Chapter 5 – Analysis of Differences

**Exercise 1**

The data below shows students’ scores before and after an intervention in a school class.

1. Explain what type of analysis is appropriate for this data set to determine whether or not there is a significant difference between the students’ scores before and after the intervention class.

A related sample design such as paired samples t-test is necessary to test the hypothesis since the grades to be compared were taken from the same students.

1. State the null and alternative hypotheses.

H0: There is no significant difference between the mean grades of students before and after the intervention class.

H1: There is a significant difference between the mean grades of students before and after the intervention class.

1. Enter the data in Jamovi (Student Grades.csv file), editing the variable name, description, and data type properly.

|  |  |  |
| --- | --- | --- |
| Student | Before Intervention | After Intervention |
| 1 | 18 | 22 |
| 2 | 21 | 25 |
| 3 | 16 | 17 |
| 4 | 22 | 24 |
| 5 | 19 | 16 |
| 6 | 24 | 29 |
| 7 | 17 | 20 |
| 8 | 21 | 23 |
| 9 | 23 | 19 |
| 10 | 18 | 20 |
| 11 | 14 | 15 |
| 12 | 16 | 15 |
| 13 | 16 | 18 |
| 14 | 19 | 26 |
| 15 | 18 | 18 |
| 16 | 20 | 24 |
| 17 | 12 | 18 |
| 18 | 22 | 25 |
| 19 | 15 | 19 |
| 20 | 17 | 16 |







1. Perform the appropriate analysis and interpret the results according to the formed hypothesis. Assume uncertainty about the direction of outcome (two-tailed test).

Use the Paired Samples T-Test dialog box, as shown in the book. For this purpose use the 'Measure 1 does not equal Measure 2' hypothesis option, also 'Mean difference', Descriptives and Descriptives plots. The results should look as follows:





The plot shows that the average grade after the intervention class is higher than the grades before the intervention class. Looking at the *p* value, the paired samples T-Test results show that there is a significant difference between the students’ grades before and after the intervention class (*p* < .05) with a mean difference of 2.

**Exercise 2**

 A study assessed the effectiveness of a teacher feedback technique designed to reduce antisocial behavior in emotionally disturbed children playing korfball. A total of 8 children are enrolled in the study and the number of incidents of illicit or aggressive behavior during a week's sessions are measured both before the introduction of the feedback method and then again for a week afterwards. The data are shown below (you will also find this in the Feedback Technique.csv file).

|  |  |  |
| --- | --- | --- |
| **Child** | **Before Technique** | **One Week of Technique** |
| 1 | 85 | 75 |
| 2 | 70 | 50 |
| 3 | 40 | 50 |
| 4 | 65 | 40 |
| 5 | 80 | 20 |
| 6 | 75 | 65 |
| 7 | 55 | 40 |
| 8 | 20 | 25 |

1. Given this small dataset, explain what is the appropriate test to see if there is a significant difference between the median of two groups.

Wilcoxon signed rank test, which is a non-parametric counterpart of paired samples t-test, is appropriate to use in this study.

1. State the null and alternative hypotheses.

H0: The average antisocial behavior scores before and after treatment are the same.

H1: The average antisocial behavior scores before and after treatment are not the same.

1. Enter the data in Jamovi. Make sure to edit the variable name, description, and data type.







1. Perform the appropriate analysis and interpret the results according to the formed hypothesis.

Use the Paired Samples T-Test dialog, as shown in the book, but selecting the Wilcoxon rank option. Use the same descriptives options as in the previous example (first hypothesis, mean difference, descriptives, descriptives plots). The results should look as shown:





The plot shows that the average score before the intervention is higher than the score after treatment, with a mean difference of 12.8. If we applied a 0.05 critical value, the Wilcoxon signed rank test results would not allow us to reject the null hypothesis that the scores are similar before and after the treatment. However, given the small number of cases, some would consider applying a critical value of 0.10, so *p* < .10; however, most would suggest using larger samples!

**Exercise 3**

A teacher of English as a Foreign Language attempts to improve grammar levels in a group of 'false beginners'. Two sets of revision classes were performed. The scores were recorded as below.

|  |  |  |  |
| --- | --- | --- | --- |
|  Adult | Before Intervention | After First Revision Class | After Second Revision Class |
| 1 | 18 | 22 | 20 |
| 2 | 21 | 25 | 22 |
| 3 | 16 | 17 | 20 |
| 4 | 22 | 24 | 26 |
| 5 | 19 | 16 | 15 |
| 6 | 24 | 29 | 30 |
| 7 | 17 | 20 | 25 |
| 8 | 21 | 23 | 26 |
| 9 | 23 | 19 | 25 |
| 10 | 18 | 20 | 23 |
| 11 | 14 | 15 | 20 |
| 12 | 16 | 15 | 14 |
| 13 | 16 | 18 | 22 |
| 14 | 19 | 26 | 24 |
| 15 | 18 | 18 | 19 |
| 16 | 20 | 24 | 23 |
| 17 | 12 | 18 | 25 |
| 18 | 22 | 25 | 27 |
| 19 | 15 | 19 | 22 |
| 20 | 17 | 16 | 20 |

1. Explain why administering multiple t-tests would not be beneficial for this scenario. Which test is appropriate?

Every time you conduct a t-test, there is a chance of rejecting a true null hypothesis (a 'type 1 error'). Thus, conducting multiple paired sample t-tests in this case will increase the error rate.

To avoid this, the one-way repeated measures ANOVA is appropriate.

1. State the null and alternative hypotheses.

H0: There is no significant difference between the average score among the three data sets.

H1: There is a significant difference between the average scores among the three data sets.

1. Load EFL Grammar.csv into JASP (unlike Jamovi, JASP has no data editing tool).



1. Perform the appropriate analysis and interpret the results related to the hypothesis. Perform the procedure for setting up Repeated Measures ANOVA in JASP as shown in the book. The dialog box should end up looking something like this:





Assuming that the critical value has been predetermined as 0.05, *p* < .05; we may reject the null hypothesis that there is no significant difference on the average grades among the three groups.

To identify any significant pairings, use the Holm post hoc test. Also, show a descriptive plot to visualize the mean grade per group and easily identify their differences.





All pairings appear to be significantly different.

**Exercise 4**

Suppose that you want to compare the scores of participants from two *different* classes of 'false beginners', where the first class was subjected to a traditional teaching method, whereas the second was an experimental class, applying a different teaching style. Do note that each separate group needs to be represented by a number (here, within the 'Class' variable); this would be the same with an ANOVA (independent samples / unrelated), only with more than 2 such numbers.

This data set is available in Teaching Methods.csv

|  |  |  |
| --- | --- | --- |
| Participant | Score | Class |
| 1 | 18 | 1 |
| 2 | 21 | 1 |
| 3 | 16 | 1 |
| 4 | 22 | 1 |
| 5 | 19 | 1 |
| 6 | 24 | 1 |
| 7 | 17 | 1 |
| 8 | 21 | 1 |
| 9 | 23 | 1 |
| 10 | 18 | 1 |
| 11 | 14 | 1 |
| 12 | 16 | 1 |
| 13 | 16 | 1 |
| 14 | 19 | 1 |
| 15 | 18 | 1 |
| 16 | 20 | 1 |
| 17 | 12 | 1 |
| 18 | 22 | 1 |
| 19 | 15 | 1 |
| 20 | 17 | 1 |
| 21 | 22 | 2 |
| 22 | 25 | 2 |
| 23 | 17 | 2 |
| 24 | 24 | 2 |
| 25 | 16 | 2 |
| 26 | 29 | 2 |
| 27 | 20 | 2 |
| 28 | 23 | 2 |
| 29 | 19 | 2 |
| 30 | 20 | 2 |
| 31 | 15 | 2 |
| 32 | 15 | 2 |
| 33 | 18 | 2 |
| 34 | 26 | 2 |
| 35 | 18 | 2 |
| 36 | 24 | 2 |
| 37 | 18 | 2 |
| 38 | 25 | 2 |
| 39 | 19 | 2 |
| 40 | 16 | 2 |

1. State the null and alternative hypotheses.

H0: There is no significant difference between the average scores of participants in the two classes.

H1: There is a significant difference between the average scores of participants in the two classes.

1. Which test is appropriate here?

The Independent samples t test. There are only two groups. The scores being compared are from different participants (who did not take both classes). An examination of the dependent variable (Score) will show that the data can be considered normal; otherwise, a Mann-Whitney test would be preferred.

1. Input the data in Jamovi. Make sure to edit the variables data type accordingly to be able to perform the test.





1. Perform the analysis and interpret the results.

Use Jamovi's Independent Sample T-Test procedure, as shown in the book. Use all three assumption checks.





Before interpreting the results, check first that no assumption has been violated. If this was not the case, use the non-parametric counterpart which is the Mann-Whitney U test.

In this case, all assumptions were satisfied. The normality plot showed no deviation from the normal line. The normality test result agrees with this since the *p* value is greater than 0.05. For the test of equality of variance (homoscedasticity), the *p* value is greater than 0.05.

Note that the null hypotheses for the normality test and equality of variance test are as follows:

H0: The data follows the assumption of normality

H0: The data follows the assumption of equality of variance

Thus, having *p* values greater than 0.05, these null hypotheses are not rejected.



Since the *p* value is greater than 0.05, we do not reject the null hypothesis. There is no significant difference between the average grades of participants in the two classes.