows the ranks of a hypothetical data set. For now, we will assume that no two observations are the same, meaning there are no ties between data points.

| Data | 26.4 | 18.1 | 16.2 | 12.0 | 25.9 | 11.8 | 26.1 |
|---|---|---|---|---|---|---|---|
| Ranks | 7 | 4 | 3 | 2 | 5 | 1 | 6 |

### Wilcoxon rank-sum statistic

The test statistic for the Wilcoxon rank-sum test is $W$, the sum of the ranks in group 1. For example, if the first three data points from our hypothetical data set (26.4, 18.1, and 16.2) were in group 1, and the rest were in group 2, our test statistic would be $W = 7 + 4 + 3 = 14$. Notice that the ranks are calculated with the two groups combined. It doesn't really matter whether you use the sum of the ranks from group 1 or the sum of the ranks from group 2, as long as you are consistant within a single analysis.

## A simple example

Using a subset of the sleep deprivation data, you are going to work through the Wilcoxon rank-sum test by hand. Suppose we have the following data:

| group | change in performance |
|---|---|
| deprived | 7.2 |
| deprived | -10.7 |
| deprived | 21.8 |
| undeprived | 25.2 |
| undeprived | 11.6 |

1. Calculate the observed test statistic, $W_{obs}$ from the data. To do this, combine the observations from both groups and then order them from smallest to largest. Assign ranks to the observations with 1=smallest and 5=largest. Then calculate $W_{obs}$ = the sum of the ranks in the deprived group.

2. List all possible assignments of the ranks to the two groups. To do this, make a table with two columns. In the first column, assign 3 ranks to the sleep deprived group. In the second column, assign the remaining 2 ranks to the undeprived group. You should have something that looks like this:

| deprived | undeprived |
|---|---|
| 1 2 3 | 4 5 |
| 1 2 4 | 3 5 |
| ⋮ | ⋮ |

3. For each possible assignment of the ranks, calculate $W$, the sum of the ranks in the sleep deprived groups. To do this, add a third column to your table:

| deprived | undeprived | $W$ |
|---|---|---|
| 1 2 3 | 4 5 | 6 |
| 1 2 4 | 3 5 | 7 |
| ⋮ | ⋮ | ⋮ |

4. Calculate a $p$-value as the proportion of possible assignments where the test statistic is "as or more extreme" than the observed test statistic. In your case, you are trying to show that the sleep deprived group performed significantly worse than the undeprived group.

5. State a conclusion about the effect of sleep deprivation on visual learning.

6. Now suppose you are trying to show that the sleep deprived group performs differently than the undeprived group. Calculate a $p$-value for this alternative hypothesis and state a conclusion.

## Using R

We can perform a Wilcoxon rank-sum test in R by adapting the code we have been using to perform permutation tests. Below you will find the code that can be used for performing this test in R. You will need to load the `gtools` library in order to use the `combinations()` function.

```
ranks = rank(performance)   ### calculates the ranks of the data

W.obs = sum(ranks[group=="deprived"])

deprived.ranks = combinations(5, 3, v=ranks, set=FALSE, repeats.allowed=FALSE)

W = rep(NA, choose(5,3))
for(i in 1:choose(5,3)) {
W[i] = sum(deprived.ranks[i,])
}

sum(W <= W.obs)/choose(5,3)
```

7. Comment on what has changed in this code compared to the code we have previously used to perform permutation tests.
8. Run the code to perform the test. Currently, it is set up to test the alternative that the sleep-deprived group performs worse than the undeprived group. State your $p$-value and a conclusion.
9. Change the code to test the alternative that the sleep deprived group performs differently than the undeprived group. Calculate a $p$-value for this alternative hypothesis and state a conclusion.

## Dealing with ties

So far we have been working under the assumption that there are no repeated observations in our data. If there are multiple observations of a single value, we must adjust the ranks to deal with these ties. To do this, we assign the average rank to the tied observation. For example, if there are two observations tied for the second smallest place, instead of giving one a rank of 2 and one a rank of 3, we give them both 2.5. Doing this preserves the overall sum of the ranks.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Data | 26 | 18 | 16 | 12 | 26 | 12 | 26 | |
| Ranks | 5 | 4 | 3 | 1 | 6 | 2 | 7 | sum of ranks = 28 |
| Adjusted ranks | 6 | 4 | 3 | 1.5 | 6 | 1.5 | 6 | sum of ranks = 28 |

The entire steep deprivation data set can be found in the file `sleepdep.txt` on the course website. Read this data into R. Don't forget to attach it! (R code: attach(sleepdep or whatever you call the data in R) )

10. Notice that this full data set does contain some tied observations. Comment on how the `rank()` function deals with ties by trying `rank(performance)`.
11. Run a Wilcoxon rank-sum test for the entire data set. Test the alternative that the sleep-deprived group performs worse than the undeprived group. Report your $p$-value and state a conclusion.
12. Run a Wilcoxon rank-sum test for the entire data set. Test the alternative that the sleep-deprived group performs better than the undeprived group. Report your $p$-value and state a conclusion.
13. Run a Wilcoxon rank-sum test for the entire data set. Test the alternative that the sleep-deprived group performs differently than the undeprived group. Report your $p$-value and state a conclusion.

## Lab Summary

Please write clearly and in complete sentences.

1. Write a paragraph summarizing your results about the relationship between sleep-deprivation and performance on a visual learning task, based on the entire data set. Report your results for 11, 12, and 13. Explain why the results from these three tests reconcile with each other.

2. Explain why using the sum of the ranks in the undeprived group as your test statistic will result in the same $p$-value as using the sum of the ranks in the deprived group.