

Psych 315, Winter 2021, Homework 8 Answer Key

Due Friday February 26th by midnight.

Name _____ ID _____

Section [AA] (Natalie), [AB] (Natalie), [AC] (Ryan), [AD] (Ryan), [AE] (Kelly), [AE] (Kelly)

Problem 1 Back before the pandemic, 123 students took Exams 1 and Exam 2 last year. The class average for Exam 1 was 82.36 with a standard deviation of 12.8, and the average for Exam 2 was 82.62 with a standard deviation of 15.95. Is this a significant change? Since each student took two tests, we can test the significance of this difference with a t-test for two dependent groups.

This is done by calculating the difference between Exams 1 and Exams 2 for each student. This gives a list of 123 differences. We then run a t-test that the mean of differences is different from zero.

The mean of these differences is $\bar{D} = 0.26$ points (which is the same as the differences of the means) and from the data I've calculated that the standard deviation of these difference is $s_D = 16.49$ points. In the following steps, test the hypothesis that the scores from Exam 1 are drawn from a population with a different mean than the scores from Exam 2 using an alpha value of 0.05.

a) State the null and alternative hypotheses.

$H_0 : \mu_{\bar{D}} = 0$, Exam 1 and Exam 2 are drawn from populations with the same mean.

$H_A : \mu_{\bar{D}} \neq 0$, Exam 1 and Exam 2 are drawn from populations with the different means.

b) Calculate your t-statistic

$$s_{\bar{D}} = \frac{16.49}{\sqrt{123}} = 1.49$$

$$t = \frac{0.26}{1.49} = 0.17$$

c) Find the critical value(s) of t

$$df = 123 - 1 = 122$$

$$t_{crit} = \pm 1.98 \text{ (using } df = 120)$$

d) What is your decision? State it as a complete sentence using APA format. Find the p-value using the t-calculator in the spreadsheet.

We fail to reject H_0 .

The mean of Exam 1 grades ($M = 82.36$, $SD = 12.8$) is not significantly different than the mean of Exam 2 grades ($M=82.62$, $SD = 15.95$), $t(122) = 0.17$, $p = 0.8653$.

e) What is the effect size? Is it small, medium or large?

$$\text{Effect size: } d = \frac{|\bar{D}|}{s_D} = \frac{0.26}{16.49} = 0.02$$

This is a small effect size

f) What is the probability of making a type I error for this hypothesis test?

$$P(\text{Type I error}) = \alpha = 0.05$$

g) Use the appropriate power curve to determine the power of this hypothesis test. Explain what this number means about last year's tests scores in complete sentences.

Using the power curve for $\alpha = 0.05$, 2-tails, 1 mean,

a sample size of 123, and an effect size of 0.02, the power is about 0.06.

This means that if we assume that the true mean difference in scores is

0.26 points, then the probability of correctly rejecting the null hypothesis

is 0.06

Problem 2 Also last year, for the 123 students, the correlation of scores between Exam 1 and Exam 2 is $r = 0.36$. Let's test the hypothesis that this correlation is significantly different from zero. We'll use an alpha value of 0.05.

a) Find the critical value of r using table G.

For a two-tailed with $df = 123 - 2 = 121$ and $\alpha = 0.05$, the critical value of r is ± 0.178

b) State your decision in a complete sentence.

We reject H_0 .

The correlation between Exam 1 and Exam 2 for scores is significantly different than zero, $r(121) = 0.36$, $p < 0.0001$.

Problem 3 For the 93 students in 2019, the correlation between Exam 1 and Exam 2 scores was 0.54. Is this significantly different from the correlation between Exam 1 and Exam 2 scores last year (see problem 2)? Use an alpha value of $\alpha = 0.05$.

a) Use table H to calculate the Fisher z' values for the two correlations:

$$\text{For } r_1 = 0.54, z' = 0.6042$$

$$\text{For } r_2 = 0.36, z' = 0.3769$$

b) Calculate the-statistic:

$$z = \frac{z'_1 - z'_2}{\sigma_{z'_1 - z'_2}}$$

where

$$\sigma_{z'_1 - z'_2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

$$\sigma_{z'_1 - z'_2} = \sqrt{\frac{1}{93 - 3} + \frac{1}{123 - 3}} = 0.1394$$

$$z = \frac{0.6042 - 0.3769}{0.1394} = 1.6306$$

c) Find the critical value of z

For a two-tailed test with $\alpha = 0.05$, the critical value of z is ± 1.96

d) State your decision in a complete sentence. For an extra challenge, use Table A to calculate the p-value. Don't forget to double the value for two-tailed test!

To calculate the p-value, we find using the z-table that $Pr(z > 1.6306) = 0.0515$
Since this is a two-tailed test, we double it to get $p = (0.0515)(2) = 0.103$

We fail to reject H_0 .

The correlation between the Exam 1 and Exam 2 for scores in 2019 (0.54) is not significantly different than the correlation for scores in 2020 (0.36), $z = 1.6306$, $p = 0.103$.

Problem 4 Use R's 'cor.test' function and your survey data to determine if there is a statistically significant correlation between your high school and UW GPAs. Remember, UW GPAs are in the field 'GPA_UW' and high school GPAs are in 'GPA_HS'.

Hint: modify Example 1 in the R script: ComparingOneCorrelation.R

```
# Comparing one correlation to the null hypothesis that rho = 0.
#
# If you have your raw x and y data, R's function 'cor.test' gives you a p-value for testing
# the hypothesis that a correlation was drawn from a population that has a correlation of zero.

# It takes in the x and y variables along with 'alternative' which can be "greater", "less"
# for a one-tailed test or "two.sided" for a two-tailed test. Here are the examples from the
# comparing_one_correlation tutorial.

# By default, cor.test converts the correlation of x and y into a t-statistic and then computes
# a t-test.

# Load in the survey data
survey <- read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# Let x be UW GPAs and y be high school GPAs
x <- survey$GPA_UW
y <- survey$GPA_HS
# Get rid of the NA's
goodvals = !is.na(x) & !is.na(y)
x <- x[goodvals]
y <- y[goodvals]
# cor.test runs the t-test for you:
out <- cor.test(x,y,alternative = "two.sided")
# 'estimate' is the correlation
out$estimate
      cor
0.1015848
# 'p.value' is the p-value
out$p.value
[1] 0.2192464

# 'statistic' is the t-statistic used in the test:
out$statistic
      t
1.233836

# with degrees of freedom:
out$parameter
      df
146

# Here's how to display your results in APA format:
sprintf('r(%g) = %4.2f, p = %5.4f', out$parameter, out$estimate, out$p.value)
[1] "r(146) = 0.10, p = 0.2192"
```