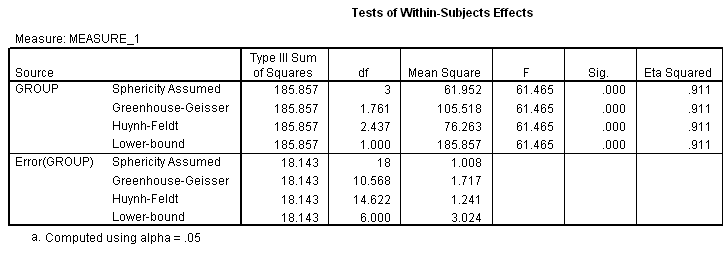
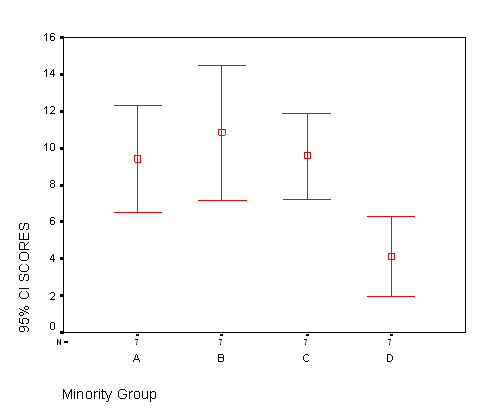
**Homework #10**

**There are a total of 40 points for this final homework**

1. When comparing the same group of individuals under different treatment conditions to determine whether there are significant differences between the treatments, a one-way ANOVA is NOT appropriate to use because:
   1. The assumption of independence is violated
   2. The assumption of sphericity is violated
   3. The error rate is inflated
   4. There is no problem with using one-way ANOVA since it is robust to violation of assumptions
2. The assumption of sphericity that must be met for repeated measures ANOVA is analogous to the homogeneity of variance assumption EXCEPT that HOV relates to equality of variances between groups while assumption of sphericity relates to:
   1. Equal variances within each group
   2. Equal variances between pairs of treatment levels
   3. Equal variances within treatment levels
   4. Equal variances between and within treatment levels
3. Which of the following tests is used to assess whether the assumption of sphericity has been met?
   1. Levene’s test
   2. Greenhouse –Geisser test
   3. Huyunh-Feldt test
   4. Mauchly’s test
4. When the assumption of sphericity is violated, you should:
   1. Adjust the model and error degrees of freedom accordingly
   2. Use a non-parametric test to analyze your data
   3. Ignore if your sample size is 30 or more
   4. Adjust your alpha level
5. Adjustments designed to correct for violations in sphericity, will not affect the F-ratio but can yield a different p-value that makes the test more conservative. (See Jane Superbrain 15.3) Which of the following corrections will yield the most conservative results in a repeated measures analysis?
   1. Assumed sphericity
   2. Greenhouse-Geisser correction
   3. Huynh-Feldt correction
6. Which of the following values for an estimation of sphericity would indicate your data is perfectly spherical?
   1. -1.0
   2. 0
   3. .75
   4. 1.0
7. Which of the following is an advantage gained when using a repeated measures ANOVA as opposed to an independent groups ANOVA?
   1. Systematic variation is reduced
   2. Between group variance is increased
   3. There is increased sensitivity for detecting experimental effects
   4. All of the above
8. In order for sphericity to be an issue in a repeated measures design, what is the minimum number of conditions required?
   1. 1
   2. 2
   3. 3
   4. None of the above
9. In a study examining attitudes to ethnic minorities, the following output was obtained via SPSS. In relation to this same experiment and the SPSS output detailed below, what does the information in the following error bar mean graph demonstrate?





1. Confidence intervals in A, B and C overlap
2. Confidence intervals in B, C and D overlap
3. Confidence intervals in D, A and B overlap
4. None of the confidence intervals overlap
5. In regards to the SPSS output in #9, you conclude that the overall ANOVA is significant and that you now must conduct follow-up pairwise comparisons to determine which group differences are significant (although you have a pretty good idea from just looking at the plots!). Although not shown in the output above, you know that sphericity was violated and you dealt with that issue already. Which of the following post-hoc tests should you use to control for Type I error but still preserve power adequately?
   1. Bonferroni
   2. Tukey
   3. Roy-Bose
   4. None of these can be used when assumption of sphericity is violated
6. You conduct an experiment where 20 participants are subjected to all 3 levels of a treatment (placebo, standard of care, experimental treatment) to see if there is any difference in function. What is the dfM ?
   1. 2
   2. 3
   3. 30
   4. 27
7. What is the dfR ? (This is also called the residual or error df)
   1. 30
   2. 38
   3. 40
   4. 42
8. When the assumption of sphericity is violated, which of the following are options to deal with the violation?
   1. Adjust the degrees of freedom using Greenhouse-Geisser or Huynh-Felt estimates of sphericity
   2. Use the non-parametric Friedman’s ANOVA
   3. Use a multivariate test statistic (MANOVA) that does not assume sphericity
   4. All of the above are options

Here are our final homework analyses for this semester. Well, here we go! [Yes, I am every bit as excited about this as you are!]

**Example 1. Use dataset titled “One-Way RM ANOVA.sav” from Module 11**

You wish to determine whether grip strength varies depending on the position of the forearm. To do this, you have recruited 18 participants and you will test each participant three times. Each time they will be seated and keeping the elbow in 90 degrees of elbow flexion, grip strength will be measured with a Jamar grip dynamometer with the forearm in 1) pronation, 2) neutral, and 3) supination. (The standardized test position for measuring grip strength is in the seated position, shoulder in zero degrees of flexion and abduction, elbow flexed 90 degrees and the forearm in neutral so the assumption is that this is the position that will produce the largest amount of force. Let’s see if that is the case.) Oh, by the way, this is bogus data so don’t assume anything!

Your null hypothesis is the following: **There will be no significant difference in grip strength values when measured with the forearm pronated, neutral or supinated.**

1. This time YOU decide what the alternate hypothesis is and state it here. There will be a significant difference in grip strength values when measuring the forearm either pronated, neutral, or supinated.

[Note: You cannot have a directional hypothesis with ANOVA. p. 413.]

Since each participant will be measured for maximal effort grip strength in all three positions, there are a few things to think about:

It is possible that strength may decrease by the third position measurement due to fatigue and it is also possible that the participants become more comfortable with practice such that they are “better” at the test due to learning effect. You should take this into account when designing your study. To remove potential variance due to fatigue or to learning, you decide to;

1. Randomize order of testing for the 3 positions
2. Allow adequate rest in between measurements

Ok now that that is taken care of let’s get down to the business of analyzing our data.

For this analysis, you will need to use the dataset titled **one-way RM ANOVA** in the Module 11 folder.

The one-way RM ANOVA has similarities to the one-way independent ANOVA that we just learned about (especially when it comes to post hoc tests and planned contrasts), but there is no “grouping” variable. We essentially have one group of people that participate at all three levels of the repeated factor “forearm position” which we will reduce to “position” in our analysis to satisfy SPSS.

**Step 1**. As always, you start out by checking assumptions. If our one-way RM ANOVA is significant, our post hoc tests of paired comparisons will essentially be related samples t-tests so we should test for normality of the distribution of difference scores rather than raw data (Field p. 344). To do that, you would need to use the “compute” function of SPSS to compute three new variables for the difference scores for the three paired comparisons you will run. As it turns out, the assumption of normality is violated for the pronated – neutral differences (Shapiro Wilk: p=.001) and we could analyze the data using the non-parametric Friedman’s test, but let’s assume that the Central Limit Theorem will correct (even though we only have a sample size of 18 and not the “magical” 30). There is no assumption for independence because we KNOW the measures are related. Homogeneity of variance also does not apply since we do not have 3 independent groups but from the reading, you know that instead we have the assumption of sphericity that must be met. Sphericity will be assessed automatically in SPSS when we run a RM ANOVA.

**Step 2.** Run the one-way RM ANOVA from the “General Linear Model” of the Analyze menu of SPSS. Running the RM ANOVA is a bit trickier than the between groups ANOVA. If you are comfortable running the RM ANOVA after doing the reading, go ahead and do so now. If you are uncertain, please view the Panopto video that I recorded titled “**One-way RM ANOVA**” on how to do this analysis, and what parts of the SPSS output are important to look at (and to paste for this question). [Note: The page number that I reference when talking about standard contrasts on the recording is on p. 405 in Field’s 5th Ed.] Also, in SPSS version 25, to select post hoc comparisons to run, you must access those by clicking on “EM Means” which stands for Estimated Marginal Means. These are not available via “Options” as indicated in Field’s text. Go ahead and select Bonferroni adjustment as the post hoc test for paired comparisons.

1. Please paste results of your one-way RM ANOVA here. (include the tables that are titled: Within-subjects Factors, Descriptive Statistics, Mauchley’s Test of Sphericity, Tests of Within-Subjects Effects, and Estimated Marginal Means (that include pairwise comparisons).

|  |  |
| --- | --- |
| **Within-Subjects Factors** | |
| Measure: Grip\_Strength | |
| Forearm\_Position | Dependent Variable |
| 1 | pron |
| 2 | neutral |
| 3 | supin |

|  |  |  |  |
| --- | --- | --- | --- |
| **Descriptive Statistics** | | | |
|  | Mean | Std. Deviation | N |
| grip strength in lbs with forearm pronated | 17.3333 | 9.84587 | 18 |
| grip strength in lbs with forearm in neutral | 27.5556 | 10.10662 | 18 |
| grip strength in labs with forearm supinated | 29.1111 | 11.00208 | 18 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Mauchly's Test of Sphericitya** | | | | | | | |
| Measure: Grip\_Strength | | | | | | | |
| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | Epsilonb | | |
| Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| Forearm\_Position | .664 | 6.540 | 2 | .038 | .749 | .805 | .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: Forearm\_Position | | | | | | | |
| 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. | | | | | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tests of Within-Subjects Effects** | | | | | | | |
| Measure: Grip\_Strength | | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
| Forearm\_Position | Sphericity Assumed | 1473.778 | 2 | 736.889 | 106.968 | .000 | .863 |
| Greenhouse-Geisser | 1473.778 | 1.498 | 984.131 | 106.968 | .000 | .863 |
| Huynh-Feldt | 1473.778 | 1.610 | 915.507 | 106.968 | .000 | .863 |
| Lower-bound | 1473.778 | 1.000 | 1473.778 | 106.968 | .000 | .863 |
| Error(Forearm\_Position) | Sphericity Assumed | 234.222 | 34 | 6.889 |  |  |  |
| Greenhouse-Geisser | 234.222 | 25.458 | 9.200 |  |  |  |
| Huynh-Feldt | 234.222 | 27.367 | 8.559 |  |  |  |
| Lower-bound | 234.222 | 17.000 | 13.778 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Pairwise Comparisons** | | | | | | |
| Measure: Grip\_Strength | | | | | | |
| (I) Forearm\_Position | (J) Forearm\_Position | Mean Difference (I-J) | Std. Error | Sig.b | 95% Confidence Interval for Differenceb | |
| Lower Bound | Upper Bound |
| 1 | 2 | -10.222\* | .862 | .000 | -12.510 | -7.935 |
| 3 | -11.778\* | 1.077 | .000 | -14.638 | -8.918 |
| 2 | 1 | 10.222\* | .862 | .000 | 7.935 | 12.510 |
| 3 | -1.556 | .627 | .072 | -3.221 | .110 |
| 3 | 1 | 11.778\* | 1.077 | .000 | 8.918 | 14.638 |
| 2 | 1.556 | .627 | .072 | -.110 | 3.221 |
| Based on estimated marginal means | | | | | | |
| \*. The mean difference is significant at the .05 level. | | | | | | |
| b. Adjustment for multiple comparisons: Bonferroni. | | | | | | |

1. Looking at your Descriptive Statistics table, what do you anticipate your SPSS analysis will show?

(If you had to guess. ☺). Based off the descriptive statistics table, I anticipate that my SPSS analysis will show that there will be a significant difference in grip strength value based on forearm position.

1. What is your Mauchly’s W statistic and associated p-value ? W = 0.664 and p = 0.038 What does this tell you about your data? (Be specific with regards to the value of Mauchley’s W as well as the p-value.) Based off the value of W, the variances between the groups are roughly equal and sphericity has been met. The p-value found with Mauchly’s W was found to be significant, since p was found to be < 0.05 at 0.038.
2. Which of the adjustments for violated sphericity provided in your in your SPSS output is considered the most conservative (makes is most difficult to find significance) and why?

The Huynh-Feldt adjustment for violated sphericity is the most conservative adjustment because it is greater than 0.75 at 0.805, which is considered a too conservative estimate.

1. Which of the adjustments for violated sphericity provided in your in your SPSS output is considered the most liberal (makes is least difficult to find significance) and why?

The Greenhouse-Geisser adjustment for violated sphericity is the most liberal adjustment because it is about 0.75 at 0.749, which is considered a fair estimate.

1. In your Tests of Within-Subjects Effects, you get an F-ratio of 106.968. How did SPSS generate this value? That is, what does it use as numerator and denominator to derive the ratio? The F-statistic is the ratio of variance explained by the model and the variance explained by unsystematic factors, i.e., the model mean squares divided by the residual mean squares.

[Hint: this should make sense to you because it represents our original basic concept that test statistics such as t and F and others are nothing more than a ratio of variance explained by the model to variance not explained by the model (ie. error). Funny how we started with this concept and we are now ending with it!]

1. What is the p-value provided by SPSS associated with the test of overall differences (sometimes called the omnibus test or test of main effects)? The p-value was found to be 0.038.
2. Did you use an adjusted value to interpret your within-subjects effects in #21? Why or why not?

No, I did not use an adjusted value because there was no departure from sphericity.

1. What do your pairwise comparisons show? (Which groups if any, were significantly different)?

Are these results consistent with what you anticipated in #16? The pronated forearm position, when compared against the neutral or supine forearm position, showed a significant difference, since the p-value for these comparisons were all < 0.001.

1. How many **different** paired comparisons were conducted in your analysis? 3
2. Are you concerned about family-wise error rate? Why or why not? No, because we conducted a Bonferroni test on the data set when conducting the ANOVA.
3. Please write your results like you would in a manuscript. Unfortunately, the examples provided in Field on p. 499 only demonstrate reporting of the overall RM ANOVA. Please write the results of your overall ANOVA (which addresses your null hypothesis) and also the pairwise comparisons (which address your alternate hypothesis). These are essentially paired t-tests which you learned how to report earlier. Don’t worry about effect size (although typically, you would report this).

When comparing the effect of different forearm positions on grip strength, there was a statistically significant difference between at least two of the three groups (supine, pronated, and neutral) F(2, 17) = 106.968, p < 0.001.

1. Speaking of effect size, take a look at the calculation of effect sizes for repeated measures

ANOVA on p. 498. Yikes! This does not look at all pleasant. Although you COULD do it, calculating omega squared is not particularly useful because it is the measure of *overall* effect size (for your overall ANOVA). Although a significant overall ANOVA is important because you cannot look at pairwise comparisons unless the overall test of differences is significant, you are usually MOST interested in knowing the effect size for differences between the treatment conditions.

So… what is the effect size for the pairwise comparisons that you ran? You know how to calculate these because we calculated effect size for paired samples t-tests a few weeks ago and that is exactly what the pairwise comparisons are…paired t-tests. Although you can calculate *r* as an effect size, Cohen’s d provides a better effect size when using t-tests to compare means. You could calculate a Cohen’s d using the formula on page 353, but since we don’t have a control group in this example, you would need to calculate a pooled variance for the denominator, which isn’t much fun at all! However, there are many effect size calculators available on line for both independent and related samples t-tests. I have provided you with one in your Homework Module. For Cohen’s d calculation for related samples, you will need to input the means, SD and correlation coefficient for each paired comparison. You have the means and SD in your Descriptive Statistics table pasted above but you will need to run correlations for the 3 conditions (pron, neutral, supin). Make sure you are using the calculator for related samples and not independent samples.

Please report your effect size (*d)* for each paired comparison and indicate if the effect size is small, medium, or large. Effect size for 1&2: *d* = 0.216 which is small

Effect size for 1&3: *d* = 0.224 which is small

Effect size for 2&3: *d* = 0.071 which is not even considered small

1. Did your effect size in #27 reflect what you found in terms of significance of differences for each paired comparison (#23)? Why or why not? No it wasn’t because the effect sizes were found to not be so different throughout the three comparisons, which is in direct contrast to the paired comparisons in #23.

**Example 2.**

So you’ve been losing sleep over that pesky violation of normality and you have decided that you will analyze your data using non-parametric statistics.

You decide that the Friedman’s ANOVA is the appropriate test and you are using the same dataset with same hypotheses as for Example 1.

1. Similar to other non-parametric tests, the Friedman’s ANOVA is based on ranked data versus the actual scores. In our example, we have 18 participants who have measurements of grip strength for 3 different conditions. Which of the following best describes how you would rank the data?
   1. Rank the scores for each condition then sum the ranks
   2. Rank the data across the three conditions for each person then sum the ranks for each condition
   3. Rank the difference between two conditions at a time then sum the ranks for each paired condition.
   4. None of the above
2. What is the test statistic that is calculated for a Friedman’s ANOVA?
   1. T
   2. Tr
   3. F
   4. Fr
3. The test statistic for a Friedman’s ANOVA approximates which of the following distributions?
   1. z-distribution
   2. t-distribution
   3. F-distribution
   4. Chi square distribution
4. If you were to look up your Friedman’s statistic in the appropriate distribution table, you would need to know the degrees of freedom. How do you determine degrees of freedom and how many degrees of freedom are there in this example? You determine degrees of freedom by subtracting 1 from the number of groups (k-1), and in this example, there are 2 degrees of freedom.
5. Please run your Friedman’s ANOVA on SPSS and paste results here. This is a very similar process as running the Kruskal-Wallis test you ran earlier. As before, you should include the Hypothesis Test Summary as well as the more detailed Friedman’s ANOVA result

Text

Description automatically generated with low confidence

A picture containing chart

Description automatically generated

1. What is the “Test Statistic” shown in your output (not the value, but what this value represents)?

The test statistic shown in the output is meant to represent Fr.

1. Since the p-value shown in your output is significant, you need to run additional tests to see which differences in the conditions are statistically significant. Please run the multiple comparisons analysis and request the “All pairwise” method. Paste your results here.

A picture containing chart

Description automatically generated

1. What is the test statistic reported in the generated output? (again this not requesting the values generated but rather what do these represent?) The test statistic is supposed to represent a t-statistic.
2. Which of the paired comparisons are significant? The paired comparisons of pronated grip strength while the forearm is in neutral and pronated grip strength while the forearm is supinated are both significant comparisons.
3. Calculate an effect size for each paired comparison and indicate if small, medium, or large as before. See page. 245 for a reminder. Don’t forget that N in this calculation is the total number of observations for each paired comparison.

Effect size of 1&2: *r* = -0.629 which is medium size

Effect size of 1&3: *r* = -0.915 which is large size

Effect size of 2&3: *r* = 0.286 which is small size

1. Please write a statement of your results showing how you would report them for a manuscript. Be sure to include the results of the overall ANOVA as well as the paired comparisons. There is a great example on page 245.

The grip strength of patients did significantly change based off the position the forearm was in, *X*2= 0.827, *p* < 0.001.

1. Do your results agree with your earlier results of the one-way RM ANOVA with regards to your null hypothesis and what you hypothesized would happen?

Yes, they both agree that grip strength did significantly changed based off the position the forearm was in.