|  |  |
| --- | --- |
| **Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **Student ID: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |

**Statistics for Psychology - PSYCH-UH 1004Q**

**Exam #2 – Advanced Inferential Statistics**

Due 12/20/23 at 11:59pm

80 points, 33% of semester grade

**Attention!** The exam consists of **four parts**.

Please make sure you have completed all four before submitting.

**Part 1: Conceptual – Correlation and Regression (17 points)**

**Instructions.** For this portion of the exam, you will answer 7 questions with short answers on this word document. These do not require long answers. Many can be answered with a single sentence or a single value. There is no need to write full paragraphs or long essays. But please be sure that your sentence or sentences completely answer the question.

You may type directly into this word document and submit it through Brightspace.

**\*\*\***

1. What is the range of values of Pearson's *r*? (2 points)
2. Which value of Pearson's *r* indicates absolutely no correlation between the two variables? (1 point)
3. Write the equation for a linear model in statistics (using the symbols *b0*, *b1*, *X* and ŷ). (5 points)
4. For each of the four terms in the equation for a linear model, tell us the information that these terms give us about our scientific theory (the properties of the world captured by our independent and dependent variables). In other words, what we learn from the model. (4 points)

*ŷ:*

*b0*:

*b1*:

*x*:

1. What does the *R2* statistic measure? (1 point)
2. List one advantage of repeated measures designs and one disadvantage of repeated measures designs. (2 points)
3. List one advantage of independent measures designs and one disadvantage of independent measures designs. (2 points)

**Part 2: Conceptual - ANOVA (24 points)**

**Instructions.** For the conceptual portion of the exam, you will answer 18 questions with short answers on this word document. These do not require long answers. Many can be answered with a single sentence or a single value. There is no need to write full paragraphs or long essays. But please be sure that your sentence or sentences completely answer the question.

1. In your own words, describe the two methods for calculating variance that are used in ANOVA. We are not looking for calculations or formulas. Just a brief description of the two methods that we have for calculating variance. (4 points)
2. In an ANOVA, does MSB vary based on whether H0 is true or false (yes or no)? (1 point)
3. In an ANOVA, does MSW vary based on whether H0 is true or false (yes or no)? (1 point)
4. Write an equation for *F* using the terms MSB and MSW. (1 point)
5. When H0 is false, what happens to *F*? (1 point)
6. In a one-way ANOVA, what are the degrees of freedom for MSB? (1 point)
7. In a one-way ANOVA, what are the degrees of freedom for MSW? (1 point)
8. In your own words, describe the multiple comparisons problem. (We don't need any formula or correction, just a description of the problem itself.) (1 point)
9. In your own words, what is a planned comparison? (1 point)
10. In your own words, what is a post-hoc comparison? (1 point)
11. If you have 3 total pairwise comparisons, and run *t*-tests for 2 of them as planned comparisons, what is the value of C that you should use when calculating a Dunn correction? (1 point)
12. If you have 3 total pairwise comparisons, and run *t*-tests for 2 of them as post-hoc comparisons, what is the value of C that you should use when calculating a Dunn correction? (1 point)
13. If someone says they ran an experiment with a 2x2 design, how many independent variables did they have in their experiment? (1 point)
14. If someone says they ran an experiment with a 2x2 design, how many levels did each factor of their experiment have? (1 point)

To answer questions 15-17, select one of the following three plots (Plots A-C):



1. Which plot shows two main effects and no interaction? (1 point)
2. Which plot shows an ordinal (or monotonic) interaction? (1 point)
3. Which plot shows a disordinal (or non-monotonic) interaction? (1 point)

For questions 18-21, you will provide the formulae for the degrees of freedom for a two-way, independent samples ANOVA. To make this easier, use Ntotal for the total number of subjects, r for the number of groups in factor 1, and c for the number of groups in factor 2.

1. What is dfW? (1 point)
2. What is df for factor 1? (1 point)
3. What is df for factor 2? (1 point)
4. What is dfinteraction? (1 point)

**Part 3: Practical – Linear Regression (14 points)**

**Instructions.** For this section, you will practice analyzing a regression study. For the study, read the description, and then perform the analysis steps or answer the questions in the numbered items below the description.

Study 1

A researcher is interested in understanding the effect of age on memory tasks. They decide to use a free recall task in which participants are shown a list of 10 words, and must recall as many as possible. The researcher has a hypothesis that older participants will perform worse on the task than younger participants. The researcher recruits 20 participants. The researcher records their age in years and asks them to complete a free recall task during a visit to the lab.

The CSV file study.1.csv contains the results of the experiment. Please use this file and the description above to perform the analyses requested.

1. Create a scatter plot of the age and recall scores. Make sure that the axis labels and axis scales are correct. Make sure to put each variable on the most appropriate axis based on the theory. You may use R or draw it by hand. Place the plot image here in the word document. You do not need to submit the R code. (3 points)
2. Calculate the correlation coefficient between age and recall. You can either use R or do it by hand. Report it here: (1 point)
3. Perform a *t*-test on the correlation coefficient. You can either use R or do it by hand. Report the following: (3 points)

The value of the test statistic:

The degrees of freedom:

The *p*-value associated with the test statistic:

1. Would you reject the null hypothesis that the correlation is 0 or not? (1 point)
2. Calculate a linear model for age and recall scores. It will probably be easiest to do this using R, but you can also use the equations in the textbook to do this. (2 points)

What is the value of *b1*:

What is the value of *b0*:

1. Use the linear model you just created. What is the predicted recall score for someone who is 80 years old? (3 points)
2. Do you think this predicted recall score makes sense? (1 point)

**Part 4: Practical – ANOVA (25 points)**

**Instructions.** For this section, you will practice analyzing two studies. For each study, read the description, and then answer the questions in the numbered items after the description.

Study 2

A researcher is interested in understanding the effect of social media on overall happiness. To test this, they randomly assign 60 participants into 3 groups of 20 each. The first group is instructed to avoid all social media for a period of two week (*none*). The second group is instructed to limit their social media consumption to 2 hours per day maximum for the same period of two weeks (*moderate*). The third group is given no restrictions on social media consumption (*high*). At the end of the two-week period, the researcher measures their overall happiness on a standardized test that yields a score between 0 (extremely unhappy) and 10 (extremely happy).

The data for this study is available in study.2.csv.

Here is a plot of the results:



To analyze these results, please run a one-way, independent samples ANOVA in R, and answer the following questions:

1. What is MSB? (1 point)
2. What is MSW? (1 point)
3. What is *F*? (1 point)
4. What is dfB? (1 point)
5. What is dfW? (1 point)
6. What is the *p*-value for the observed *F*? (1 point)
7. What is your statistical decision? (1 point)
8. There are 3 possible pairwise comparisons in this experiment. If we run these pairwise comparisons, and if we set the alpha criterion to .05 when deciding if these comparisons are significant, what would the experimentwise error rate be? (1 point)
9. Run each of the pairwise comparisons as (uncorrected) *t*-tests using R. Report the (uncorrected) *p*-value for each comparison here: (3 points)

none vs moderate:

none vs. high:

moderate vs high:

1. Now use the Dunn correction to hold the experimentwise error rate to .05. What is the alpha criterion that you must use? (1 point)
2. Which of the pairwise comparisons are statistically significant after Dunn correction? (1 point)

Study 3

A researcher is interested in understanding the effect of social media on overall happiness. To test this, they constructed a 2x2 design with two independent variables: how often the participants will view social media (condition, with two levels: low, high) and which social media platform they will view (company, with two levels: twitter, facebook). The researcher randomly assigned 12 participants to each group. At the end of a two-week period, the researcher measures their overall happiness on a standardized test that yields a score between 0 (extremely unhappy) and 10 (extremely happy).

The data for this study is available in study.3.csv.

Here is a plot of the results:



To analyze these results, please run a two-way, independent samples ANOVA in R, and answer the following questions:

1. From the plot, would you call this an ordinal (monotonic) or disordinal (non-monotonic) interaction? (1 point)
2. What is MSW? (1 point)
3. What is MSB for the factor we have called condition? (1 point)
4. What is MSB for the factor we have called company? (1 point)
5. What is MSinteraction? (1 point)
6. What is the *F* for the factor we have called condition? (1 point)
7. What is the *F* for the factor we have called company? (1 point)
8. What is the *F* for the interaction? (1 point)
9. What is the *p*-value (rounded to 4 digits after the decimal) for the factor we have called condition? (1 point)
10. What is the *p*-value (rounded to 4 digits after the decimal) for the factor we have called company? (1 point)
11. What is the *p*-value (rounded to 4 digits after the decimal) for the interaction? (1 point)
12. What is your statistical decision for the interaction? (1 point)