Repeated Measures ANOVA

Repeated measures ANOVA, also known as within-subjects ANOVA, are an extension of Paired T-Tests. Like T-Tests, repeated measures ANOVA gives us the statistic tools to determine whether or not changed has occurred over time. T-Tests compare average scores at two different time periods for a single group of subjects. Repeated measures ANOVA compared the average score at multiple time periods for a single group of subjects.

The first step in solving repeated measures ANOVA is to combine the data from the multiple time periods into a single time factor for analysis. The different time periods are analogous to the categories of the independent variable is a one-way analysis of variance. The time factor is then tested to see if the mean for the dependent variable is different for some categories of the time factor. If the time factor is statistically significant in the ANOVA test, then Bonferroni pair wise comparisons are computed to identify specific differences between time periods.

For a problem where the dependent variable is measured at three time periods, there are three paired comparisons: time 1 versus time 2, time 2 versus time 3, and time 1 versus time 3. These three paired comparisons can produce eight possible combinations which we use to interpret the results of the analysis.

<table>
<thead>
<tr>
<th>Pattern 1</th>
<th>Time 1 versus Time 2</th>
<th>Time 2 versus Time 3</th>
<th>Time 1 versus Time 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern 2</td>
<td>Significant difference</td>
<td>Significant difference</td>
<td>Non-significant difference</td>
</tr>
<tr>
<td>Pattern 3</td>
<td>Significant difference</td>
<td>Non-significant difference</td>
<td>Significant difference</td>
</tr>
<tr>
<td>Pattern 4</td>
<td>Significant difference</td>
<td>Non-significant difference</td>
<td>Non-significant difference</td>
</tr>
<tr>
<td>Pattern 5</td>
<td>Non-significant difference</td>
<td>Significant difference</td>
<td>Significant difference</td>
</tr>
<tr>
<td>Pattern 6</td>
<td>Non-significant difference</td>
<td>Significant difference</td>
<td>Non-significant difference</td>
</tr>
<tr>
<td>Pattern 7</td>
<td>Non-significant difference</td>
<td>Non-significant difference</td>
<td>Significant difference</td>
</tr>
<tr>
<td>Pattern 8</td>
<td>Non-significant difference</td>
<td>Non-significant difference</td>
<td>Non-significant difference</td>
</tr>
</tbody>
</table>

Stating the finding for each of these patterns clearly and succinctly is a challenge. We can facilitate this task by visualizing what each of these patterns looks like in a chart, such as the ones below. It is customary to show change over time as a line chart with time on the horizontal axis and the average measure on the vertical axis. Each point represents a time period, and the lines connecting the point show the direction of the relationship.

These charts presume that a difference of 1.0 or more between the means two time periods is statistically significant. Our interpretation of this difference will be characterized as a “substantial” difference. If the difference between the means for two time period is less than 1.0, we will presume that it is not statistically significant and will be characterized as a “minimal” difference. The red vertical error bars help us evaluate the differences between time periods. If the error bars overlap, the difference is less than one. If the error bars do not overlap, the difference between time periods is 1.0 or greater.

The lines between the first time period and the second and between the second time period and the third, can either be in the same direction (both increasing or decreasing) or they can be in opposite directions (one increasing and the other decreasing). Whether the direction changes or not will have an impact on the overall pattern from time 1 to time 3.
Pattern 1: Substantial changes between all time periods

In this chart, the change from time 1 to time 2 is significant, as is the change from time 2 to time 3. If both of these changes are significant and in the same direction (i.e. the mean at time 3 is less than the mean at time 2 which is less than the mean at time 1, the change from time 1 to time 3 has to be significant as well.

![Graph showing significant changes T1-T2, T2-T3, T1-T3.]

In my interpretation, I would consider the change from time 1 to time 3 to be redundant, providing little additional information, so I would only mention the changes from time 1 to time 2 and from time 2 to time 3.

The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time one to time two as well as from time two to time three.*

Pattern 2: Substantial change in consecutive time periods in opposite directions

The only way to have the change from time 1 to time 3 be non-significant when the changes from time 1 to time 2 and time 2 to time 3 are significant is to change the direction of the relationship.

![Graph showing significant changes T1-T2, T2-T3, with T1-T3 not significant.]

If the direction of the change from time 2 to time 3 is opposite from the direction of change from time 1 to time 2, the overall pattern from time 1 to time 3 may not be significant.
The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time one to time two, but this was offset by a substantial increase from time two to time three*

**Pattern 3: Substantial change followed by minimal change in same direction**

If the change from the first time period to both the second and third time periods is significant, but the change from the second time period to the third is not significant, the combined pattern is shown in the following chart. As long as the change from the second time period to the third time period is in the same direction as the change from time 1 to time 2, it will be larger than the change from the first to the second, and therefore, statistically significant. We do not necessarily need to include the information about the change from time 1 to time 3 in our interpretation.

![Pattern 3 Chart](image1)

The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time one to time two, but was followed by a minimal decrease from time two to time three*

**Pattern 4: Substantial change followed by minimal change in opposite direction**

In the next chart, only the change from the first time period to the second is significant. If the change from time 2 to time 3 did not change direction, the change from time 1 to time 3 would be even larger and therefore would also be statistically significant.

![Pattern 4 Chart](image2)
The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time one to time two, followed by a minimal increase from time two to time three*

**Pattern 5: Minimal change followed by substantial change in same direction**

Thus far, the minimal change has occurred in the second time period. It can also appear in the first time period, and we can think of the significant change not appearing until the second time period as a delayed reaction. If time 2 to time 3 is a significant change, and if the change between the two preceding time periods is the same direction, the overall pattern will also be significant.

**Significant Change: T1-T3, T2-T3**
**T1-T2 Not Significant**

The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time two to time three, following a minimal decrease from time one to time two*

**Pattern 6: Minimal change followed by substantial change in opposite direction**

If the change in the first time period were minimal and we wanted to depict significant change from time two to time three without significant change in the overall time period from time one to time two, the direction of change for the two periods would have to be in the opposite direction.

**Significant Change: T2-T3**
**T1-T2, T1-T3 Not Significant**
The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time two to time three, following a minimal increase from time one to time two.*

**Pattern 7: Minimal change between time periods, but collectively substantial**

If the change from time one to time three were significant, but neither of the intervening time periods were significant, we would have the pattern shown in the following chart. This could represent the situation where change occurs slowly, and the change from time 1 to time 2 and from time 2 to time 3 was too small to make an individual substantial contribution, but collectively they made a difference.

![Pattern 7 Chart]

The pattern of changes can be summarized with a statement like the following: *The average score decreased substantially from time one to time three, but the changes from both time one to time two and from time two to time three were minimal.*

**Pattern 8: No change**

The final possibility is that the changes were minimal whether viewed individually or collectively.

![Pattern 8 Chart]

The pattern of changes can be summarized with a statement like the following: *The average score did not change across the three time periods at time one, time two, and time three.*
The Requirements for Repeated Measures ANOVA

Level of Measurement and Sample Size

In these problems, the repeated measures ANOVA uses a single measure taken at different points in time. For the OMAHA.SAV data set which we will use for these problems, each variable is measured at three points in time: one week after the domestic violence incident, six months after the incident, and twelve months after the incident. All of the measures are Likert scales which we will treat as interval level variables.

I did not find any consistent recommendations for a minimum sample size, so I adapted a recommendation that Tabachnick and Fidell make for MANOVA (Multivariate Analysis of Variance), which is similar to Repeated Measures ANOVA in that both involve multiple dependent variables. The minimum sample size for these problems is $10 + \text{the number of dependent variables}$.

The Assumption of Normality

Like other ANOVA tests, the dependent variable is assumed to be normally distributed. We test for this by computing skewness and kurtosis, each of which should fall between $-1.0$ and $+1.0$ to satisfy the assumption. There is widespread consensus that violations do not seriously affect the result.

The Assumption of Sphericity

In repeated measures ANOVA, the same subjects are used at each time period, so there is an expectation that the measures will be correlated across subjects. If the correlations across time periods are similar. If they are not similar, e.g. suppose they are getting stronger over time, this will make the differences between time periods stronger than they really are.

If sphericity is violated, the computed significance for the F-tests are too low and we may be rejecting the null hypothesis when we shouldn’t. The solution usually employed is to use a correction factor to increase the degrees of freedom for the test, which in turn reduces the significance level to a more accurate number.

SPSS provides Mauchly’s Test of Sphericity and its significance. If we violate the assumption, we interpret the output for the Greenhouse-Geisser correct factor.

You can find a more thorough discussion of the issue at A Bluffer’s Guide to Sphericity.

Interpreting the Effects

The test for a main effect of change over time is the same that we have used for other types of ANOVA problems. Similarly, the Bonferroni pairwise comparison’s are interpreted for significance, but given the types of interpretation for change over time presented above.
To check for sample size requirements (and the assumption of homogeneity of group variances), we run the univariate general linear model procedure.

The introductory statement identifies the variables for the analysis and the significance levels to use. Note that we use a more conservative alpha (.01) for diagnostic statistics than we do for the statistics that answer our research questions.

Level of Measurement

Repeated-measures analysis of variance requires that the variable be interval level. The variables "Have positive attitude of myself (1 week)" [se1_6], "Have positive attitude of myself (6 months)" [se6_6] or "Have positive attitude of myself (12 months)" [se12_6] are ordinal level. However, we will follow the common convention of treating ordinal variables as interval level. We should consider including the use of this convention as a limitation of the analysis.
To check for compliance with sample size requirements, we run the univariate general linear model procedure. This procedure will give us the correct number of cases in each cell, taking into account missing data for the three time period variables in the analysis.

The minimum sample size requirement that we will use for repeated measures ANOVA is 10 + the number of dependent variables, adapted from Tabachnick and Fidell.

Select **General Linear Model > Repeated Measures** from the *Analyze* menu.
First, change the name of the *Within-Subject Factor Name* from the default of *factor1* to *time*. We can select any name that is meaningful to us.

Second, enter the number of levels in the text box. In this example, we have measures at 1 week, 6 months, and 12 months so we enter 3.

Third, click on the *Add* button to add this factor to the list.

First, highlight the *time(3)* factor in the list box.

Second, click on the *Define* button to tell SPSS what variables make up the time factor.
First, select the variable for the first time period, se_1_6.

Second, click on the right arrow button to move the variable to the first time slot.

The variables must be added in the correct sequence, since SPSS will assume that the order entered is correct.

First, select the variable for the second time period, se_6_6.

Second, click on the right arrow button to move the variable to the second time slot.
First, select the variable for the third time period, se12_6.

Second, click on the right arrow button to move the variable to the third time slot.

The variables for the three time periods have been added to the list box.

SPSS now has sufficient information to create the time factor.

Click on the Options button to request additional output.
Finally, click on the **Continue** button to close the dialog box.

**First**, copy the time factor from **Factors and Factor Interactions** to the **Display Means for list box**.

**Second**, mark the check box **Compare main effects**. This will compute the post hoc tests for the main effects.

**Third**, select **Bonferroni** from the **Confidence interval adjustment** drop down menu. This will hold the error rate for our multiple comparisons to the specified alpha error rate.

**Next**, mark the check boxes for:
- Descriptive statistics,
- Estimates of effect size,
- and
- Observed power.
To support the interpretation of the pattern of change, click on the **Plots** button to request a line chart.

**First**, move the time factor to the **Horizontal Axis** text box.

**Second**, click on the **Add** button to add this entry to the list of **Plots**.
Click on the **Continue** button to close the Profile Plots dialog.

Click on the **OK** button to obtain the output.
Evaluating the Sample Size

Marking the Statement for the Level of Measurement and Sample Size Requirement

The table of **Descriptive Statistics** contains the number of cases in each time period. The number of cases available for the analysis was 414. Since this number is greater than 10 + the number of levels in the repeated factor, the minimum sample size requirement is satisfied.

Since we satisfied both the level of measurement and the sample size requirements for analysis of covariance, we mark the first checkbox for the problem.
The next statement in the problem focuses on the assumption of normality, using the skewness and kurtosis criteria that both statistical values should be between -1.0 and +1.0.

Skewness and kurtosis are calculated in several procedures. We will use Descriptive Statistics.

Select Descriptive Statistics > Descriptives from the Analyze menu.
First, move the variables for the three time periods to the Variable(s) list box.

Second, click on the Options button to specify the statistics we want computed.

First, mark the check boxes for Kurtosis and Skewness.

Second, click on the Continue button to close the dialog box.
Computing Skewness and Kurtosis to Test for Normality – 4

Evaluating the Assumption of Normality - 1

Have positive attitude of myself (6 months) [se6_6] satisfied the criteria for a normal distribution. The skewness of the distribution (-.571) was between -1.0 and +1.0 and the kurtosis of the distribution (.669) was between -1.0 and +1.0.

Have positive attitude of myself (1 week) [se1_6] satisfied the criteria for a normal distribution. The skewness of the distribution (-.651) was between -1.0 and +1.0 and the kurtosis of the distribution (.081) was between -1.0 and +1.0.

Have positive attitude of myself (6 months) [se6_6] satisfied the criteria for a normal distribution. The skewness of the distribution (-.571) was between -1.0 and +1.0 and the kurtosis of the distribution (.669) was between -1.0 and +1.0.
Evaluating the Assumption of Normality

All three variables satisfied the skewness and kurtosis requirements for normality.

"Have positive attitude of myself (12 months)" [se12_6] satisfied the criteria for a normal distribution. The skewness of the distribution (-.501) was between -1.0 and +1.0 and the kurtosis of the distribution (.044) was between -1.0 and +1.0.

Since the skewness and kurtosis was between -1.0 and +1.0 for all three variables, the assumption of normality is satisfied and the check box is marked.

Marking the Statement for the Assumption of Normality
The Assumption of Sphericity

Evaluating the Assumption of Sphericity - 1

The next statement in the problem focuses on the assumption of sphericity, which is part of the output for Repeated Measures ANOVA.

The probability of Mauchly's Test of Sphericity (Mauchly's W(2) = .99, p = .201) was less than or equal to the alpha for diagnostic tests (.01).

The assumption of sphericity was satisfied.
Marking the Statement for the Assumption of Sphericity

Since we satisfied the assumption of sphericity, we mark the check box.

Interpretation of the main effect

The next statement indicates that there is a difference among the means for the three time periods, i.e. the average, or typical, score changed.
Implication of Satisfying the Assumption of Sphericity

Since we satisfied the assumption of sphericity, we will interpret the output on rows labeled Sphericity Assumed. Had we violated the assumption, we would interpret output on the rows labeled Greenhouse-Geisser, which corrects for the violations of the assumption.

The probability of the F statistic ($F(2, 826) = 5.28$, $p = .005$, partial eta squared = .01) was less than or equal to the alpha of (.05).

The null hypothesis that the average for all time periods is the same was rejected. The main effect of change over time was statistically significant.
Marking the Statement for the Main Effect

The main effect of change over time was statistically significant, so we mark the check box.

Interpretation of the Post Hoc Tests

The next four statements are possible interpretations of the pattern of change over time supported by the Bonferroni Post Hoc tests. Since each statement is a summary of the post hoc tests, only one will be correct (though it is possible than none are correct).

To identify the pattern of change, we first look at the significance of each pair wise comparison.
The difference between time period 1 and time period 2 is not significant. Any difference between the two group means would be characterized as a minimal difference.

For the statement "Have positive attitude of myself", the mean from 1 week ((M=3.16, SD=.75)) to 6 months ((M=3.24, SD=.64)) increased. The difference (-.08) had a probability (p = .121), greater than alpha (.05). The null hypothesis that these two means were equal was not rejected.

The difference between time period 1 and time period 3 is statistically significant. The difference between the two group means would be characterized as a substantial difference.

For the statement "Have positive attitude of myself", the mean from 1 week ((M=3.16, SD=.75)) to 12 months ((M=3.28, SD=.65)) increased. The difference (-.12) had a probability (p = .005), less than or equal to alpha (.05). The null hypothesis that these two means were equal was rejected.
The difference between time period 2 and time period 3 is not significant. Any difference between the two group means would be characterized as a minimal difference.

Based on estimated marginal means

For the statement "Have positive attitude of myself", the mean from 6 months ((M=3.24, SD=.64)) to 12 months ((M=3.28, SD=.65)) increased. The difference (-.04) had a probability (p = .803), greater than alpha (.05). The null hypothesis that these two means were equal was not rejected.

The only difference available to interpret is between time periods one and three.

Looking at the table of means, we see that the mean increased from time one (3.159) to time three (3.275).
Evaluating the Post Hoc Tests - 5

The profile plot supports an interpretation that the means for this question are increasing over time.

Marking the Statements for Post Hoc Tests - 1

Based on our evaluation of post hoc results, we are looking for an interpretative statement that states that there was a substantial increase from the first time period to the third, and only minimal changes from the first to the second time periods and from the second to the third time period.

The first three options all identify a substantial increase from either time 1 to time 2 or from time 2 to time 3, so they are not correct.
Marking the Statements for Post Hoc Tests - 2

The fourth option identifies a substantial change from the first time period (1 week) to the third time period (12 months), and only minimal changes from time 1 (1 week) to time 2 (6 months) and from time 2 (6 months) to time 3 (12 months). This is the correct option.

The Problem Graded in BlackBoard

When this assignment was submitted, BlackBoard indicated that all marked answers were correct, and we received full credit for the question.
Logic Diagram for Repeated Measures ANOVA Problems – 1

SPSS Univariate General Linear Model for Repeated Measures

Level of measurement and sample size ok? (> 10 + number of dv’s)

- Yes → Mark check box for correct answer
- No → Do not mark check box

Mark check box for correct answer

Ordinal dv's?

- Yes → Mention convention in discussion of findings
- No → Do not mark check box

SPSS Descriptives to get skewness and kurtosis for the repeated measures

Assumption of normality satisfied? (skewness, kurtosis between +/-1)

- Yes → Mark check box for correct answer
- No → Do not mark check box

Do not mark check box

Assumption of sphericity satisfied? (Mauchly's W Sig > diagnostic alpha)

- Yes → Interpret Sphericity Assumed Output
- No → Interpret Greenhouse-Geisser Output

Do not mark check box

State Greenhouse-Geisser correction used in findings
Logic Diagram for Repeated Measures ANOVA Problems – 2

Main effect statistically significant? (Sig < alpha)

Yes →

Compare pattern of change over time based on pair wise comparisons and means to interpretative statements

No → Do not interpret pair wise comparisons → Stop

Yes →

Statement matches pair wise comparison?

No → Do not mark check box for statement

Yes → Mark check box for statement