Test for ordered alternative in nonparametric location tests

Jonckheere-Terpstra Test

Tao Xu

1. Introduction

A parametric statistical model means the distribution of model involves several unknown constants which is called parameters. However, in statistics, non-parametric statistics (also called “distribution free statistics”) means the statistics do not assume data or population have any parameters (such as mean and variance) or characteristic structure (such as normal distribution) [1]. Non-parametric methods are often the only method available for data that simply specify order or counts of numbers of events or of individuals in various categories since those populations do not have normal distribution.

In the parametric test, we have one-sample t-test, two-sample t-test, and one way ANOVA. Same in the non-parametric test, we have Wilcoxon Signed-Rank test, Wilcoxon Rank-Sum test, and Kruskal-Wallis test correspondingly. The Kruskal-Wallis test is an omnibus test and is used to compare population location parameters among two or more groups based on independent samples. If a treatment represents, for example, steadily increasing doses of a stimulant we may want to test hypotheses about mean or median ($\theta_i$) for $H_0$: all $\theta_i$ are equal against $H_A$: $\theta_1 \leq \theta_2 \leq \theta_3 \leq \ldots \leq \theta_k$ with at least one of the inequalities is strict. In this case, a special case of Kruskal-Wallis test which is called Jonckheere-Terpstra test might be used.

2. Example of Jonckheere-Terpstra test

As I mentioned in the introduction, the hypothesis of Jonckheere-Terpstra test can be described as follow.

$H_0$: $\theta_1 = \theta_2 = \theta_3 = \ldots = \theta_k$

$H_A$: $\theta_1 \leq \theta_2 \leq \theta_3 \leq \ldots \leq \theta_k$ (with at least one strict inequality)

The way to calculate test statistics of Jonckheere-Terpstra test will be explained in the following example.

Example [2]:
In the 1989, Hinkley measured braking distances taken by motorists to stop when motorcycles were travelling at different speed. One subset of his data is shown in the table 1.

<table>
<thead>
<tr>
<th>Speed (mph)</th>
<th>Braking distances (feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>48</td>
</tr>
<tr>
<td>25</td>
<td>33 59 48 56</td>
</tr>
<tr>
<td>30</td>
<td>60 101 67</td>
</tr>
<tr>
<td>35</td>
<td>85 107</td>
</tr>
</tbody>
</table>

Table 1. Results of braking distances

Test statistics calculation

First, rank the braking distance from short to long (see table 2).

<table>
<thead>
<tr>
<th>Distance</th>
<th>33 48 48 56 59 60 67 85 101 107</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>1 2.5 2.5 4 5 6 7 8 9 10</td>
</tr>
</tbody>
</table>

Table 2. Rank of braking distance in the data set

Hypothesis of this case:

H$_0$: braking distance is independent of initial speed.

H$_A$: braking distance increases when initial speed increases.

The hypothesis seems to be very obvious since it is common sense that higher initial speed will cause longer braking distance. But let’s look at the result from Kruskal-Wallis test (table 3).

<table>
<thead>
<tr>
<th>Kruskal-Wallis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
</tr>
<tr>
<td>DF</td>
</tr>
<tr>
<td>Pr &gt; Chi-Square</td>
</tr>
</tbody>
</table>

Table 3. The test statistic and p-value of Kruskal-Wallis test

From the Kruskal-Wallis test, we cannot reject the null hypothesis (p-value is larger than 0.05) which means there is no evidence that increasing initial speed will cause increasing braking distance. This conclusion is not what we expected. Now, let’s try Jonckheere-Terpstra test and to see if there are any differences on results.

If there are k samples (in this case, it is 4), we can calculate the sum, U. $U_r$ is relevant to the rth sample ($r=1, 2, \ldots, k-1$) and any sample s for $s>r$. In Hinkley’s data, we will have
U_{12}, U_{13}, U_{14}, U_{23}, U_{24}, and U_{34}. To calculate the U_{rs}, we need get the sum of the number of sample s values that exceeds each sample r value. Take U_{12} as an example, 56 and 59 (group 2) is larger than 48 (group 1) which we count as 2. Also, 48 (group 2) is equal to 48 (group 1) which we can count 0.5 (same as in the Mann-Whitney statistic). So, U_{12} equals 2.5. Do the same thing on others, we got:

U_{13}=3; U_{14}=2; U_{23}=12; U_{24}=8; U_{34}=5.

Adding all U_{rs} together, we got U equals 32.5.

The test statistics in the Jonckheere-Terpstra test, Z, can be calculated by equation:

\[ Z = \frac{U - E(U)}{\sqrt{Var(U)}} \]

For large N and the individual n_i not too small, the distribution of test statistic Z is approximately standard normal.

\[ E(U) = \frac{N^2 - \sum n_i^2}{4} \]

\[ Var(U) = \frac{N^2(2N+3) - \sum [n_i^2(2n_i+3)]}{72} \]

Where N is the total sample size and n_i is the sample size in group i.

So, for our example, we can calculate:

\[ E(U) = \frac{10^2 - (1^2+4^2+3^2+2^2)}{4} = 17.5 \]

\[ Var(U) = \frac{10^2(2*10+3) - [1^2*(2*1+3)+4^2*(2*4+3)+3^2*(2*3+3)+2^2*(2*2+3)]}{72} = 27.9 \]

\[ Z = \frac{32.5 - 17.5}{\sqrt{27.9}} = 2.84 \]

From the z-score table, the one-tail p-value is 0.0023. Here, use the one tail p-value is because the hypothesis is directional.

The Jonckheere-Terpstra test shows the significant difference which means the increasing initial speed increases braking distance.

To prove the calculation I did manually, SAS is used. In the SAS, this test can be done by frequency table with “jt” function. The results of Jonckheere-Terpstra test by SAS are shown in the table 4.

| Jonckheere-Terpstra Test |
### Jonckheere-Terpstra Test

<table>
<thead>
<tr>
<th>Statistic</th>
<th>32.5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z )</td>
<td>2.8489</td>
</tr>
<tr>
<td>One-sided ( Pr &gt; Z )</td>
<td>0.0022</td>
</tr>
<tr>
<td>Two-sided ( Pr &gt;</td>
<td>Z</td>
</tr>
</tbody>
</table>

Table 4. The test statistic and p-value of Jonckheere-Terpstra test

Same test can be done in the R with “clinfun” package and it gives the same results.

R results

Jonckheere-Terpstra test
data:

\[ JT = 32.5, \text{ p-value} = 0.002263 \]

alternative hypothesis: increasing
data: speed and distance

Kruskal-Wallis chi-squared = 7.2599, df = 3, p-value = 0.06406

### 3. Discussion

From the example I showed above, we can see that Kruskal-Wallis test does not show the clear evidence of increasing speed increases braking distance. However, the Jonckheere-Terpstra does. The difference of results pointed out that when the increasing speed and increasing braking distance are relevant, the Jonckheere-Terpstra test is generally more powerful than Kruskal-Wallis test. The Jonckheere-Terpstra test has higher probability that the test will reject the null hypothesis when the alternative hypothesis is true for ordered alternatives.

### 4. References

5. Supplementary materials

*SAS code*

```sas
data breaking;
input speed distance @@;
cards;
20 48
25 33
25 59
25 48
25 56
30 60
30 101
30 67
35 85
35 107;
proc npar1way wilcoxon data=breaking;
class speed;
var distance;
title1 'Kruskal-Wallis Test';
run;
proc freq;
table speed*distance / jt nopercent nocol norow;
run;
```

*R code*

```r
> install.packages("clinfun")
> library(clinfun)
> speed20<-c(48)
> speed25<-c(33,59,48,56)
> speed30<-c(60,101,67)
> speed35<-c(85,107)
> pieces<-list(speed20, speed25, speed30, speed35)
> n<-c(1,4,3,2)
> grp<-as.ordered(factor(rep(1:length(n),n)))
> jonckheere.test(unlist(pieces),grp,alternative="increasing")
```