Lesson 10
Part 1

Confidence Intervals and Hypothesis Tests for Means from Paired data

Lesson 10 Overview

- Statistical inference from a sample to the population has been limited to data from one sample so far
  - Lesson 8: One sample confidence intervals and hypothesis tests of the population mean
  - Lesson 9: One sample confidence intervals and hypothesis tests of the population proportion
- Lesson 10 extends inference methods about the mean to paired data and data from two samples
  - Paired data – estimation and tests of mean difference
  - Two samples – estimation and tests of difference of means.

Paired Data

- Paired data are the result of subjects being measured twice: typically before and after an intervention
- The research question is whether or not the intervention made a difference – is there a change?
- The parameter of interest is the mean of the differences.

Hypothesis Test of the Mean Difference

- The hypothesis test of the Mean difference is called a Paired t-test
  - The null hypothesis is that there is no change:
    - H_o: the mean difference = 0
  - The alternative hypothesis is that there is a change
    - H_A: the mean difference ≠ 0 for a two-tailed test
    - H_A: the mean difference < 0 or >0 for a one-tailed test

Paired t-test Procedure

1. State the hypotheses
2. Calculate test statistic
2. Calculate the p-value
3. State the conclusion of the test

Example

- The 6 minute walk test can be used to measure physical stamina: How far (in ft) can the patient walk in 6 minutes?

- A study was designed to evaluate whether a new respiratory therapy intervention for COPD (Chronic Obstructive Pulmonary Disease) patients resulted in a change in 6 minute distance walked

Example

- Distance walked in 6 minutes was measured for a random sample of 16 COPD patients before (Pre) and after (Post) 90 days on the new therapy.
  - Assume that distribution of the mean difference of distance walked is approximately normally distributed
- Conduct a Paired t-test to test whether there was a significant change in 6 min. distance

Pre	Post	DIFF
1025	1050	25
630	620	-10
860	875	15
675	675	0
1120	1150	30
590	600	10
540	650	110
615	625	10
745	775	30
730	750	20
945	950	5
630	625	-5
770	775	5
1030	1050	20
530	550	20
1120	1150	30

6 Minute walk data was collected for 16 COPD patients

$$\bar{d} = 19.69$$

$$s_d = 27.11$$

1. State the hypotheses

- The null and alternative hypotheses are about the population mean difference notated with the Greek letter delta: δ
- H_0: $\delta = 0$
- H_A: $\delta \neq 0$

(Note: Use a one-sided test if you know ***before*** collecting data that change can only be in one direction or if interest is in one direction only).

Null hypothesis for paired t-test

- The null hypothesis that $\delta = 0$ is the same as stating that, on average, there is no change. Why?

Step 2. Calculate test statistic

- The test statistic for a paired t-test is a t-statistic

$$t = \frac{\bar{d} - 0}{SE(diff)}$$

$\bar{d}$ is the sample mean difference

0 is the hypothesized difference

$$SE(diff) = \frac{S_d}{\sqrt{n}}$$

$n =$ number of subjects

Step 2. Calculate the test statistic

$$t = \frac{\bar{d} - 0}{S_d / \sqrt{n}} =$$

Step 3: P-value

- The p-value of the test statistic is the area beyond ± 2.905 under the t-distribution with 15 df
- TDIST(________,_____,_____) = 0.011

Step 4. State the Conclusion of the test

Inference for Paired Data

- In addition to the paired t-test, Confidence intervals of the population mean difference can be constructed
- The confidence interval provides information about the *precision* of the estimate of the mean difference.
- The confidence interval can also be used to evaluate change. If the 95% confidence interval does not contain the value 0, the change is significant at the alpha = 0.05 level.

Confidence Interval of Mean Difference

- The confidence interval of the population mean difference is constructed like a confidence interval for a population mean

$$\text{Mean difference} \pm \text{Confidence Coefficient} * \text{SE (diff)}$$

$$\bar{d} \pm t^{*}_{n-1}\left(S_d \big/ \sqrt{n}\right)$$

Confidence Interval of Mean Difference

- 95% confidence interval for mean difference in 6 minute walk distance:
- 19.69 = average difference
- Confidence Coefficient: TINV(0.05, 15) = 2.13

$$SE(diff) = {S_d}/{\sqrt{n}} =$$

CI Interpretation

- The 95% confidence interval does not contain the value 0.
- Conclusion?

Requirements for Paired t-test

- The sample data consist of two measurements on each subject
- The mean is an appropriate summary statistic of the data (data are quantitative)
- The subjects are randomly selected from the population of interest
- Data are approximately normally distributed or sample size is large (25-30)

EXAMPLE: ATHEROSCLEROSIS

- A radiologist assess the viability of carotid B-mode ultrasonography to evaluate artherosclerosis in 20 randomly selected patients.
- Two tests are conducted, 48 hours apart. The same technician performs all tests. The space between carotid lesions, known as the minimum residual lumen (MRL), is measured by the same radiologist. The radiologist was unaware of the purpose of the study.
- The mean difference is 0.65 mm (standard deviation of the difference is 0.43). Assume relevant distributions are normal.

[Adapted from: Diane Essex-Sorlie, Medical Biostatistics and Epidemiology, Page 187]

The 99% CI for δ : (0.375, 0.925). Test Ho: No Mean difference vs. Ha: A Mean difference exists.

1. 0 doesn't lie within the 99% CI, we would reject H_0 (α=0.01).
2. 0 doesn't lie within the 99% CI, we would reject H_0 (α=0.01).
3. 0 lies within the 99% CI, we fail to reject H_0 (α=0.01).

What type of error might you have made based on your conclusion?

1. No error, my conclusion is correct.
2. Type I error
3. Type II error
4. Power too high: a statistically significant result that is not scientifically relevant.

Suppose the two sided hypothesis test was designed with the probability of a type I error at 5% (instead of 1%) , What conclusion can you make?

1. Fail to reject the null.
2. Accept the null.
3. Reject the null.
4. Accept the Alternative.

Over-the-Counter Medication Effects

- A researcher is interested in whether an over-the-counter multi-symptom cold medicine raises systolic blood pressure as an undesirable side effect.
- SBP is measured in 15 individuals, immediately prior to taking the recommended dose of the cold medicine, and again 30 minutes later.
- Average increase in SBP was 6.40 mm Hg with standard deviation in the increase of 8.458 mm Hg.
- Is there evidence to show a significant increase in SBP after taking this medicine (α = 0.05)?

What is the appropriate approach to solve this problem?

1. 95% CI for a paired difference.
2. Normal Approximation to the Binomial for two proportions.
3. T test for the difference of two independent means
4. T test for a single mean or a mean of a paired difference.

What is the p-value for this test?

1. $P(T_{14}>2.93)$
2. $2*P(Z>2.93)$
3. $P(Z >2.93)$
4. $2*P(T_{14}< -2.93)$
5. $2*P(T_{29}>2.93)$

What is the p-value for this test?

1. $P(T_{14}>2.93)$
2. $2*P(Z>2.93)$
3. $P(Z >2.93)$
4. $2*P(T_{14}< -2.93)$
5. $2*P(T_{29}>2.93)$

The p-value of the test = 0.011, what does this mean?

1. The probability the null is true = 0.011
2. If the null is indeed true, the probability of observing a difference of 6.4 is 0.011.
3. The probability the alternative is true given the difference of 6.4.
4. We should accept the alternative.

Readings and Assignments

- Reading: Chapter 5 pages 114 – 118
- Work through Exercise 1 on the Lesson 10 Practice exercises
- Excel Module 10 examples
  - Use the Data Analysis tool for the Paired t-test
  - Select: t-test Paired two sample for Means
- Begin Homework 7 Problems