Application and Procedure of Repeated Measures ANOVA in Statistics with SPSS

Somanpreet Singh, Mohammad Muqarram
Asst. Prof., Department of Physical Education, Central University of Kashmir, Ganderbal, J & K, India

Abstract

A repeated measure is a term used when the same entities take part in all conditions of an experiment. Sphericity is the condition where the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of the independent variable) are equal. Sphericity is met when these variances are roughly equal. The effect of violating sphericity is a loss of power (i.e. an increased probability of a Type II error) and a test statistic (F- ratio) that simply cannot be compared to tabulated values of the F-distribution (Field, 2009; 2013). Hypothetical data to illustrate the calculation of the variance of the differences between conditions are described. The description of data entering followed by the different steps for the analysis has also been described in the research work with a hypothetically illustration. The analysis processed has been completed were followed by the description of the output.

KEYWORDS: Repeated Measures, ANOVA, SPSS etc

Repeated measures design is a research design that involves multiple measures of the same variable taken on the same or matched subjects either under different conditions or over two or more time periods. For instance, repeated measurements are collected in a longitudinal study in which change over time is assessed. A repeated measure is a term used when the same entities take part in all conditions of an experiment. So, for example, a researcher might want to test the effects of different durations of playing the recreational activities on the enjoyment level of the students. In this type of experiment it is important to control for individual differences in tolerance to the participation in the physical activities with different durations: some students can play for a longer duration without really feeling the fatigue, whereas others starts to feel bored and fatigue after a particular duration. To control for these individual differences researcher actually test the whole populations in all conditions of the experiment, so researcher would test each subject after their involvement in recreational activities with different durations (i.e. 30 minutes of hours of recreational activities, 45 minutes of recreational activities and 60 minutes of recreational activities). After each duration the participant might be given a questionnaire assessing their enjoyment of the recreational activities. Therefore, every participant provides a score representing their enjoyment before the study (no recreational activities), after 30 minutes of recreational activities, after 45 minutes of recreational activities, and so on. This design is said to use repeated measures.

Further, In statistical analysis, all parametric tests assume some certain characteristic about the data, also known as assumptions. Violation of these assumptions changes the conclusion of the research and interpretation of the results. Sphericity is an important assumption of repeated-measures ANOVA. It is the condition where the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of the independent variable) are equal.

What is Sphericity?

Generally it is seen that parametric tests based on the normal distribution assume that data points are independent. This is not the case in a repeated measures design because data for different conditions have come from the same entities. This means that data from different experimental conditions will be related; because of this researcher have to make an additional assumption to those of the independent ANOVAs. Put simply (and not entirely accurately), researcher assume that the relationship between pairs of experimental conditions is similar (i.e. the level of dependence between pairs of groups is roughly equal). This assumption is known as the assumption of sphericity.

The more accurate but complex explanation is as follows. Table 1 shows data from an experiment with three conditions as the example cited above as the involvement of students in the recreational activities with different durations (i.e. 30 minutes of hours of recreational activities, 45 minutes of recreational activities and 60 minutes of recreational activities).. Imagine the researcher calculated the differences between pairs of scores in all combinations of the treatment levels. Having done this, the researcher calculated the variance of these differences. Sphericity is met when these variances are roughly equal. In these data there is some deviation from sphericity because the variance of the differences between conditions A and B is greater than the variance of the differences between A and C and between B and C. However, these data have local circularity (or local sphericity) because two of the variances of differences are very similar.

Table 1: Hypothetical data to illustrate the calculation of the variance of the differences between conditions

Condition A	Condition B	Condition C	A−B	A−C	B−C
10	12	8	−2	2	4
15	15	12	0	3	3
25	30	20	−5	5	10
35	30	28	5	7	2
30	27	20	3	10	7
		Variance:	15.7	10.3	10.7

Pattern of the data and experimental protocol

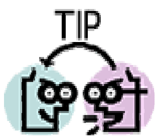
Subjects	Time/Condition		
	T_1	T_2	T_3
S_1	S_1	S_1	S_1
S_2	S_2	S_2	S_2
S_3	S_3	S_3	S_3
S_4	S_4	S_4	S_4
S_5	S_5	S_5	S_5
S_6	S_6	S_6	S_6

Furthermore, some of the literature quoted that what is the effect of violating the Assumption of Sphericity?

The effect of violating sphericity is a loss of power (i.e. an increased probability of a Type II error) and a test statistic (F- ratio) that simply cannot be compared to tabulated values of the F-distribution (Field, 2009; 2013).

Assessing the Severity of Departures from Sphericity. Departures from sphericity can be measured in three ways: Greenhouse and Geisser (1959) and Huynh and Feldt (1976).

The Lower Bound estimate (the lowest possible theoretical value for the data) The Greenhouse-Geisser and Huynh-Feldt estimates can both range from the lower bound (the most severe departure from sphericity possible given the data) and 1 (no departure from sphjercitiy at all). Field (2013) or Girden (1992).While in the real world setting or calculating the collected data from an experiment, SPSS also produces a test known as Mauchly's test, which tests the hypothesis that the variances of the differences between conditions are equal.



- ❖ If Mauchly's test statistic is significant (i.e. has a probability value less than .05) we conclude that there are significant differences between the variance of differences: the condition of sphericity has not been met.
- ❖ If, Mauchly's test statistic is nonsignificant (i.e. $p > .05$) then it is reasonable to conclude that the variances of differences are not significantly different(i.e.theyareroughlyequal).
- ❖ If Mauchly's test is significant then we cannot trust the *F*-ratios produced by SPSS.
- ❖ Remember that, as with any significance test, the power of Mauchley's test depends on the sample size. Therefore, it must be interpreted within the context of the sample size because:
- ❖ In small samples large deviations from sphericity might be deemed non- significant.

One-Way Repeated Measures ANOVA using SPSS

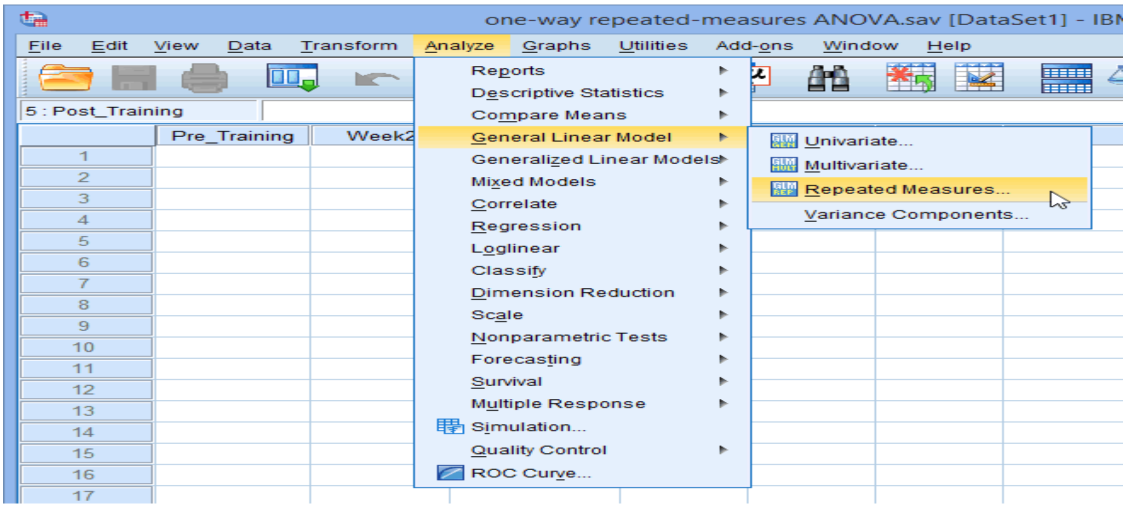
Entering the Data

From the above said example which a researcher might want to test the effects of different durations of playing the recreational activities on the enjoyment level of the students. So, researcher would test each subject after their involvement in recreational activities with different durations (i.e. 30 minutes of hours of recreational activities, 45 minutes of recreational activities and 60 minutes of recreational activities).

The independent variable is the different durations of recreational activities and the dependent variable is the enjoyment level of the student.

To conduct an ANOVA using a repeated measures design, activate the define factors dialog box by selecting Analyze, General Linear Model Repeated Measures...

.



In the Define Factors dialog box (Figure 2), a name should be entered for the within-subject (repeated-measures) variable. If the design has several repeated measures variables then researcher can add more factors to the list.

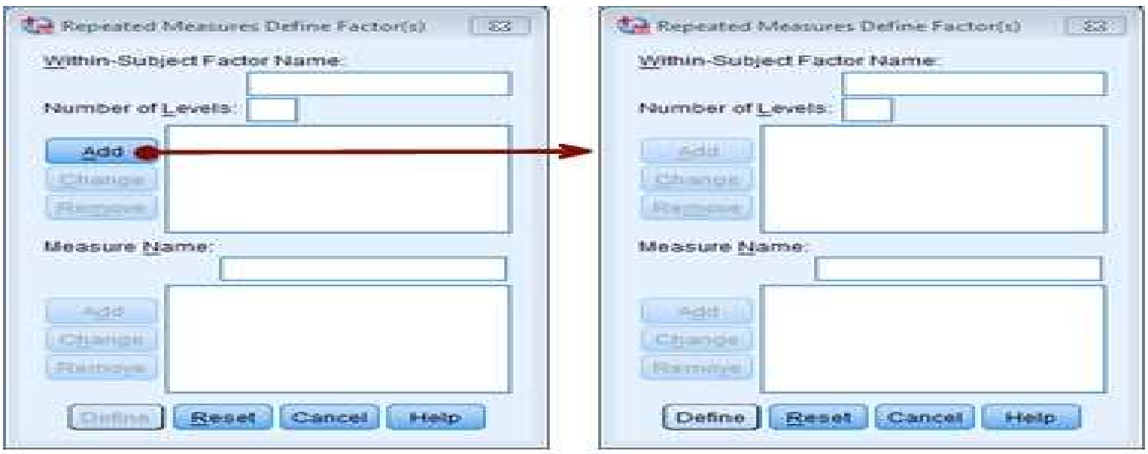


Figure 2: Define Factors dialog box for repeated measures ANOVA

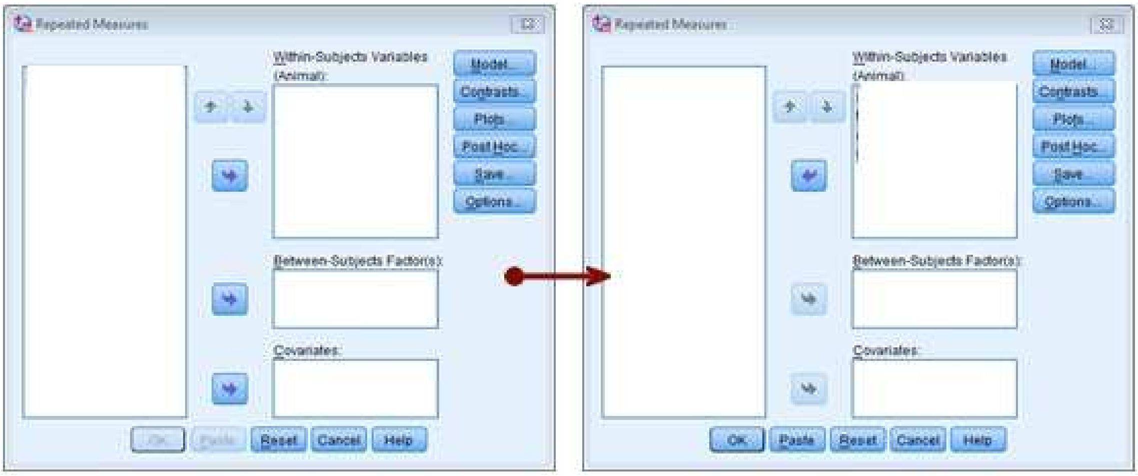


Figure 3: Main dialog box for repeated measures ANOVA

The main dialog box (Figure 3) has a space labelled within subjects variable list that contains a list of 4 question marks proceeded by a number. These question marks are for the variables representing the 4 levels of the independent variable. The variables corresponding to these levels should be selected and placed in the appropriate space. It has only 4 variables in the data editor, so it is possible to select all four variables at once (by clicking on the variable at the top, holding the mouse button down and dragging down over the other variables). The selected variables can then be transferred by dragging them or clicking on .

When all four variables have been transferred, researcher can select various options for the analysis. There are several options that can be accessed with the buttons at the bottom of the main dialog box. These options are similar to the ones, have already encountered.

Post Hoc Tests

There is no proper facility for producing post hoc tests for repeated measures variables in SPSS. However, It can produce a basic set of post hoc tests clicking Options... in the main dialog box. To specify post hoc tests, select the repeated measures variablefrom the box labelled Estimated Marginal Means: Factor(s) and Factor Interactions and transfer it to the box labelled Display Means for by clicking on  (Figure 4).

Once a variable has been transferred, the box labelled Compare main effects (☑ Compare main effects) becomes active and you should select this option. If this option is selected, the box labelled Confidence interval adjustment becomes active and click on LSD(none) to see a choice of three adjustment levels. The default is to have no adjustment and simply perform a Tukey LSD post hoc test (this is not recommended). The second option is a Bonferronicorrection and the final option is a Sidak correction, which should be selected if you are concerned about the loss of power associated with Bonferroni corrected values. When It has to be selected the options of interest, click on Continue to return to the main dialog box, and then click on OK to run the analysis.

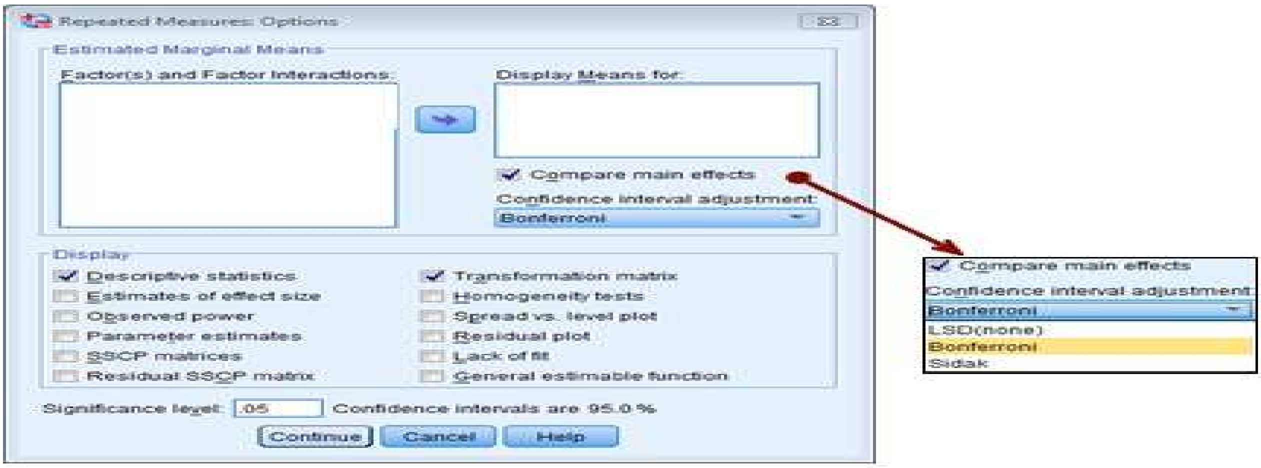


Figure 4: Options dialog box

OUTPUT FOR REPEATED MEASURES ANOVA

Output 1 Descriptive statistics and other Diagnostics

Descriptive Statistics

	Mean	Std. Deviation	N
A	8.13	2.232	8
B	4.25	1.832	8
C	4.13	2.748	8
D	5.75	2.915	8

Output 1 shows two tables. The first lists the variables that represent each level of the independent variable, which is useful to check that the variables were entered in the correct order. It provides basic descriptive statistics for the four levels of the independent variable.

Output 2Mauchly's test

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Enjoment level	.136	11.406	5	.047	.533	.666	.333

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: Animal

b. May be 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.

Output 2 shows Mauchly's test for these data, and the important column is the one containing the significance value. The significance value is .047, which is less than .05, so it must accept the hypothesis that the variances of the differences between levels were significantly different. In other words the assumption of sphericity has been violated. We could report Mauchly's test for these data as the Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(5) = 11.41$, $p = .047$.

Output 3 The Main ANOVA

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Enjoyment	Sphericity Assumed	83.125	3	27.708	3.794	.026
	Greenhouse-Geisser	83.125	1.599	52.001	3.794	.063
	Huynh-Feldt	83.125	1.997	41.619	3.794	.048
	Lower-bound	83.125	1.000	83.125	3.794	.092
Error(Enjoyment)	Sphericity Assumed	153.375	21	7.304		
	Greenhouse-Geisser	153.375	11.190	13.707		
	Huynh-Feldt	153.375	13.981	10.970		
	Lower-bound	153.375	7.000	21.911		

Output 3 shows the results of the ANOVA for the within-subjects variable. The above table look slightly different, but it has represented in a simplified a bit. The above table can be read much the same as for One-way independent ANOVA. The significance of *F* is .026, which is significant because it is less than the criterion value of .05. We can, therefore, conclude that there was a significant difference in the enjoyment level in different time durations.

Output 4 Post Hoc Tests

Given the main effect was not significant, It should not follow this effect up with post hoc tests.

Pairwise Comparisons

Measure: MEASURE_1

Duration	Duration (J)	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	3.875*	.811	.012	.925	6.825
	3	4.000*	.732	.006	1.339	6.661
	4	2.375	1.792	1.000	-4.141	8.891
2	1	-3.875*	.811	.012	-6.825	-.925
	3	.125	1.202	1.000	-4.244	4.494
	4	-1.500	1.336	1.000	-6.359	3.359
3	1	-4.000*	.732	.006	-6.661	-1.339
	2	-.125	1.202	1.000	-4.494	4.244
	4	-1.625	1.822	1.000	-8.249	4.999
4	1	-2.375	1.792	1.000	-8.891	4.141
	2	1.500	1.336	1.000	-3.359	6.359
	3	1.625	1.822	1.000	-4.999	8.249

Based on estimated marginal means

*. The mean difference is significant at the

b. Adjustment for multiple comparisons: Bonferroni.

However, just to illustrate how it would interpret the SPSS, the difference between group means is displayed, the standard error, the significance value and a confidence interval for the difference between means. By looking at the significance values it is cleared that the only significant differences between the groups.

Reference

Barnard, G. A. (1963). Ronald Aylmer Fisher, 1890–1962: Fisher's contributions to mathematical statistics. Journal of the Royal Statistical Society, Series A (General), 126, 162–166.

Bock, R. D. (1975).Multivariate statistical methods in behavioural research. New York: McGraw-Hill.

Boik, R. J. (1981). A priori tests in repeated measures designs:Effects of nonsphericity. Psychometrika, 46(3), 241–255.

Bowerman, B. L., & O'Connell, R. T. (1990). Linear statistical models: An applied approach (2nd ed.). Belmont, CA:Duxbury.

Field, A. P. (1998). A bluffer's guide to sphericity. Newsletter of the mathematical, statistical and computing section of the British Psychological Society, 6(1)13–22.

Field, A. P. (2000). Discovering statistics using SPSS for Windows: Advanced techniques for the beginner. London: Sage.

Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. Journal of Educational Statistics, 5(3), 269–287.

Mitzel, H. C., & Games, P. A. (1981). Circularity and multiple comparisons in repeated measures designs. British Journal of Mathematical and Statistical Psychology, 34, 253–259.

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 69, 894–908.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. Psychological Bulletin, 83, 579–586.

Olson, C. L. (1979). Practical considerations in choosing a MANOVA test statistic: A rejoinder to Stevens. Psychological Bulletin, 86, 1350–1352.

Pearson, E. S., & Hartley, H. O. (1954).Biometrika tables for statisticians, volume I. New York: Cambridge University Press.

Pearson, K. (1894). Science and Monte Carlo. The Fort nightly Review, 55, 183–193.

Stevens, J. P. (1980). Power of the multivariate analysis of variance tests. Psychological Bulletin, 88, 728–737.

Stevens, J. P. (2002). Applied multivariate statistics for the social sciences (4th ed.). Hillsdale, NJ: Erlbaum.

Verma J.P. and Ghufran Mohammad (2012), Statistics for Psychology-A Comprehensive Text,New Delhi:McGraw Hill Publication,p.450

Zabell, S. L. (1992). R. A. Fisher and fiducial argument. Statistical Science, 7(3), 369–387.