EDUC/PSY 6600 | Name: | **Unit 3 Assignment**

Research Design & Analysis I | **Hypothesis Tests for 2 Measures per Subject**

---

*Please complete the following exercises. Feel free to work with classmates, but each student must turn in **UNIQUE** work, not photocopies or identical replicates. When applicable, use **APA format** in communicating your results in text. **Show your work!** If any question involves any math at all, show your work. When it doubt, write it out. Always show more than you think you need.*

## 1) WRITE-UP - Textbook Problems

| Cohen Chap | | Exercises | Pts | Off |
|---|---|---|---|---|
| 9 | A | 1, 2, *7 | 3 | |
| | B | *5, *6 | 2 | |
| | C | 1, 2, 3, 4 | 4 | |
| 10 | A | *7, *8, 9 | 3 | |
| | B | 6, *9, *10, *15 | 4 | |
| | C | 1, 2 | 2 | |
| 11 | A | *2, *3, 7, *8 | 4 | |
| | B | 3, *8, 9, *11, *13 | 5 | |
| | C | 1, 2, 3 | 1 | |

## 2) SUMMARY – Supplementary Reading

| Increased arterial stiffness parameters in panic disorder patients | | Pts | Off |
|---|---|---|---|
| Half Page | Read the Unit 3 Journal Article on Canvas. Summarize any mention or use/abuse of the concepts in the above chapters. | 5 | |

## 3) R SYNTAX – Section B: Various data set – add to the skeleton R notebook and knit to .pdf & upload

| Cohen Chap | | Exercises | Pts | Off |
|---|---|---|---|---|
| 9 | B | *5, *6 | 2 | |
| 10 | B | 6, *9, *10 | 2 | |
| 11 | B | 3, *8, 9 | 3 | |

## 4) R SYNTAX – Section C: Ihno's data set – add to the skeleton R notebook and knit to .pdf & upload

| Cohen Chap | | Exercises | Pts | Off |
|---|---|---|---|---|
| 9 | C | 1, 2, 3, 4 | 5 | |
| 10 | C | 1, 2 | 2 | |
| 11 | C | 1, 2, 3 | 3 | |

| Grading | | Earned | Possible |
|---|---|---|---|
| ***CORRECTNESS*** | *a subset of spot-checked items: must show work, especially items from back of book or done in class* | | 50 |
| ***COMPLETENESS*** | *more than one item is missing or skipped: 25/50*<br>*roughly half the assignment is completed: 10/50* | | 50 |
| | | | 100 |

| 9 A | 1. Correlation: positive vs. negative |
|---|---|

Describe a realistic situation in which two variables would have a **high positive correlation**.

Describe another situation for which the correlation would be **highly negative**.

| 9 A | 2. Association does NOT imply causation, in observational studies |
|---|---|

A recent medical study found that the moderate consumption of alcoholic beverages is **associated with** the fewest heart attacks (as compared to heavy drinking or no drinking).

It was suggested that the alcohol caused the **beneficial effects**.

Devise an **explanation** for this relationship that assumes there is **no direct causal link** between drinking alcohol and having a heart attack. *( Hint : Consider personality.)*

## 9 A *7. Low Pearson's r

A psychologist is studying the relationship between the reported vividness of visual imagery and the ability to rotate objects mentally. A sample of graduate students at a leading school for architecture is tested on both variables, but the Pearson's r turns out to be **disappointingly low**.

Which of the following is the most likely explanation for why Pearson's r was **not higher**?

- ☐ a.) One or both of the variables has a restricted range.
- ☐ b.) The relationship between the two variables is curvilinear.
- ☐ c.) The number of degrees of freedom is too small.
- ☐ d.) One variable was just a linear transformation of the other.

## 9 B *5. Test for Association: Pearson's r — Code: R notebook

A psychiatrist has noticed that the schizophrenics who have been in the hospital the longest score the lowest on a mental orientation test. The data for 10 schizophrenics are listed in the following table:

| Years of Hospital (X) | Orientation Test (Y) |
|---|---|
| 5 | 22 |
| 7 | 26 |
| 12 | 16 |
| 5 | 20 |
| 11 | 18 |
| 3 | 30 |
| 7 | 14 |
| 2 | 24 |
| 9 | 15 |
| 6 | 19 |

a) Calculate **Pearson's r** for the data.

**r = _____**

b) Test for statistical significance at the **.05** level (**two**-tailed). (SPSS)

- ☐ Evidence of linear association
- ☐ No such evidence

**2-tail: p = _____**

## 9 B *6. Reliability: Pearson's r for test-retest scores — Code: R notebook

If a test is reliable, each participant will tend to get the same score each time he or she takes the test. Therefore, the correlation between two administrations of the test (test-retest reliability) **should be high**. The **reliability** of the verbal GRE score was tested using five participants, as shown in the following table:

| Verbal GRE (1) | Verbal GRE (2) |
|---|---|
| 540 | 570 |
| 510 | 520 |
| 580 | 600 |
| 550 | 530 |
| 520 | 520 |

a) Calculate **Pearson's r** for the test-retest reliability of the verbal GRE score.

**r = _____**

b) Test for statistical significance at the **.05** level (**one**-tailed).

- ☐ Evidence of linear association
- ☐ No such evidence

**1-tail: p = _____**

Would this correlation be significant with a **two**-tailed test?

- ☐ Evidence of linear association
- ☐ No such evidence

**2-tail: p = _____**

| 9 C | 1. Scatterplot: estimate Pearson's r | Code: R notebook |
|---|---|---|

a) Create a scatter plot of **phobia** (X) versus **statquiz** (Y).

From looking at the plot, do you think the **Pearson's r** will be:

☐ positive -or- ☐ negative | ☐ Large -or- ☐ medium -or- ☐ small?

b) Create a scatter plot of **baseline anxiety** (X) versus **postquiz anxiety** (Y).

From looking at the plot, do you think the **Pearson's r** will be:

☐ positive -or- ☐ negative | ☐ Large -or- ☐ medium -or- ☐ small?

| 9 C | 2. Calculate Pearson's r | Code: R notebook |
|---|---|---|

a) Compute the Pearson's r between **phobia** (X) versus **statquiz** (Y), for ALL students.

**r = _____**

Also, compute the Pearson's r between **baseline anxiety** (X) versus **postquiz anxiety** (Y).

**r = _____**

b) Use **dplyr::filter()** to **delete** any student whose baseline anxiety is **over 29**, and repeat part b of the first exercise:

Create a scatter plot of **baseline anxiety** (X) versus **postquiz anxiety** (Y).

From looking at the plot, do you think the **Pearson's r** will be:

☐ positive -or- ☐ negative | ☐ Large -or- ☐ medium -or- ☐ small?

Also, re-run the Pearson's r between **baseline anxiety** (X) versus **postquiz anxiety** (Y).

**r = _____**

What **happened** to the Pearson's r?

Use the **change in the scatter plot** to explain the change in the correlation coefficient.

| 9 C | 3. Calculate Pearson's r, report APA style | Code: R notebook |
|---|---|---|

a) Compute Pearson's r s among the **three measures of anxiety**. Write up the results in **APA style.**

b) Compute the **average of the three measures of anxiety**, and then compute the correlation between each measure of anxiety and the average, so that the output contains a **single column** of correlations.

| Anxiety Measure | Average Anxiety |
|---|---|
| Baseline | **r =** _____ |
| Pre-quiz | **r =** _____ |
| Post-quiz | **r =** _____ |

| 9 C | 4. Pearson's r: Missing values | Code: R notebook |
|---|---|---|

a) Compute Pearson's r for the following list of variables:
**Mathquiz, Statquiz**
**Phobia**
(na.rm = FALSE)

| | Mathquiz | Statquiz | phobia |
|---|---|---|---|
| Mathquiz | | | |
| Statquiz | **r =** _____ | | |
| phobia | **r =** _____ | **r =** _____ | |

b) Repeat part a after selecting **Exclude cases**
**listwise**
(na.rm = TRUE)

| | Mathquiz | Statquiz | phobia |
|---|---|---|---|
| Mathquiz | | | |
| Statquiz | **r =** _____ | | |
| phobia | **r =** _____ | **r =** _____ | |

Which correlation was **changed**? Explain **why**.

## 10 A *7. Regression Equation: calculate from summary stats

For a hypothetical population of men, **<u>waist size</u>** is **positively correlated** with **<u>height</u>**, such that:

- Pearson's r = + .6
- The mean **height** ($\mu_X$) for this group is 69 inches with $\sigma_X = 3$
- The mean **waist** measurement ($\mu_Y$) is 32 inches with $\sigma_Y = 4$.

a) What is the **slope** of the regression line predicting waist size from height? (formula 10.3A)

**slope = _____**

b) What is the value of the **Y intercept**? (formula 10.3B)

**y-intercept = _____**

c) Does the value found in part **<u>b</u>** above **make any sense**?

d) Write the raw-score **regression equation** predicting waist size from height.

## 10 A *8. Regression Equation: make predictions

Based on the regression equation found in Exercise 7:

a) What waist size would you **predict** for a man who is **6 feet tall**?

**waist = _____ inches**

b) What **waist** size would you **predict** for a man who is **62 inches tall**?

**waist = _____ inches**

c) **How tall** would a man have to be for his predicted **waist** size to be **34 inches**?

**height = _____ inches**

## 10 A 9. Regression Equation: variance measures

a) In Exercise 7, what is the value of the **coefficient of determination**?

**$r^2$ = _____**

b) How large is the **coefficient of nondetermination**? (formula 10.8A)

**$k^2$ = _____**

c) How large is the **variance of the estimate** ("**residual variance**")? (formula 10.8B)

**$\sigma^2_{est\ Y}$ = _____**

| 10 B | 6. Regression: swap X and Y | Code: R notebook |
|---|---|---|

A cognitive psychologist is interested in the relationship between spatial ability (e.g., ability to rotate objects mentally) and mathematical ability, so she measures 12 participants on both variables. The data appear in the following table:

| ID | Spatial Ability Score | Math Score |
|---|---|---|
| 1 | 13 | 19 |
| 2 | 32 | 25 |
| 3 | 41 | 31 |
| 4 | 26 | 18 |
| 5 | 28 | 37 |
| 6 | 12 | 16 |
| 7 | 19 | 14 |
| 8 | 33 | 28 |
| 9 | 24 | 20 |
| 10 | 46 | 39 |
| 11 | 22 | 21 |
| 12 | 17 | 15 |

a) Find the regression equation for predicting the **math** score **from** the **spatial** ability score.

b) Find the regression equation for predicting the **spatial ability** score **from** the **math** score.

c) According to your answer to **part a**, what **math score** is predicted from a spatial ability score of 20? *(by hand)*

**math score = _____**

d) According to your answer to **part b**, what **spatial ability score** is predicted from a math score of 20? *(by hand)*

**Spatial ability score = _____**

| 10 | B | *9. Regression: Predictions & residuals | Code: R notebook |
|---|---|---|---|

A cognitive psychologist is interested in the relationship between spatial ability (e.g., ability to rotate objects mentally) and mathematical ability, so she measures 12 participants on both variables. The data appear in the following table:

a) Find the regression equation for **predicting shoe size** from age.

b) Find the regression equation for **predicting reading level** from age.

c) Use the equations from parts a and b to make shoe size and reading level predictions for each child. Subtract each prediction from its actual value to find the residual.

| | | Shoe Size | | | Reading Level | | |
|---|---|---|---|---|---|---|---|
| Child | Age | Actual | Predicted | Residual | Actual | Predicted | Residual |
| 1 | 8 | 5.2 | | | 1.7 | | |
| 2 | 6 | 4.7 | | | 1.5 | | |
| 3 | 7 | 7.0 | | | 2.7 | | |
| 4 | 8 | 5.8 | | | 3.1 | | |
| 5 | 9 | 7.2 | | | 3.9 | | |
| 6 | 10 | 6.9 | | | 4.5 | | |
| 7 | 11 | 7.7 | | | 5.1 | | |
| 8 | 12 | 8.0 | | | 7.4 | | |

| 10 | B | *10. Regression: Predictions | Code: R notebook |
|---|---|---|---|

a) Calculate **Pearson's r** for **shoe size** and reading level using the data from Exercise 9.

**r = _____**

b) Calculate **Pearson's r** for the two sets of **residuals** you found in part c of Exercise 9.

**r = _____**

c) **Compare** your answer in **part b** with your answer to **part a**. The correlation in part b is the partial correlation between shoe size and reading level after the confounding effect of age has been removed from each variable (see Chapter 17 for a much easier way to obtain partial correlations).

| 10 B | *15. Regression Equation: effect size |
|---|---|

According to the guidelines suggested by J. Cohen (1988), **d = .8** is a large effect size; any effect size much larger would probably be too obvious to require an experiment.

a) What **proportion of population variance** is accounted for when d reaches this value? (Formula 10.15)

$\omega^2$ = _____

b) What **proportion of population variance** is accounted for when d is moderate in size, i.e., **d = .5**?

$\omega^2$ = _____

c) How high does d have to be for **half** of the **population variance to be accounted** for?

d = _____

| 10 C | 1. Regression | Code: R notebook |
|---|---|---|

Perform a linear regression to predict **statquiz** from **phobia** , and write out the raw-score **regression formula**.

Do the slope and Y intercept differ **significantly from zero**? **Explain** how you know.

| SLOPE: | Y-INTERCEPT: |
|---|---|
| ☐ diff from zero -or- ☐ no such evidence<br>Explain… | ☐ diff from zero -or- ☐ no such evidence<br>Explain… |

| | |
|---|---|
| **What stats quiz score** would be **predicted** for a student with a phobia rating of 9? *(by hand)*<br><br>**Stats quiz = _____** | Approximately **what phobia rating** would a student need to have in order for her predicted statquiz score to be 7.2? *(by hand)*<br><br>**Phobia rating = _____** |

| 10 | C | 2. Regression | Code: R notebook |
|---|---|---|---|

a) Perform a linear regression to predict **prequiz anxiety** from **phobia** , and write out the raw-score regression formula.

b) Repeat part a separately for men and women. (use SPSS)

| MEN | WOMEN |
|---|---|
| | |

For each gender, **what prequiz anxiety rating** would be predicted for someone reporting a phobia rating of 8? *(by hand)*

| MEN | WOMEN |
|---|---|
| **Prequiz anxiety = _____** | **Prequiz anxiety = _____** |

For which gender should you really not be making predictions at all? ☐ Men -or- ☐ Women
**Explain**.

| 11 A | *2. Independent groups: test difference in means |
|---|---|

Can the depression of psychotherapy patients be reduced by treating them in a room painted in bright primary colors, as compared to a room with a more conservative look with wood paneling? **Ten** patients answered depression questionnaires after receiving therapy in a primary-colored room, and 10 patients answered the same questionnaire after receiving therapy in a traditional room. Mean depression was lower in the colored room ( $\bar{X}_{color} = 35$) than the traditional room ($\bar{X}_{trad} = 39$) ; the standard deviations wers$_{color} = 7$ and $s_{trad} = 5$, respectively.

a) Calculate the **t value** for the test of two **independent** means (Formula 7.8)

**t( ___ ) = _____**

b) Is this t value **significant** at the .**05 (two**-tailed) level? *(check df)*

☐ YES, evidence of a difference -or- ☐ No evidence of a difference

**$t_{cv}$ = _____**

| 11 A | *3. Matched pairs: test difference in means |
|---|---|

Suppose that the patients in Exercise 2 had been **matched in pairs**, based on general depression level, before being assigned to groups.

a) If the correlation were only **.1**, how high would the **matched t value** be? (Formula 11.2)

**t( ___ ) = _____**

b) Is this matched t value **significant** at the **.05 (two**-tailed) level? *(check df)*

☐ YES, evidence of a difference -or- ☐ No evidence of a difference

**$t_{cv}$ = _____**

**Explain** any discrepancy between this result and the decision you made in part b of Exercise 2.

c) How high would the **matched t value** be if the correlation were **.3**?

**t( ___ ) = _____**

d) If the correlation were **.5**?

**t( ___ ) = _____**

| 11 A | 7. Matched pairs experiments |
|---|---|

a) Design an experiment for which it would be reasonable for the researcher to match the participants into pairs

b) Design an experiment in which it would be difficult to match participants into pairs.

| 11 A | *8. Matched pairs: very large t |
|---|---|

Suppose that the matched **t value** for a before-after experiment turns out to be **15.2**

Which of the following can be concluded?

- ☐ a.) The before and after scores must be highly correlated.
- ☐ b.) A large number of participants must have been involved.
- ☐ c.) The before and after means must be quite different (as compared to the standard deviation of the difference scores).
- ☐ d.) The null hypothesis can be rejected at the .05 level.
- ☐ e.) No conclusion is possible without more information.

| 11 B | 3. Matched pairs vs. Direct Difference | Code: R notebook |
|---|---|---|

a) Using the data from Exercise 9B6, which follows, determine whether there is a **significant** tendency for verbal GRE scores to **improve** on the second testing. Calculate the **matched t** in terms of the Pearson correlation coefficient already calculated for that exercise.
( paired t-test)

**t( ___ ) = _____** **2-tail: p = _____**

b) Recalculate the **matched t** test according to the **direct-difference method…**
(compute differences & do a 1-sample t-test)

| Verbal GRE (1) | Verbal GRE (2) | Direct Difference |
|---|---|---|
| 540 | 570 | |
| 510 | 520 | |
| 580 | 600 | |
| 550 | 530 | |
| 520 | 520 | |

**t( ___ ) = _____** **2-tail: p = _____**

… and compare the result to your answer for part a.

## 11 B *8. Matched pairs t-test & Confidence interval **Code: R notebook**

A cognitive psychologist is testing the theory that short-term memory is **mediated** by subvocal rehearsal. This theory can be tested by reading aloud a string of letters to a participant, who must repeat the string correctly after a brief delay. If the theory is correct, there will be more errors when the list contains letters that sound alike (e.g., G and T) than when the list contains letters that look alike (e.g., P and R). Each participant gets both types of letter strings, which are **randomly mixed** in the same experimental session. The number of errors for each type of letter string for each participant are shown in the following table:

| ID # | Letters that SOUND alike | Letters that LOOK alike |
|---|---|---|
| 1 | 8 | 4 |
| 2 | 5 | 5 |
| 3 | 6 | 3 |
| 4 | 10 | 11 |
| 5 | 3 | 2 |
| 6 | 4 | 6 |
| 7 | 7 | 4 |
| 8 | 11 | 6 |
| 9 | 9 | 7 |

a) Perform a **matched** t test ( α = **.05, one**-tailed) on the data above. (paired t-test)

…and state your **conclusions**.

**t( ___ ) = _____** **1-tail: p = _____**

b) Find the **95% confidence interval** for the **population <u>difference</u>** for the two types of letters.

**95% CI: ( _____ , _____ )**

## 11 B 9. Matched pairs: t-test for mean differences vs. correlation

Use R to find the **correlation coefficient** and the **regression slope** in Exercise 10B6: **Code: R notebook**

**r(____ ) = _____ , p = _____** **Slope's p = _____ (pg 7)**

a) Calculate the **matched t value** to test whether there is a significant difference ( α = **.05, two**-tailed) between the spatial ability and math scores. (paired t-test)

**t( ___ ) = _____** **2-tail: p = _____**

b) Explain how the **Pearson r** for paired data can be very **high** and statistically **significant**, while the **matched t test** for the same data **fails** to attain significance.

| 11 B | *11. Matched pairs: power |
|---|---|

Imagine that an experiment is being planned in which there are two groups, each containing 25 participants. The (unmatched) effect size ( d ) is estimated to be about .4.

a) If the groups are to be **matched**, and the **correlation** is expected to be .5, what is the **power** of the matched t test being planned, with alpha = **.05** and a **two**-tailed test? (formula 11.5 & 8.10, table a.3)

**power = _____**

b) If the correlation in the preceding example were **.7**, and all else remained the same, what would the power be?

**power = _____**

| 11 B | *13. Matched pairs: sample size estimations |
|---|---|

A matched t test is being planned to evaluate a new method for learning foreign languages. From previous research, an (**unmatched**) **effect size** of **.3**, and a **correlation** of **.6** are expected.

a) How **many participants** would be needed in each matched group to have **power** = .75, with a **two**-tailed test at **alpha** = .05? (table a.4, formula 11.5 & 8.11)

**n = _____**

b) What would your answer to part a be if **alpha** were changed to .01?

**n = _____**

| 11 C | 1. Matched pairs t-test | Code: R notebook |
|---|---|---|

A) Perform a matched-pairs **t test** to determine whether there is a significant ***increase*** in **heart rate** from **baseline** to **pre quiz**.

**t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

B) Repeat the paired t test separately for **Men** and **Women**.

**Men:** **t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

**Women:** **t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

| 11 C | 2. Matched pairs t-test | Code: R notebook |
|---|---|---|

A) Perform a matched-pairs **t test** to determine whether there is a significant ***increase*** in **anxiety** from **baseline** to **pre quiz**.

**t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

B) Perform a matched-pairs **t test** to determine whether there is a significant ***decrease*** in **anxiety** from **pre quiz** to **post quiz**.

**t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

| 11 C | 3. Matched pairs t-test | Code: R notebook |
|---|---|---|

A) Perform a matched-pairs **t test** to determine whether there is a significant difference in mean scores between the **experimental stats quiz** and the **regular stats quiz**.

**t( ___ ) = _____** **2-tail: p = _____**

☐ YES, evidence of a difference
☐ No evidence of a difference

B) Is the **correlation** between the two quizzes statistically significant?

**r = _____** **2-tail: p = _____**

☐ YES, evidence of an association
☐ No evidence of an association

Explain any **discrepancy** between the significance of the **correlation** and the significance of the matched **t test**.