# Chapter 14
# Automatic Nonparametric Testing (30 Patients)

## 1 General Purpose

If your data are pretty complex and involve both repeated outcomes and different types of predictors including categorical ones, then multivariate methods (Chaps. 17 and 18) would be required for an overall analysis. However, with small samples, power is little, and an optimized univariate analysis testing the outcomes separately is an alternative. Automatic nonparametric testing chooses the best tests based on the data. Also, it takes account of nongaussian outcomes.

## 2 Schematic Overview of Type of Data File

| outcome 1 | outcome 2 | predictor 1 | predictor 2 | pre... |
|---|---|---|---|---|
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |



T.J. Cleophas, A.H. Zwinderman, *SPSS for Starters and 2nd Levelers*,
DOI 10.1007/978-3-319-20600-4_14

## 3 Primary Scientific Question

Can automatic nonparametric testing simultaneously assess the effect of multiple predictors including categorical ones on repeated outcomes and at the same account nonnormality in the outcomes.

## 4 Data Example

In a parallel-group study with three predictors (treatment 0, 1, and 2 correspondingly given to the groups 0, 1, and 2), and two continuous outcomes (hours of sleep and levels of side effects), assess whether the treatments are significantly different from one another.

| Outcome efficacy | Outcome side effect | Predictor gender | Predictor comorbidity | Predic.. group |
|---|---|---|---|---|
| 6,00 | 45,00 | ,00 | 1,00 | 0 |
| 7,10 | 35,00 | ,00 | 1,00 | 0 |
| 8,10 | 34,00 | ,00 | ,00 | 0 |
| 7,50 | 29,00 | ,00 | ,00 | 0 |
| 6,40 | 48,00 | ,00 | 1,00 | 0 |
| 7,90 | 23,00 | 1,00 | 1,00 | 0 |
| 6,80 | 56,00 | 1,00 | 1,00 | 0 |
| 6,60 | 54,00 | 1,00 | ,00 | 0 |
| 7,30 | 33,00 | 1,00 | ,00 | 0 |
| 5,60 | 75,00 | ,00 | ,00 | 0 |

Only the first ten patients are shown. The entire data file is in extras.springer.com and is entitled "chap14automaticnonparametrictesting". Automatic nonparametric tests is available in SPSS 18 and up. Start by opening the above data file.

## 5 Automatic Nonparametric Testing

For analysis the statistical model Independent Samples in the module Nonparametric Tests is required.

Command:

Analyze. . ..Nonparametric Tests. . ..Independent Samples. . ..click Objective. . .. mark Automatically compare distributions across groups. . ..click Fields. . ..in Test fields: enter "hours of sleep" and "side effect score". . ..in Groups: enter "group". . ..click Settings. . ..Choose Tests. . ..mark "Automatically choose the tests based on the data". . ..click Run.

In the interactive output sheets the underneath table is given. Both the distribution of hours of sleep and side effect score are significantly different across the three

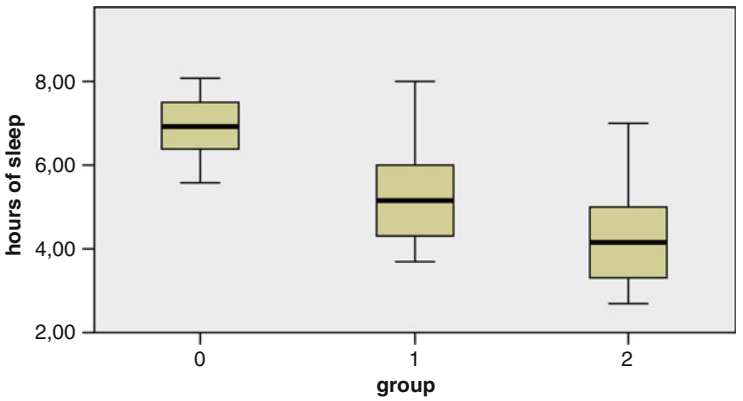
categories of treatment. By double-clicking the table you will obtain an interactive set of views of various details of the analysis, entitled the Model Viewer.

Hypothesis test summary

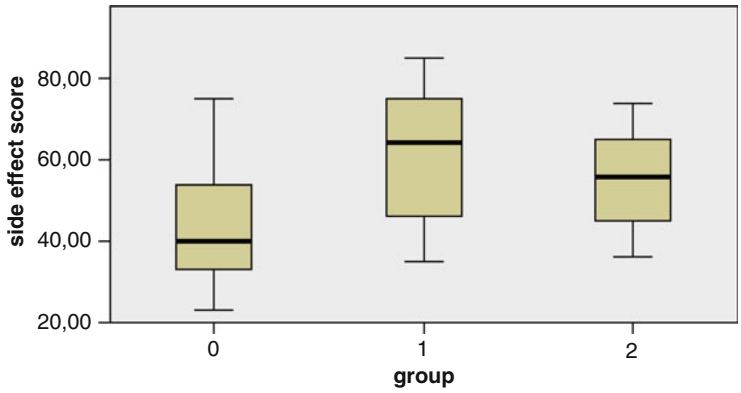
| | Null hypothesis | Test | Sig. | Decision |
|---|---|---|---|---|
| 1 | The distribution of hours of sleep is the same across categories of group. | Independent-samples Kruskal-Wallis test | ,001 | Reject the null hypothesis. |
| 2 | The distribution of side effect score is the same across categories of group. | Independent-samples Kruskal-Wallis test | ,036 | Reject the null hypothesis. |

Asymptotic significances are displayed. The significance level is, 05

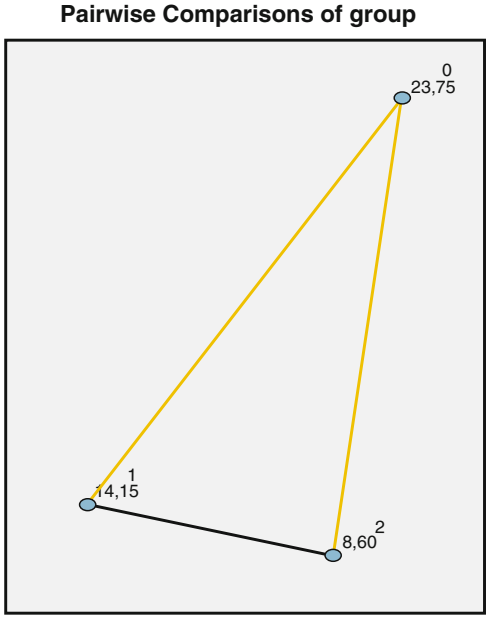
One view provides the box and whiskers graphs (medians, quartiles, and ranges) of hours of sleep of the three treatment groups. Group 0 seems to perform better than the other two, but we don't know where the significant differences are.




Also the box and whiskers graph of side effect scores is given. Some groups again seem to perform better than the other. However, we cannot see whether 0 vs 1, 1 vs 2, and /or 0 vs 2 are significantly different.




In the view space at the bottom of the auxiliary view (right half of the Model Viewer) several additional options are given. When clicking Pairwise Comparisons, a distance network is displayed with yellow lines corresponding to statistically significant differences, and black ones to insignificant ones. Obviously, the differences in hours of sleep of group 1 vs (versus) 0 and group 2 vs 0 are statistically significant, and 1 vs 2 is not. Group 0 had significantly more hours of sleep than the other two groups with p = 0,044 and 0,0001.



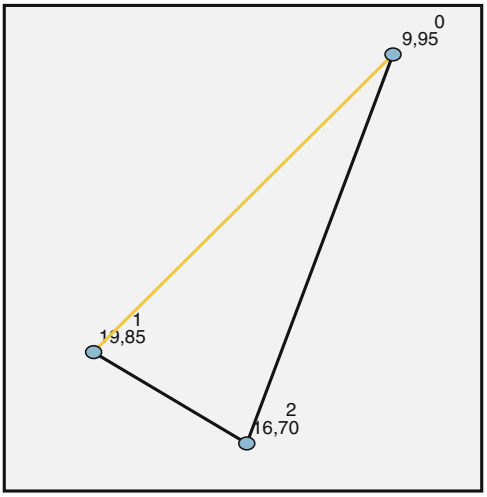
Each node shows the sample average rank of group.

| Sample1-Sample2 | | Test Statistic | Std. Error | Std. Test Statistic | Sig. | Adj.Sig. |
|---|---|---|---|---|---|---|
| 2- | 1 | 5,550 | 3,936 | 1,410 | ,158 | ,475 |
| 2- | 0 | 15,150 | 3,936 | 3,849 | ,000 | ,000 |
| 1- | 0 | 9,600 | 3,936 | 2,439 | ,015 | ,044 |

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is, 05.

As shown below, the difference in side effect score of group 1 vs 0 is also statistically significant, and 1 vs 0, and 1 vs 2 are not. Group 0 has a significantly better side effect score than the 1 with p = 0,035, but group 0 vs 2 and 1 vs 2 are not significantly different.

**Pairwise Comparisons of group**




Each node shows the sample average rank of group.

| Sample1-Sample2 | | Test Statistic | Std. Error | Std. Test Statistic | Sig. | Adj.Sig. |
|---|---|---|---|---|---|---|
| 0- | 2 | -6,750 | 3,931 | -1,717 | ,086 | ,258 |
| 0- | 1 | -9,900 | 3,931 | -2,518 | ,012 | ,035 |
| 2- | 1 | -3,150 | 3,931 | ,801 | ,423 | 1,000 |

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is, 05.

## 6 Conclusion

If your data are pretty complex and involve both repeated outcomes and different types of predictors including categorical ones, then multivariate methods (Chaps. 17 and 18) would be required for an overall analysis. However with small samples power is little, and an optimized univariate analysis testing the outcomes separately is an alternative. Automatic nonparametric testing chooses the best tests based on the data. Also it takes account of nongaussian outcomes. If you wish to report the above data as a whole, then Bonferroni adjustments for multiple testing should be performed (Statistics applied to clinical studies 5th edition, Chaps. 8 and 9, Springer Heidelberg Germany, 2012, from the same authors).

## 7 Note

More background theoretical and mathematical information of the Kruskal-Wallis test is given in Statistics applied to clinical trials 5th edition, Chap. 2, Springer Heidelberg, 2012, from the same authors.