HOMEWORK #7

**Note that for this homework, there are 36 items but there are a total of 40 possible points because some of the questions might ask for more than one piece of information. I have tried to denote those items.**

* + - 1. A researcher was interested in stress levels of lecturers during lectures. She took the same group of eight lecturers and measured their anxiety (out of 15) during a normal lecture and again in a lecture in which she had paid students to be disruptive and misbehave. The data were not normally distributed. Which test should she use to compare her experimental conditions?

1. Paired *t*-test
2. Mann–Whitney test
3. Wilcoxon signed-rank test
4. Wilcoxon rank-sum test
   * + 1. 2. The non-parametric equivalent of the paired-samples *t*-test is the:
   1. Mann–Whitney test
   2. Wilcoxon signed-rank test
   3. Friedman test
   4. Kruskal–Wallis test
      * 1. 3. The non-parametric equivalent of the independent *t*-test is the
           1. Mann–Whitney test
           2. Wilcoxon sign test
           3. Friedman test
           4. Kruskal–Wallis test

4. If the null hypothesis is true and we run the Mann–Whitney test on data, the expectation is that the \_\_\_\_\_ in the two groups will be approximately equal.

* 1. means
  2. ranks
  3. variances
  4. mean ranks
     + 1. As long as parametric assumptions are met, you prefer to analyze your data using parametric tests versus non-parametric tests because of which of the following reasons?

a. Non-parametric tests in general are harder to carry out.

b. Non-parametric tests in general are less powerful.

c. Parametric tests are less robust.

d. Parametric tests are more likely to show causal effects.

6. Ten visitors to an art gallery are asked to rate each of two sculptures for aesthetic value on a scale of 1 (low value) to 5 (high value) What procedure would be used to test the hypothesis that the ratings on each of the sculptures are not significantly different?

* 1. Mann–Whitney test
  2. Wilcoxon-signed-rank test
  3. Wilcoxon rank-sum test
  4. Kruskal–Wallis test

7. What symbol represents the test statistic for the Mann-Whitney test?

a. T

b. p

c. U

d. M

8. In a study investigating whether there is a difference in age at diagnosis of type 2 diabetes for young adult men and women, the result obtained for a two-tailed probability is *U* =3, *p* = .09). What does this result signify?

* 1. Indeterminate result
  2. No real evidence that age at diagnosis is different for young adult men and women
  3. Significant and therefore there is evidence of a difference in age at diagnosis
  4. None of the above

9. What would this same result for the study described in #8 be if a one-tailed probability were required?

a. .045

b. .45

c. .09

d. None of the above

10. After running a Mann-Whitney test to determine whether differences found are significant, you wish to calculate effect size. Which of the following will you need to compute effect size?

a. T and sample size

b. U and sample size

c. z-score of the test statistic and sample size

d. It is not possible to compute effect sizes for non-parametric tests

Let’s look at a couple of examples.

**Example 1.**

11. The data below reflects amount of temperature change in degrees in a muscle after 20 minutes of two different interventions: 1) hot packs applied to the muscle, 2) exercising the muscle by cycling. Please rank the data [as described in Field 7.4.1, page 216-217] in order to test significance of differences between the 2 independent groups using a non-parametric test. Worth 1 pts.

|  |  |  |  |
| --- | --- | --- | --- |
| **Hot Pack** | **rank** | **Cycling** | **rank** |
| **6** | **4** | **8** | **4** |
| **2** | **3** | **7** | **3** |
| **8** | **5** | **12** | **7** |
| **9** | **6** | **3** | **1.5** |
| **-3** | **1** | **11** | **5.5** |
| **0** | **2** | **3** | **1.5** |
| **12** | **7** | **11** | **5.5** |

**Example 2.**

Please use dataset titled “EMG Biofeedback” found in your folder for this example. This is a small dataset

showing values of change in EMG activity (in microvolts), as biofeedback for relaxation techniques was performed in 2 different positions (supine, or sitting). Participants were randomly assigned to the supine or sitting position. The researchers hypothesized that the positions would facilitate different levels of relaxation as measured by change in EMG activity.

12. Please state the null hypothesis for this study.

The two positions, supine and sitting, will not facilitate different levels of relaxation when measured by changes in EMG activity.

13. The alternate hypothesis (as hypothesized by researchers) is stated above. Is this a directional or non-directional hypothesis?

The two positions, supine and sitting, will facilitate different levels of relaxation when measured by changes in EMG activity.

14. Since the two randomized groups are independent (each group gets a different level of the treatment), and because the sample size is so small, the best way to analyze this data is with the Mann-Whitney test (even though the data is at the ratio level of measurement). Please remember that ordinarily, you always start by examining your data and testing for assumptions needed for parametric tests of significance. However, in this case, with an n=5 in one group and n=6 in the other group, there really is no point of testing for normality!

Looking at the data in the “EMG Biofeedback,” rank the data for the two groups. Remember that when ranking the data, BOTH groups are considered together, then ranks are summed for each group. Field p.216-217. (1 pt. each)

What is R (supine)? 42

What is R (sitting)? 24

15. The test statistic for the Mann-Whitney test is “U”. In fact, it is frequently referred to as the “Mann-Whitney U” test. Calculate U for each group (sitting and supine) using the formula: 1 pt. each

Usupine = *n*1 *n*2 + *n*1(*n*1+ 1) - R supine **See p. 218. Formula may be clearer in Field.**

2

What is your calculated U supine? 3

Now do the same for U sitting. What is your calculated U sitting? 27

Use: Usitting = *n*1 *n*2 + *n*2(*n*2+ 1) - R sitting

2

16. You could calculate the z-score which would be the difference between our ranks in each group divided by the Standard Error (SE), ( just like we have done in the past when calculating z-scores), but the formula is fairly cumbersome so we will let SPSS do that for us! Go ahead and run the Mann-Whitney test on the EMG data on SPSS at this time. Questions 17-23 relate to the resulting output. Please cut and paste the results of your Mann-Whitney test here. **Remember that you must double click on the Hypothesis Test Summary in order to get the specific statistics and graph.** Paste both to your document here. [NOTE: the Hypothesis Test Summary will be shown in the Output viewer and you can copy and paste from here into your Word document as you have done before. However, to get the additional test statistics that you see when you double click on the Test Summary, you must go to ‘Edit” and select “Copy Auxiliary View”, then paste it into the Output viewer. From here you can proceed as before.]

A picture containing table

Description automatically generated

Chart

Description automatically generated

17. What is the U-statistic reported by SPSS? 3 . This value should correspond to what you calculated for U supine. [Note: the U-test statistic is always the smallest of the calculated U1 or U2.] The value of U supine could then be compared to a critical value of U for a one-tailed or two-tailed test (depending on whether you have a directional or non-directional hypothesis) at a particular alpha level (usually .05). SPSS then calculates the z-score for U supine and to make a decision about whether the value is statistically significant or not. Remember that calculating a z-score standardizes the statistic so that you can compare to a normal, standard curve.

18. What is the z-score calculated by SPSS that is reported in your output? -2.191

19. Is the z-score statistically significant at an alpha level of .05? Explain.

Yes, because the z-score is less than -1.96 which is the standard negative z-score parameter at an alpha level of 0.05.

20. You have two p-values reported in your output. Although they are very close, which should you use to determine statistical significance of your calculated Mann-Whitney U test?\_\_\_\_\_\_\_\_. Why?

I would use the p-value of 0.030 because it is the exact significant value.

21. What is your decision, based on the calculated p-value regarding the null hypothesis?

I would reject the null hypothesis.

One final point regarding the calculated U. Why did SPSS use U supine  rather than U sitting as the test statistic? The reference used is always the smaller of the calculated U-values (in this case U supine). Because of the way U is calculated, this will be the U that comes from the largest sum of ranks. Remember that when we rank our data, larger values get a higher ranking so it is based on the group with the highest rankings.

22. Calculate the effect size observed in this experiment. (show your work). -2.191/11^(1/2) = -0.661

23. Please write a statement of your results showing how you would report them for a manuscript. When writing your result, don’t forget that the ranks that were used to calculate the Mann-Whitney U were based on how much the EMG values were reduced, because this would represent the amount of relaxation achieved. See p. 223 for reporting the results of Mann-Whitney U tests.

Biofeedback for relaxation techniques in the supine position (*Mdn* = 37) did significantly differ from the biofeedback for relaxation techniques in the sitting position (*Mdn* = 15) after administering an EMG, *U* = 3, *z* = -2.191, *p* = 0.030, *r* = -0.661.

**Example 3.**

Please use dataset titled “Extensor Strength Paired” found in your folder for this example. This is a small dataset showing the effects of two different knee flexion positions (90° and 15°) on knee extensor strength measurements. Muscle strength was assessed on a scale of 0-12 with 0 signifying no muscle activity and 12 signifying normal strength. (OK, I’ve never heard of this scale for assessing strength, but let’s go with it for the purposes of this example!) Ten individuals were tested for knee extensor strength in BOTH knee positions on the same day, though the order of testing position was randomized.

The researchers hypothesized that muscle strength values would be different depending on the position of the knee.

24. Please state the null hypothesis for this study. Muscle strength values will not differ depending on the position of the knee, either 90° or 15°.

25. Is the alternate hypothesis (as hypothesized by researchers) as stated above a directional or non-directional hypothesis? Non-directional hypothesis

26. All participants were tested in both positions and we have a small sample size (n=10) so we are thinking non-parametric analysis but let’s test assumptions for parametric analysis and look for outliers anyway. Remember that when assessing normality in related samples, we are interested in whether the **difference scores** are normally distributed, not whether the raw data is normally distributed. To assess assumptions of normality and to check for outliers, you will need to compute difference scores between the scores in 90° knee flexion and those in 15°knee flexion. Please paste your K-S and Shapiro Wilk tests of normality here.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tests of Normality** | | | | | | |
|  | Kolmogorov-Smirnova | | | Shapiro-Wilk | | |
| Statistic | df | Sig. | Statistic | df | Sig. |
| Extensor strength at 90 | .312 | 10 | .007 | .871 | 10 | .102 |
| Extensor strength at 15 | .335 | 10 | .002 | .662 | 10 | .000 |
| a. Lilliefors Significance Correction | | | | | | |

27. Well that’s interesting! Turns out our difference scores are normally distributed and we have no outliers but the best way to analyze this data is still with the Wilcoxon Signed-Ranks test because our strength measure is ordinal data.

Looking at the data in the “Extensor Strength Paired” dataset, you will note that each subject’s data is located in the same row and there is no grouping variable. Ranking the data for related samples works a little differently so let’s go ahead and rank our data.

On the dataset, please enter data for the difference score, including signs.

When ranking the difference scores, you ignore all zero differences (anything with a difference score of 0 is excluded). Of the remaining scores, rank all differences scores with the lowest score ranked as “1” and proceeding to the highest rank as before. Tied scores are treated the same as before. Once all difference scores are ranked, assign the appropriate + or – sign. [Hint: There will be seven ranks once you drop all zero differences and only ONE will be a negative rank.] Now sum all the positive ranks (T+ ) and sum the negative ranks (T-). You do not need to turn in the dataset with the additional ranked columns. It is just helpful to use for ranking. 1 pt. each

What is T+ : 27 What is T- : 1

28. The test statistic for the Wilcoxon Signed-ranks is T+. How easy is that?! Enter your test statistic here. 27

Once again, we could calculate a z-score and associated p-value to determine significance of differences but we will let SPSS do that!

29. Go ahead and run the Wilcoxon Signed-ranks test on the strength data on SPSS at this time. When you select the fields to analyze, you need to put in the two variables that will be paired/compared. It will work better if you select and move the 15 degree variable first, then the 90 degree variable. (You will get the same results except that your z-score will be positive.)

Please cut and paste the results of your Wilcoxon test here. As before, you must double click on the Hypothesis Test Summary in order to get the specific statistics and graph. Paste both to your document here.

Table

Description automatically generated with low confidence

Chart, histogram

Description automatically generated

Questions 30-36 relate to your SPSS output.

30. What is the test statistic (T) reported by SPSS? 27

31. What is the z-score calculated by SPSS that is reported in your output? 2.217

32. What does the positive value for the z-score tell you about the data? (look at your data to answer this) 1 pt. The positive value for the z-score tells us that when subtracting the values of the extensor strengths at 15° from 90°, most of the differences are positive (i.e., the frequency of positive differences is more than the frequencies of negative differences, and you can see this on the histogram).

NOTE: If you had entered your variables into SPSS in the opposite order (90 degree variable, then 15 degree variable), you would have gotten the same z-score but would have a negative sign. The indication below your generated bar graph will show “15-90”. This seems counter intuitive but…that’s the way it works.

33. Is the z-score statistically significant at an alpha level of .05? Explain.

Yes, because the z-score is more than 1.96 which is the standard positive z-score parameter at an alpha level of 0.05.

34. Note that you only have one p-value reported in your output this time, so you have no choice as to which to pick! What is the calculated p-value and what is your decision, based on the calculated p-value regarding the null hypothesis?

The calculated p-value is 0.027, and I would reject the null hypothesis.

Note that if you were to look up a z-score of 2.217 in Appedix A1, you would get a p of .01355 (Smaller portion) and since we have a non-directional hypothesis with 2 tails, we would multiply by 2 to get p=.0271. Exactly what SPSS calculated! Whew, good thing.

35. Calculate the effect size observed in this experiment. (show your work). 2.217/20^(1/2) = 0.496

36. Please write a statement of your results showing how you would report them for a manuscript. See page 230.

Muscle strength values at a 90° position (*Mdn* = 9.3) did differ significantly from muscle strength values at a 15° position (*Mdn* = 7.8) after testing, *Ws* = 27, *z* = 2.217, *p* = 0.027, *r* = 0.496.