stcp-marquier-FriedmanR

> The following resources are associated:
> The R dataset 'Video.csv' and R script 'Friedman.R'

Friedman test in R
(Non-parametric equivalent to repeated measures ANOVA)

Dependent variable: Continuous (scale) but not normally distributed or ordinal

Independent variable: Categorical (Time/ Condition)

Common Applications: Used when several measurements of the same dependent variable are taken at different time points or under different conditions for each subject and the assumptions of repeated measures ANOVA have not been met. It can also be used to compare ranked outcomes.

Data: The dataset 'Video' contains some results from a study comparing videos made to aid understanding of a particular medical condition. Participants watched three videos (A, B, C) and one product demonstration (D) and were asked several Likert style questions about each. These were summed to give an overall score for each e.g. TotalAGen below is the total score of the ordinal questions for video A.

TOTALA_GEN	TOTALB_DOC	TOTALC_OLD	TOTALD_DEMO		RANKA_GEN	RANKB_DOC	RANKC_OLD	RANKD_DEMO
25	23	13	22		4	3	1	2
20	17	14	23		3	2	1	4
25	22	Friedman ranks each participants responses			2	1	3	
24	22	23	25		3	1	2	4

The Friedman test ranks each person's score from lowest to highest (as if participants had been asked to rank the methods from least favourite to favourite) and bases the test on the sum of ranks for each column. For example, person 1 gave C the lowest Total score of 13 and A the highest so SPSS would rank these as 1 and 4 respectively. As the raw data is ranked to carry out the test, the Friedman test can also be used for data which is already ranked e.g. the ranked example columns RANKA – RANKD. There should be one column per condition/ time point being compared containing the score or rank for that condition. If the Friedman test is significant, post hoc tests to locate which pairs are different are needed.

Research question: Which method is best for understanding the medical condition?
The dependent variable is the overall score and this is recorded in a separate variable for each method.

The sample sizes are small and normality checks showed that the assumption of normality had not been met so a repeated measures ANOVA is not suitable for comparing the groups. The **Friedman** test is the non-parametric equivalent to repeated measures ANOVA.

Steps in R
In order to import the file into R use the command
```
videoR<-read.csv('C:\\...\\stcp-Rdataset-Video.csv',header=T,sep=';')
```
attaching enables you to refer to the variables within the data set directly.
```
attach(videoR)
```

**Summary Statistics:** As we are carrying out a non-parametric test, use medians to compare the scores for the different methods.
To find the medians per category use the *median( )* command. The argument is the label of the variable you want to have the median from:
```
> median(TotalAGen)
[1] 24
> median(TotalBdoc)
[1] 23
> median(TotalCOld)
[1] 14
> median(TotalDDEMO)
[1] 22
```
The median scores show that Video C has a much lower median than the others. A box-plot is also useful for assessing differences:

```
boxplot(TotalAGen,TotalBdoc,TotalCOld,
TotalDDEMO,names=c("A","B","C","D"))
```
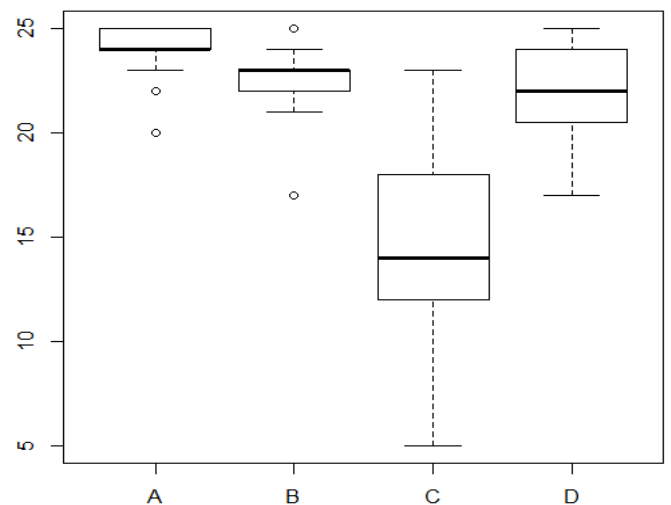
Carrying out the Friedman test
The format is similar to the ANOVA command. We use the `friedman.test()` command. The argument will contain the table of scores for the 4 videos.
We need to create a proper table (or matrix) containing our 4 samples of videos. We call that table "Sample". Each column of the table "Sample" will represent the scores for a certain type of video.
```
Sample<-matrix(c(TotalAGen,TotalBdoc,TotalCOld,TotalDDEMO),ncol=4)
```

The Friedman test is then applied to the table containing our 4 videos scores:

```
> friedman.test(Sample)

        Friedman rank sum test

data:  Sample
Friedman chi-squared = 41.3724, df = 3, p-value = 5.452e-09
```

Here the test statistic (41.3724) and degrees of freedom (3) are reported for the Friedman test. The Friedman test was significant (p = 5.452e-09 < 0.001) so the distributions of the scores for the videos being compared are different.

Pairwise Comparisons
To find out which pairs are different, you will need to install the package 'PMCMR' and load the library after that:
```
library(PMCMR)
```
You will then need to conduct the Nemenyi Post-hoc tests to compare all the pairs. They will be presented as a table.
```
posthoc.friedman.nemenyi.test(Sample)
```

Alternatively, you can carry out pairwise Wilcoxon tests (slightly less conservative) using the `wilcoxon.test()`. You should then adjust the results by multiplying by the number of pairs (Bonferroni correction). In our example, since we have 4 samples, we have 6 possible pairs. So we will need to multiply each p-value by 6:
```
wilcox.test(TotalAGen,TotalBdoc)$p.value*6
```
If the adjusted p-value is more than 1, then this p-value will take value 1 (since a p-value is a value between 0 and 1).

Reporting the results
A Friedman test was carried out to compare the total understanding scores for the four methods. There was found to be a significant difference between the methods, Nemenyi post hoc tests were carried out and there were significant differences between the Old video C and the Doctors video B (p < 0.001), the demonstration D (p <0.001) and video A (p<0.001). There were no significant differences between any other methods.