Lesson 10
Part 2

Confidence Intervals and Hypothesis Tests for Means from Two samples

Lesson 10 Part 2 Overview

- Lesson 10 Part 1 covered Confidence Intervals of the mean difference from paired data and the Paired t-test
  - The methods for inference of the mean difference are similar to methods used for one-sample
- Lesson 10 Part 2 covers inference about the difference between the means from two independent samples.
  - Confidence intervals of the difference of means
  - Two sample t-tests

Hypothesis Testing: Comparing Two Means

- A common goal of inference is to compare the average responses in ***two*** groups.
- Each group is considered to be a sample from a distinct population.
- The average responses in each group are independent of those in the other group.
- The sample sizes in the two groups need not be the same.

Recall: Steps in Hypothesis Testing

1. State the null hypothesis H_0 and the alternative hypothesis H_a.
2. Calculate the value of the test statistic on which the test will be based.
3. Find the p-value for the observed data.
4. State a conclusion.

Step 1: State the Hypotheses

- Usually the null hypothesis is a statement of “no effect” or “no difference”.
- H_a is the statement of what we hope or suspect is true.
- Hypotheses refer to some population or model and are stated in terms of population parameters.

Step 1: One sided or 2-sided test?

- If you do not have a specific direction firmly in mind in advance (before looking at the data), use a 2-sided alternative hypothesis.
- Some statisticians argue that you should always use a 2-sided alternative, but we will be working with both types.

Step 2: Calculate the Test Statistic

- When H_0 is true we would expect the sample statistic to take a value near the parameter estimate specified by H_0.
- Values of the sample statistic far from the parameter value specified by H_0 gives evidence against H_0.
- The alternative hypothesis determines which directions count against H_0 (one or two-sided).

Step 2: Test Statistic

- A test statistic measures compatibility between the null hypothesis and the sample data.

- We calculate the test statistic *assuming that the null hypothesis is true*.

- When we have a one-sided hypothesis we use the value of the parameter in H_0 that is closest to H_a to calculate our test statistic.

Step 3: Calculating the P-value

- A test of significance assesses the evidence against the null hypothesis in terms of probability.

- The p-value is the probability, computed assuming H_0 is true, that the test statistic would take a value as extreme or more extreme than that actually observed.

- The smaller the p-value the more evidence against H_0.

Step 4: Making a Conclusion

- We need to compare our p-value with a fixed value that we regard as decisive. This value determines how much evidence against H_0 we will require to reject H_0 and we call it the significance level (α).

- With a significance level set at 0.05 we are requiring that the data give evidence against H_0 so strong that it would happen no more than 5% of the time when H_0 is true.

- If the p-value is as small or smaller than α, we say that the data are statistically significant at level α and we would reject H_0.

Hypothesis Testing: Comparing Two Means

- Identify two independent populations.
- Draw a simple random sample of size n_1 from population 1 and a simple random sample of size n_2 from population 2.
- Compute the mean for each sample $\bar{x}_1$, $\bar{x}_2$.
- Formulate hypothesis test based on the difference of the means $\bar{x}_1-\bar{x}_2$.

Hypothesis Testing: Comparing Two Means

- Given two normally distributed populations with mean μ_1 and μ_2, and standard deviations σ_1 and σ_2.

$$(x_1 - x_2) \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

Hypothesis Testing: Comparing Two Means

- Step 1: State your hypotheses
  - H_0: $\mu_1 - \mu_2 = u_o$
  - H_a: $\mu_1 - \mu_2 \neq u_o$ (two-sided)

    Or
  - H_a: $\mu_1 - \mu_2 < u_o$ (one-sided)

    Or
  - H_a: $\mu_1 - \mu_2 > u_o$ (one-sided)

Hypothesis Testing: Comparing Two Means

- Step 2 : Calculate your test statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_o)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

OR

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_o)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Known σ_1 and σ_2

Unknown σ_1 and σ_2

Degrees of freedom = smaller of n_1-1 or n_2-1

Hypothesis Testing: Comparing Two Means

- Step 3: Calculate the p-value

$$2 * p(|z| \geq z^*)$$ OR $$2 * p(|t| \geq t^*)$$

$$p(z \leq z^*)$$ $$p(t \leq t^*)$$

$$p(z \geq z^*)$$ $$p(t \geq t^*)$$

Degrees of freedom = smaller of n_1-1 or n_2-1

- Step 4 : Make a conclusion

If p-value > α, then do **not** reject H_o

If p-value < α, then reject H_o

Hypothesis Testing Comparing Two Means: An Example

- The effect of environmental exposure to lead on intellectual development is investigated using two randomly selected samples of 7 year old children from similar backgrounds but with different lead exposures.

Hypothesis Testing Comparing Two Means: An Example

- Serum lead levels in group 1 > 30 ug/dL
- Serum lead levels in group 2 > 10 ug/dL

- Conflicting publishing results lead investigators to hypothesize that a difference in intelligence scores exists between these two groups.

Hypothesis Testing Comparing Two Means: An Example

- Does a significant difference exist between the mean intelligence test score in these two groups?

- The data for intelligence test score is summarized below:

$n_1=61$	$\bar{x}_1=94$	$s_1=17$
$n_2=41$	$\bar{x}_2=101$	$s_2=8$

Hypothesis Testing Comparing Two Means: An Example

- Step 1: State your hypotheses (set α=.01)
  - H_0:
  - H_a:
- Step 2 : Calculate your test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_o)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} =$$

Hypothesis Testing Comparing Two Means: An Example

- Step 3: Calculate the p-value

$$2 * p(|t| \geq 2.789) < 0.008$$

$$= tdist(\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_)$$

Degrees of freedom =

- Step 4 : Conclusion?

Two sample Problem with σ_1 and σ_2 unknown and assumption $\sigma_1 = \sigma_2$.

- Inference is based on 2 independent SRS, one from each population.

Population	Sample Size	Sample Mean	Sample Standard Deviation
1	n_1	$\bar{x}_1$	s_1
2	n_2	$\bar{x}_2$	s_2

POOLED Two sample Problem assumption $\sigma_1 = \sigma_2$.

$$\bar{x}_1 \sim N(\mu_1, {\sigma_1}^2) \qquad \bar{x}_2 \sim N(\mu_2, {\sigma_2}^2)$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_o}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{with} \quad n_1 + n_2 - 2 \quad \text{degrees of freedom}$$

$$\text{with } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

To Pool or Not to Pool

Two methods can be used to test for equality of variance.

- If the ratio of the largest SD/smallest SD < 2, the standard deviations are considered equal and the equal variance method is use.

- The F-test is a test of equality of variances. If the F-test is not significant (p>0.05) the variances are considered equal. Excel can be used to find the p-value of the F-test (more details on this later).

Two-sample t-test Example

- A study is designed to test whether or not there is a difference in mean pain scores between patients on standard treatment (control) and patients on a new pain medication.
- 50 patients were randomly assigned to treatment or control. There were 25 patients in each group.
- Test the null hypothesis of no difference in mean pain scores between the treatment group and the control group
- This is a two-tailed two sample t-test

Two-sample t-test Steps 1-2

Step 1. State the null and alternative hypotheses

Ho:

Ha:

Step 2. Identify the appropriate test statistic

A t-statistic will be used for this test to compare mean pain scores between the two groups

Use the pooled standard deviation if the variances are approximately equal – this will be checked before calculating the test statistic

Collect the Data

- Pain scores were obtained for the 25 patients in each group
- A higher pain score indicates more severe pain
- Data are summarized in the table

	Control	Treatment
Sample size (n)	25	25
Sample Mean	8	3.92
Sample Standard deviation (SD)	4.601	2.644

Check for equality of variance

- Check for equality of variance using the ratio of standard deviations:

Pooled or Un-pooled?

Test Statistic

Step 2. Calculate the t-statistic

$$S_p = \sqrt{\frac{(n_c - 1)S_c^2 + (n_T - 1)S_T^2}{n_C + n_T - 2}} = \sqrt{\frac{(25-1)(4.601^2) + (25-1)(2.644^2)}{25 + 25 - 2}} = 3.75$$

and

$$t = \frac{(\overline{x}_C - \overline{x}_T) - 0}{S_p \sqrt{\left(\frac{1}{n_C} + \frac{1}{n_T}\right)}} = \frac{(8 - 3.92)}{3.75\sqrt{\left(\frac{1}{25} + \frac{1}{25}\right)}} = 3.84$$

Pvalue and Conclusion

- Test decision using the p-value method

This test statistic has a t-distribution with 48 degrees of freedom.

P-value =**TDIST**(3.84,48,2)=0.00035

Note: For TDIST you always use the positive value of the t-statistic.

Since 0.00035 < 0.05, the null hypothesis of no difference in mean pain score between the two groups is rejected.

Conclude that mean pain score is significantly different in the control group.

Confidence Interval of Difference between two means

- The other inference method for means from two samples is a confidence interval for the difference between two means

$$(\bar{x}_1 - \bar{x}_2) \pm \text{Confidence Coefficient} * SE\ (diff)$$

Confidence Interval of Difference between two means- Unpooled

$$(\bar{x}_1 - \bar{x}_2) \pm t_{df} * \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

df = smaller of n1 – 1 or n2 - 1

Confidence Interval of Difference between two means - Pooled

$$(\bar{x}_1 - \bar{x}_2) \pm t_{df} * \mathrm{s_p}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}, where$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$

df = n1 + n2 -2

95% Confidence Interval for difference in mean pain score

- Use the pain score data to calculate the 95% confidence interval for the difference in mean pain scores between the control and treatment groups

$$(8 - 3.92) \pm 2.011 * 1.061 = (1.95, 6.21)$$

- Control mean = 8, Treatment mean = 3.92
- Confidence coefficient = 2.011
- SE(diff) is calculated using the pooled SD = 1.061

Confidence Interval Interpretation

- The 95% confidence interval does not contain 0 so we can conclude that this difference is significant at the alpha = 0.05 level

Paired data or Independent Groups?

The following can help identify if data are paired or not

- Indications that the data are paired
  - 'Before / After' or 'Pre / Post' measurements are indicated
  - Each subject is measured twice
  - Subjects matched
- Indications that data are from independent groups
  - Study description identifies two populations of interest

Duration of cold Symptoms study

- Set up a hypothesis test to test the null hypothesis of equal mean duration of cold symptoms in the two groups.
- Allow for a difference in means in either direction

Notation for mean duration of cold symptoms:

Group	Population Mean	Sample Mean
Vit C	μ_C	$\bar{x}_C$
Multi-vitamin	μ_M	$\bar{x}_M$

The appropriate test to use is

1. One-sample z-test of proportion
2. Paired t-test of mean difference
3. Two sample t-test of difference in means

The correct null hypothesis for this test:

1. $\mu_C = \mu_M = 0$
2. $\mu_C - \mu_M = 0$
3. $\mu_C = \mu_M$
4. Both 1 and 3
5. Both 2 and 3

The correct alternative hypothesis is

1. $\mu_C - \mu_M \neq 0$
2. $\mu_C - \mu_M > 0$
3. $\mu_C \neq \mu_M$
4. Both 1 and 3
5. Both 2 and 3

Study Data

	Vit. C Group	Multi-vitamin group
Sample size	100	100
Sample Mean	11	12
Sample SD	3	4

Should you used the pooled or un-pooled test statistic?

1. Matched pairs pooled
2. Un-pooled
3. Pooled
4. Cannot be determined from the information given.

What is the p-value for this test?

$$t = \frac{\bar{x}_C - \bar{x}_M}{SE(diff)} = \frac{11-12}{0.5} = \frac{-1}{0.5} = -2$$

1. $P(t_{99} < -2)$
2. $2 * P(t_{99} < -2)$
3. $2 * P(t_{199} < -2)$
4. $2 * P(t_{198} < -2)$
5. $-2 * P(t_{198} > 2)$

Suppose you reject the null hypothesis, what type of error might have occurred?

1. Type I
2. Type II
3. 1-Type II
4. None. The correct decision was made.
5. Fail to reject the null.

If you constructed a 95% confidence interval for the difference in mean duration between Vit C group and Multi-vitamin group: $\mu_C - \mu_M$, the CI limits would:

1. Both be positive
2. Both be negative
3. Have a negative lower limit and a positive upper limit

Readings and Assignments

- Reading: Chapter 6 Pages 135 - 143
- Lesson 10 Practice Exercises: Exercise 2
- Lesson 10 Excel Module Examples
- Complete Homework 8 and submit by due date.
  - You will use the Data Analysis tool for some of the homework problems