Lab 7: Paired Data Analysis

Objective: In this lab, you will learn how to perform a hypothesis test in the case when we have two quantitative variables collected in pairs, called a **paired *t*-test**. You will make a confidence interval for and test hypotheses about the population mean difference, μ_d. Using these, you will be able to provide a statement about how confident you are regarding your interval estimate or in your decision.

Application: Mackenzie believes that college students can run a mile faster in the afternoon then they can in the morning. She has ten of her friends run a mile early in the morning and also late in the afternoon and she records their time in seconds. For each of her ten friends Mackenzie computes a difference: time to run the morning mile (in seconds) – time to run the afternoon mile (in seconds).

Runner	AM	PM	Difference (AM – PM)
Runner 1	633.8	618.9	14.9
Runner 2	588.9	569.6	19.3
Runner 3	619.4	630.9	-11.5
Runner 4	640.9	628.5	12.4
Runner 5	613.9	574.2	39.7
Runner 6	590.0	627.3	-37.3
Runner 7	613.4	603.3	10.1
Runner 8	568.6	593.1	-24.5
Runner 9	637.0	596.2	40.8
Runner 10	648.6	613.4	35.2

Overview:

Matched or paired data results from a deliberate experimental design scheme. Mackenzie's scenario is one example of a paired design. Another example of a paired design is an experiment where rats are matched by weight, where one rat in each match receives a new diet and the other rat in the match receives a control diet.

These types of design are called **paired data designs**. Note that paired designs can occur when you have two measurements on the same individual OR when you have two individuals that have been matched or paired prior to administering a treatment.

The inference procedures for a paired data design are based on the one-sample *t*-test procedures from the previous lab. The change is that we want to estimate or test hypotheses about the population mean difference μ_d, which is generally compared to a hypothesized value of zero, indicating no difference on average.

The assumptions are similar to those assumptions made for the one sample t-test. Going back to the application, Mackenzie now must assume the sample of the ten differences in running time is a random sample. This can be verified by creating a time plot of those ten sample differences and checking for stability. She must also assume that the population of differences is normally distributed. Mackenzie can check this assumption by creating a QQ plot of those ten sample differences.

The test statistic t is created by taking the sample mean difference and dividing by the standard error of the sample mean difference, $t = \frac{\bar{d}}{se(\bar{d})}$.

If the null hypothesis is true, this test statistic has a t distribution with n-1 degrees of freedom, where n is the number of pairs. For Mackenzie's 10 runners, the degrees of freedom will be 10 – 1 = 9.

Formula Card

Population Mean		
Parameter	μ	
Statistic	$\bar{x}$	
Standard Error $\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}$		
Confidence Interval $\bar{x} \pm t^{*}\text{s.e.}(\bar{x})$		$\text{df} = n - 1$
Paired Confidence Interval $\bar{d} \pm t^{*}\text{s.e.}(\bar{d})$		$\text{df} = n - 1$
One-Sample *t*-Test $t = \frac{\bar{x} - \mu_0}{\text{s.e.}(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$		$\text{df} = n - 1$
Paired *t*-Test $t = \frac{\bar{d} - 0}{\text{s.e.}(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}}$		$\text{df} = n - 1$

Warm-Up: Check Your Understanding

Eight cars were run to determine their mileage, in miles per gallon (MPG). Then each car was given a tune-up, and run again to measure the mileage a second time. The difference in mileage was computed as difference = After MPG minus Before MPG. Assume that the selection of 8 cars represents a random sample of cars.

1. Select the appropriate alternative hypothesis to assess if on average mileage significantly improves after a tune up.

   a. H_a: $\mu_d > 0$

   b. H_a: $\mu_d < 0$

   c. H_a: $\mu_d \neq 0$

2. The researcher of course is hoping that the results of the experiment are statistically significant. What type of *p*-value would the researcher want to obtain?

   a. A large *p*-value.

   b. A small *p*-value.

   c. The magnitude of a *p*-value has no impact on statistical significance

ILP: Do Books Purchased from Barnes and Noble (In-store) Cost More on Average than If Purchased at Amazon.com (Online)?

Background: The popularity of purchasing books online has increased dramatically, and the conventional bookstore no longer dominates the sales of books. The most influential factor that sways customers into purchasing online is lower prices. A group of statistics students decided to perform a comparison of the Amazon.com prices versus local Barnes and Noble bookstore prices based on a sample of 40 books, selected from a wide range of categories. For Amazon, a standard ground shipping of $4.29 and local state tax were included in the cost. The corresponding costs are available in the R data set called **books.Rdata** (Source: Statistics group project). Do the data provide sufficient evidence to conclude that, on average, *Barnes and Noble (in-store) books are <u>more expensive</u> than Amazon.com books?*

Task: Perform a test regarding the mean difference in book price, μ_D, where the differences are computed as "*Barnes and Noble less Amazon*" (i.e. 'price at Barnes and Noble' minus 'price on Amazon').

Procedure: The appropriate inference procedure for this scenario is a paired t-test, and the specific parameter of interest is ____________________ . Think about why this is this a paired procedure.

Hypothesis Test:

1. **State the Hypotheses:** H_0: _____________________ and H_a: ______________________ ,

   where the parameter___________ represents:

   Determine Alpha: We were told the significance level was 5%.

Remember: Your hypotheses and parameter definition should always be a statement about the **population(s)** under study.

2. **Checking the Assumptions**
   a. For this scenario, we need to assume that the sampled differences are a random sample. Since the data was collected at one time point, a time plot is not appropriate to make. However we would want to learn more about how the books were selected.

   b. We also need to assume that the population of differences is normally distributed. To check this assumption, we would make a __________ plot of the ____________________. (What you would hope to see?)

3. **Compute the test statistic and calculate the *p*-value**
   Test statistic
   a. Let's create a variable for the *differences* so we can get summary statistics needed to compute the test statistic. First, we need to compute a difference for each pair. We can do this in R fairly easily by going to **Data > Manage variables in active data set > Compute new variable**. Call the new variable **difference**, and enter the difference of the two variable names in the **Expression to compute** box. You can double click on a variable name to insert it into the expression box.

b. Now, we will compute those summary statistics. Go to **Statistics > Summaries > Numerical summaries** to generate these, making sure you select your difference variable, and fill out the table below. Be sure to check the standard error option under the **Statistics** tab.

Summary Statistics			
Mean diff ($\bar{d}$)	**Std. Dev (s_d)**	**Sample size (n)**	**Std. Error**

c. Generate the *t*-test output using **Statistics > Means > Paired T Test**. Select the two variables, and specify the correct direction for the alternative hypothesis. (This should match your answer in question 1.) Use the output to fill out the following output table for this test.

Paired T Results		
t	***df***	***p-value***

d. Our sample mean difference is __________ standard errors below the hypothesized mean difference of zero.

e. What is the distribution *of the test statistic* if the null hypothesis is true?

Note: This is not the same as the distribution *of the population that the data were drawn from*, and will be the model used to find the *p*-value.

f. Now, try conducting a one mean test on the difference variable we created earlier, using the same hypotheses from problem 1. This can be found by going to **Statistics > Means > One Mean T Test**. Are the results similar to those you found in part c? Why is this the case?

One Sample T Results for the Differences		
t	***df***	***p-value***

Visualize the *p*-value:

g. Draw a picture of the *p*-value,
with labels for the distribution and x-axis.

4. Evaluate the *p*-value and Conclusion

Evaluate the *p*-value:
What is your decision at a 5% significance level? **Reject H_0** **Fail to Reject H_0**

Remember:	Reject H_0	⇔ Results statistically significant
	Fail to Reject H_0	⇔ Results not statistically significant

Conclusion:
What is your conclusion *in the context of the problem*?

Note: Conclusions should always include a reference to the population parameter of interest.
Conclusions should not be too strong; you can say that you have sufficient evidence,
but do NOT say that we have *proven* anything true or false.

Cool-Down: Setting Up Hypotheses and Writing Conclusions

The Major League Baseball Organization would like to assess whether or not players have the same batting averages during the nighttime games as during the daytime games using a 5% significance level. Eight players are selected at random and their nighttime and daytime batting averages are collected for a given period of time.

Let μ_d represent the population mean difference between the nighttime minus the daytime batting average. Write out the corresponding null and alternative hypothesis using the appropriate notation

Hypothesis Test: **H_0**:__________________________ **H_a**:__________________________

Why is this a paired t-test?

The *p*-value for this paired t-test was 0.028. Make a decision for this test and write a conclusion in context of the problem.