Chapter 9
Repeated Measures Analysis of Variance, Friedman (10 Patients)

1 General Purpose

Just like paired t-tests (Chap. 2), repeated-measures-analysis of variance (ANOVA) can assess data with more than a single continuous outcome. However, it allows for more than two continuous outcome variables. It is, traditionally, used for comparing crossover studies with more than two treatment modalities.

2 Schematic Overview of Type of Data File

Outcome 1	outcome 2	outcome 3
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.

The above repeated-measures-ANOVA does not include predictor variables, and the effects of a predictor on the outcomes can, therefore, not be assessed. Instead significances of differences between the paired observations can be tested. Gaussian frequency distributions of the outcomes are assumed.

T.J. Cleophas, A.H. Zwinderman, *SPSS for Starters and 2nd Levelers*,
DOI 10.1007/978-3-319-20600-4_9

3 Primary Scientific Question

Do three different pills produce significantly different clinical outcome effects.

4 Data Example

In a crossover study of three different sleeping pills the significance of difference between hours of sleep between the different treatments was assessed.

Hours of sleep after sleeping pill

one	two	three
6,10	6,80	5,20
7,00	7,00	7,90
8,20	9,00	3,90
7,60	7,80	4,70
6,50	6,60	5,30
8,40	8,00	5,40
6,90	7,30	4,20
6,70	7,00	6,10
7,40	7,50	3,80
5,80	5,80	6,30

5 Analysis, Repeated Measures ANOVA

The data file is in extras.springer.com, and is entitled "chapter9repeatedmeasuresanova". Open the data file in SPSS. For analysis the module General Linear Model is required. It consists of 4 statistical models:

Univariate,
Multivariate,
Repeated Measures,
Variance Components.
We will use here Repeated Measures.

Command:
Analyze....General Linear Model....Repeated Measures....Repeated Measures Define Factors....Within-subject Factor name: treat....Number of Levels: 3....click Add....click Define: Within-Subjects Variables (treat): enter treatmenta, treatmentb, treatment3....click OK.

The output sheets show a series of tables starting with the multivariate tests table. This is to check the correlation of the predictors that are transiently made

dependent. The nullhypothesis is no significance of difference between the repeated measures.

Mauchlys Test of Sphericity[a]

Measure:MEASURE 1							
					Epsilon[b]		
Within subjects effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
treat	,096	18,759	2	,000	,525	,535	,500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix

[a]Design: Intercept within subjects design: treat

[b]Maybe used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the tests of within-subjects effects table

Tests of within-subjects effects

Measure:MEASURE 1						
Source		Type III sum of squares	df	Mean square	F	Sig.
treat	Sphericity assumed	24,056	2	12,028	10,639	,001
	Greenhouse-Geisser	24,056	1,050	22,903	10,639	,009
	Huynh-Feldt	24,056	1,070	22,489	10,639	,008
	Lower-bound	24,056	1,000	24,056	10,639	,010
Error (treat)	Sphericity assumed	20,351	18	1,131		
	Greenhouse-Geisser	20,351	9,453	2,153		
	Huynh-Feldt	20,351	9,627	2,114		
	Lower-bound	20,351	9,000	2,261		

The repeated-measures ANOVA tests whether a significant difference exists between three treatments. An important criterion for validity of the test is the presence of sphericity in the data, meaning that all data come from Gaussian distributions. It appears from the above upper table that this is not true, because based on this table we are unable to reject the null-hypothesis of non-sphericity. This means that an ANOVA test corrected for non-sphericity has to be performed. There are three possibilities: the Greenhouse, Huynh, and Lower-bound methods. All of them produce a much larger p-value than the uncorrected method, but the result is still statistically highly significant with p-values of 0,009, 0,008, and 0,010. A significant difference between the treatments has, thus, been demonstrated. However, we do not yet know whether the significant difference is located between the treatments 1 and 2, between the treatments 1 and 3, or between the treatments 2 and 3. In order to find out three separate paired t-tests have to be performed. Note, that with multiple t-tests it is better to reduce the cut-off level for statistical

significance to approximately 0.01 (more information about the adjustments for multiple testing including the Bonferroni procedure is given in the textbook "Statistics applied to clinical trials", 5th edition, the Chaps. 8 and 9, 2012, Springer Heidelberg Germany, from the same authors).

6 Alternative Analysis: Friedman Test

If the outcome data do not follow Gaussian patterns, or if your data are pretty small, it will be more safe to perform a test, that allows for nonnormal data. The Friedman test is adequate, but can also be applied with normal data. So, it is an excellent choice, either way. For analysis the statistical model K Related Samples in the module Nonparametric Tests is required.

Command:
Analyze....NonparametricTests....Legacy Dialogs....K Related Samples.... Test Variables: enter treatmenta, treatmentb, treatmentc....Mark: Friedman.... click OK.

Test statistics[a]

N	10
Chi-Square	7,579
df	2
Asymp. Sig.	,023

[a]Friedman test

The result is significant, but the p-value is markedly larger than the p-value of the ANOVA, i.e., 0,023. Just like with the above ANOVA we will have to perform additional tests to determine, where the difference of the three treatments is located. For that purpose three Wilcoxon's tests could be performed (and adjustment for multiple testing can be done similarly to the above procedure: using either a p-value of 0,01 or a Bonferroni adjustment, see textbook "Statistics applied to clinical studies", the Chaps. 8 and 9, 5th edition, 2012, Springer Heidelberg Germany, from the same authors).

7 Conclusion

In a crossover study of multiple different treatment modalities the significance of difference between the outcomes of the different treatments can be tested with repeated-measures ANOVA. The test result is an overall result, and does not tell you where the difference is. E.g., with three treatments it may be a difference between treatment 1 and 2, 2 and 3, or 1 and 3 or some combination of these three

possibilities. In order to find out where it is additional paired t-tests or Wilcoxon tests adjusted for Bonferroni inequalities have to be performed, and one might consider to skip the overall tests and start with the paired t-tests or Wilcoxon tests from the very beginning.

8 Note

More background, theoretical and mathematical information of repeated measures ANOVA is given in Statistics applied to clinical studies 5th edition, Chap. 2, Springer Heidelberg Germany, 2012, from the same authors.