

Nonparametric Tests App User Guide

Ivy(Ruxin) Tong

Goal of the application

The app is designed for users to apply their data for Kruskal-Wallis (KW) rank sum test and Jonckheere-Terpstra (JT) test. The app is designed for users to test if there are statistically mean differences between multiple groups and to test if there is an ascending order of the mean differences among the groups. This app can produce the boxplot of your data, give the JT and KW test p value, and generate the JT star histogram.

File Requirements

It is required that file uploaded should be in the text format. Additionally, it should be formatted in such a way that groups should be in the rows and blocks should be in columns. It is also required that each group has the same sample size.

	Block 1	Block 2	Block 3	...
Treatment 1
Treatment 2
Treatment 3
...

Contents of the application

1.1 Nonparametric Tests are used in the app

The app employs two nonparametric tests:

Kruskal-Wallis rank sum test can be used to identify statistically significant mean differences between two or more groups of an independent variable.

Jonckheere-Terpstra test tests for the significant ordered differences in mean. The app is specifically designed to test the ascending order of mean differences.

1.2 Backend Function

KW Test: The Kruskal-Wallis test determines whether two or more groups have significant difference in means.

$$H_0: \mu_1 = \mu_2 = \dots = \mu_n$$

H_A : at least one of them is different from each other

The data should be formatted with each treatment separated by row. The test reads the data and determines how many treatment groups there are. It then combines all the data into a single list, finds the total number of data points, and then ranks the data. From this, the expected mean rank is calculated, as well as the average rank per group. The KW statistic is calculated by taking the difference of each treatments mean rank and the expected mean rank. The test then squares these values, multiplies them by the number of data points in a treatment, and sums the new values. Finally this sum is scaled, providing the KW observed statistic. In order to calculate the p-value for the data, the test runs the number of permutations, R, you selected. A sample is taken from the combined data and R KW statistics are calculated and stored. The p-value is calculated by finding the number of those new KW statistics that are larger than KW observed and dividing that quantity by the permutation size R. The app then displays the p-value.

JT Test: The Jonckhere-Terpstra tests if there is an order to the mean results of different treatments or groups. The hypothesis for this test in this specific application is:

$$H_0: \tau_1 = \tau_2 = \dots = \tau_K$$

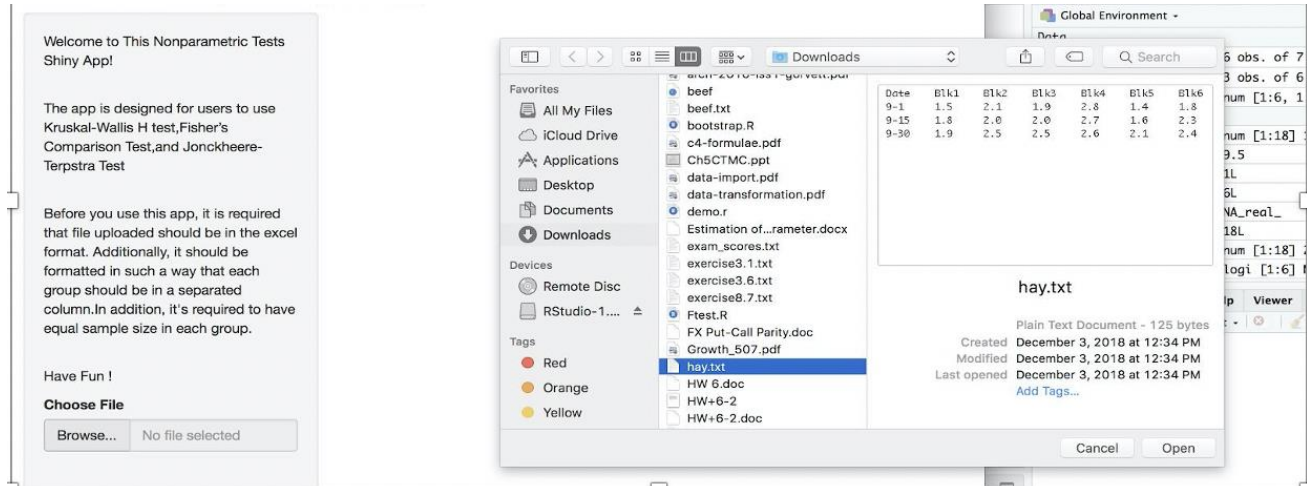
$$H_A: \tau_1 \leq \tau_2 \leq \dots \leq \tau_K$$

As per the other tests, the data should be formatted with the groups or treatment as the rows. The test first finds the number of observations in one group that are greater than the observations in all of the preceding groups. This is done for all groups, and that total will be JTobs. The test then creates a function that does the same thing, but for any data, and creates the JTstar, which is calculated the same way as JTobs. Lastly, based on your permutation size R selection, the test will permute the data R times, and finds all JTstars, which is the distribution of JT under the null distribution. The app also shows a histogram of the null distribution and where the JTobs lies on that distribution. The p-value is calculated by dividing the number of JTstars that are greater than or equal to JTobs by the number of permutations. The app displays this p-value.

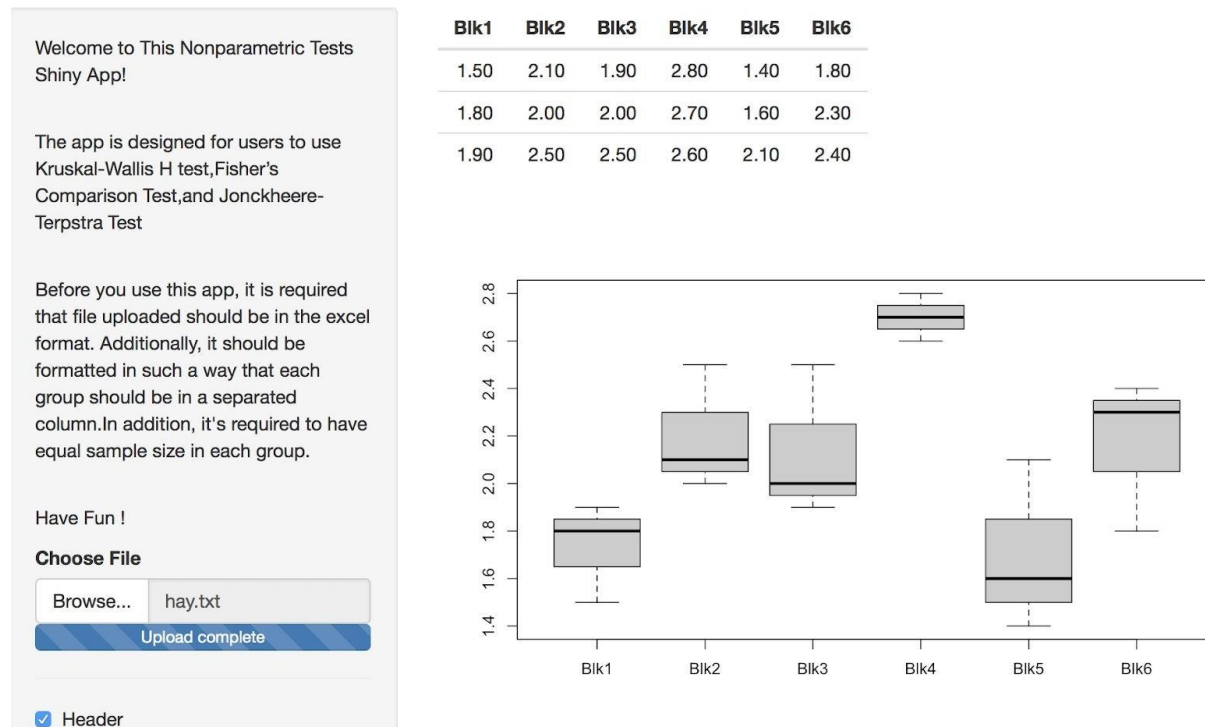
An applied example

File used: "hay.txt"

1: Upload the txt.file.



2: Select Separator, Quote, Display based on the data structure. We selected separator to be "tab" And " Double Quote"



☒ Header

Separator

☐ Comma

☐ Semicolon

☒ Tab

Quote

☐ None

☒ Double Quote

☐ Single Quote

Display

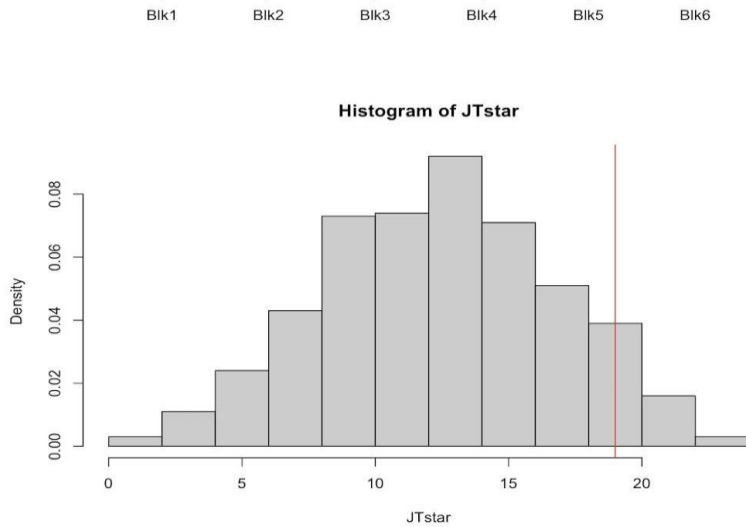
☒ Head

☐ All

Simulation Sample Size:

500 5,000

500 950 1,400 2,300 3,200 4,100 5,000



The p-value for the KW test is 0.206
The p-value for the JT test is 0.116

You are able to see the data, the boxplot of the data, the histogram of the JT star, the p value of JT test, and the p value of the KW test. As can be seen from the P value output, both p values are greater than alpha level 0.05, so we cannot reject the null hypothesis and we can conclude there is no significant mean differences and there is no significant ascending mean differences.

3. You could use the sample size slider to choose how many permutation samples size. Relevant graphs and P value outputs will change based on it.

Separator

☐ Comma

☐ Semicolon

☒ Tab

Quote

☐ None

☒ Double Quote

☐ Single Quote

Display

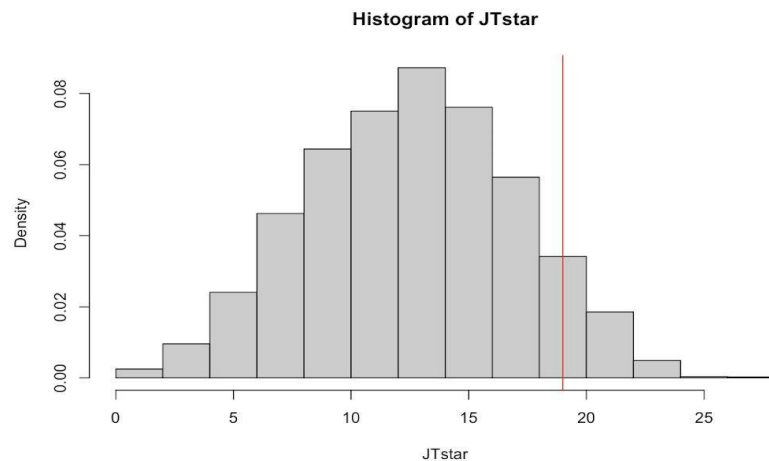
☒ Head

☐ All

Simulation Sample Size:

500 4,171 5,000

500 950 1,400 2,300 3,200 4,100 5,000



The p-value for the KW test is 0.203788060417166
The p-value for the JT test is 0.116518820426756