Comparison of Two Independent Samples

• Many times in the sciences it is useful to compare two groups
  – Male vs. Female
  – Drug vs. Placebo

\[ \mu_1, \sigma_1 \quad \text{Population 1} \]
\[ \mu_2, \sigma_2 \quad \text{Population 2} \]
\[ \bar{y}_1, s_1 \quad \text{Sample 1} \]
\[ \bar{y}_2, s_2 \quad \text{Sample 2} \]

Comparison of Two Independent Samples (cont')

• Two Approaches for Comparison
  – Confidence Intervals
    • we already know something about CI's
  – Hypothesis Testing
    • this will be new

• What seems like a reasonable way to compare two groups?
• What parameter are we trying to estimate?

RECALL: The sampling distribution of \( \bar{y} \) was centered at \( \mu \), and had a standard deviation of \( \sigma/\sqrt{n} \).

We’ll start by describing the sampling distribution of \( \bar{y}_1 - \bar{y}_2 \).

– Mean of \( \mu_1 - \mu_2 \)
– Standard deviation of \( \sqrt{s_1^2/n_1 + s_2^2/n_2} \)

• What seems like appropriate estimates for these quantities?

Standard Error of \( \bar{y}_1 - \bar{y}_2 \)

• We know \( \bar{y}_1 - \bar{y}_2 \) estimates \( \mu_1 - \mu_2 \),
• What we need to describe next is the precision of our estimate, \( SE_{\bar{y}_1 - \bar{y}_2} \)

\[
SE_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{SE_1^2 + SE_2^2}
\]

Standard Error of \( \bar{y}_1 - \bar{y}_2 \) (cont')

Example: A study is conducted to quantify the benefits of a new cholesterol lowering medication. Two groups of subjects are compared, those who took the medication twice a day for 3 years, and those who took a placebo. Assume subjects were randomly assigned to either group and that both groups data are normally distributed. Results from the study are shown below:

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication</td>
<td>209.8</td>
<td>44.3</td>
<td>14.0</td>
</tr>
<tr>
<td>Placebo</td>
<td>224.3</td>
<td>46.2</td>
<td>14.6</td>
</tr>
</tbody>
</table>
Standard Error of $\bar{y}_1 - \bar{y}_2$ (cont')

Example: Cholesterol medicine (cont')

Calculate an estimate of the true mean difference between treatment groups and this estimate’s precision.

- First, denote medication as group 1 and placebo as group 2

$$
\bar{y}_1 - \bar{y}_2 = 209.8 - 224.3 = -14.5
$$

<table>
<thead>
<tr>
<th>Medication</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>10</td>
</tr>
<tr>
<td>$s$</td>
<td>44.3</td>
</tr>
<tr>
<td>SE</td>
<td>14.0</td>
</tr>
</tbody>
</table>

$$
SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{44.3^2}{10} + \frac{46.2^2}{10}} = 20.24
$$

Pooled vs. Unpooled

- $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ is know as an “unpooled” version of the standard error
- There is also a “pooled” SE

- First we describe a “pooled” variance, which can be thought of as a weighted average of $s_1^2$ and $s_2^2$

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

CI for $\mu_1 - \mu_2$

- RECALL: We described a CI earlier as:
  - the estimate $\pm$ (an appropriate multiplier)$\times$SE

- A 100(1-α)% confidence interval for $\mu_1 - \mu_2$ p.227

$$
(\bar{y}_1 - \bar{y}_2) \pm t(df) \times SE(\bar{y}_1 - \bar{y}_2)
$$

where $df = \frac{(SE_1^2 + SE_2^2)^2}{\left(\frac{SE_1^2}{(n_1 - 1)} + \frac{SE_2^2}{(n_2 - 1)}\right)}$

Pooled vs. Unpooled (cont')

- Then we use the pooled variance to calculate the pooled version of the standard error

$$
SE_{pooled} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)
$$

- NOTE: If $n_1 = n_2$ or if $s_1 = s_2$ the pooled and unpooolled will give the same answer for $SE(\bar{y}_1 - \bar{y}_2)$

- It is when $n_1 \neq n_2$ that we need to decide whether to use pooled or unpooolled:
  - if $s_1 = s_2$ then use pooled (unpooolled will give similar answer)
  - if $s_1 \neq s_2$ then use unpooolled (pooled will NOT give similar answer)

CI for $\mu_1 - \mu_2$ (cont')

- RESULT: Because both methods are similar when $\lambda = s_1$
  - and the pooled version is not valid when $n_1 \neq n_2$, we choose to always use the unpooolled version ☺
  - Why all the torture? This will come up again in chapter 11

Example: Cholesterol medication (cont')

Calculate a 95% confidence interval for $\mu_1 - \mu_2$.

We know $\bar{y}_1 - \bar{y}_2$ and $SE(\bar{y}_1 - \bar{y}_2)$ from the previous slides. Now we need to find the t multiplier

$$
df = \frac{(14^2 + 14.6^2)^2}{(10 - 1)} = \frac{167411.9056}{9317.023} = 17.97 = 17
$$

Round down to be conservative

- NOTE: Calculating that df is not really that fun, a quick rule of thumb for checking your work is:
  $$
n_1 + n_2 - 2
$$
CI for $\mu_1 - \mu_2$ (cont')

$(\bar{x}_1 - \bar{x}_2) \pm t(\text{df}) \frac{S_{\bar{x}_1 - \bar{x}_2}}{\sqrt{n}}$

$= -14.5 \pm t(17) \frac{20.24}{\sqrt{21}}$

$= -14.5 \pm 2.110(20.24)$

$= (-57.21, 28.21)$

CONCLUSION: We are highly confident at the 0.05 level, that the true mean difference in cholesterol between the medication and placebo groups is between -57.02 and 28.02 mg/dL.

Note the change in the conclusion of the parameter that we are estimating. Still looking for the 5 basic parts of a CI conclusion (see slide 38 of lecture set 5).

Hypothesis Testing: The independent t test

• The idea of a hypothesis test is to formulate a hypothesis that nothing is going on and then to see if collected data is consistent with this hypothesis (or if the data shows something different)
  - Like innocent until proven guilty

• There are four main parts to a hypothesis test:
  - hypotheses
  - test statistic
  - p-value
  - conclusion

Hypothesis Testing: #1 The Hypotheses

• There are two hypotheses:
  - Null hypothesis (aka the "status quo" hypothesis)
    • denoted by $H_0$
  - Alternative hypothesis (aka the research hypothesis)
    • denoted by $H_a$

Example: Cholesterol medication (cont')

Suppose we want to carry out a hypothesis test to see if the data show that there is enough evidence to support a difference in treatment means.
Find the appropriate null and alternative hypotheses.

$H_0: (\mu_1 - \mu_2) = 0$
(no statistical difference in the true means of the medication and placebo groups)

$H_a: (\mu_1 - \mu_2) \neq 0$
(a statistical difference in the true means of the medication and placebo groups, medication has an effect on cholesterol)
Hypothesis Testing: #2 Test Statistic

• A test statistic is calculated from the sample data
  – it measures the "disagreement" between the data and the null hypothesis
    • if there is a lot of "disagreement" then we would think that the data provide evidence that the null hypothesis is false
    • if there is little to no "disagreement" then we would think that the data do not provide evidence that the null hypothesis is false

\[ t_s = \frac{(\bar{y}_1 - \bar{y}_2) - 0}{SE_{\bar{y}_1 - \bar{y}_2}} \]

subtract 0 because the null says the difference is zero

Example: Cholesterol medication (cont')

Calculate the test statistic

\[ t_s = \frac{(209.8 - 224.3) - 0}{20.24} = -0.716 \]

– Great, what does this mean?
  • \( \bar{y}_1 \) and \( \bar{y}_2 \) differ by about 0.72 SE's
  • this is because \( t_s \) is the measure of difference between the sample means expressed in terms of the SE of the difference

How far is far?

• How do we use this information to decide if the data support Ho?
  – Perfect agreement between the means would indicate that \( t_s = 0 \), but logically we expect the means do differ by at least a little bit.
  • The question is how much difference is statistically significant?
    – If Ho is true, it is unlikely that \( t_s \) would fall in either of the far tails
    – If Ho is false it is unlikely that \( t_s \) would fall near 0

Hypothesis Testing: #3 P-value

• What this means is that we can think of the p-value as a measure of compatibility between the data and Ho
  – a large p-value (close to 1) indicates that \( t_s \) is near the center (data support Ho)
  – a small p-value (close to 0) indicates that \( t_s \) is in the tail (data do not support Ho)
Hypothesis Testing: #3 P-value (cont’)

- Where do we draw the line?
  - How small is small for a p-value?
- The threshold value on the p-value scale is called the significance level, and is denoted by \( \alpha \)
  - The significance level is chosen by whomever is making the decision (BEFORE THE DATA ARE COLLECTED!)
  - Common values include 0.1, 0.05 and 0.01
- Rules for making a decision:
  - If \( p \leq \alpha \) then reject \( H_0 \) (statistical significance)
  - If \( p > \alpha \) then fail to reject \( H_0 \) (no statistical significance)

Example: Cholesterol medication (cont’)

Find the p-value that corresponds to the results of the cholesterol lowering medication experiment

We know from the previous slides that \( t = -0.716 \) (which is close to 0)

This means that the p-value is the area under the curve beyond \( \pm 0.716 \) with 18 df.

Example: Cholesterol medication (cont’)

Using the t table we can find the area under the curve beyond \( \pm 0.716 \) with 18 df to be:

\[ p > 2(0.2) = 0.4 \]

This is called bracketing the p-value. We can get an exact p-value from the computer.

Hypothesis Testing and your HW

- On your HW problems you need to include all 4 parts of a hypothesis test
  - Even if the book doesn’t ask for it
  - A hypothesis test is a test with 4 parts
  - How can you identify a hypothesis test problem?
- Important parts of Hypothesis test conclusions:
  1. Decision (significance or no significance)
  2. Parameter of interest
  3. Variable of interest
  4. Population under study
  5. (optional but preferred) P-value