T-test for dependent Samples
(ak.a., Paired samples t-test, Correlated Groups Design, Within-Subjects Design, Repeated Measures, ........)

The t Test for Dependent Samples

• Repeated-Measures Design
  – When you have two sets of scores from the same person in your sample, you have a repeated-measures, or within-subjects design.
  – You are more similar to yourself than you are to other people.

Difference Scores

• The way to handle two scores per person, or a matched pair, is to make difference scores.
  – For each person, or each pair, you subtract one score from the other.
  – Once you have a difference score for each person, or pair, in the study, you treat the study as if there were a single sample of scores (scores that in this situation happen to be difference scores).

A Population of Difference Scores with a Mean of 0

• The null hypothesis in a repeated-measures design is that on the average there is no difference between the two groups of scores.
• This is the same as saying that the mean of the sampling distribution of difference scores is 0.

The t Test for Dependent Samples

• You do a t test for dependent samples the same way you do a t test for a single sample, except that:
  – You use difference scores.
  – You assume the population mean is 0.

\[
t = \frac{\overline{D} - \mu_{D_{hyp}}}{s_D} \quad t = \frac{\overline{D} - \mu_{D_{hyp}}}{s_D}
\]

\[
s_D = \sqrt{\frac{nD^2 - (\overline{D})^2}{n(n-1)}}
\]

The t Test for Dependent Samples
The *t* Test for Dependent Samples: An Example

Eight individuals indicated their attitudes toward socialized medicine before and after listening to a pro-socialized medicine lecture. Attitudes were assessed on a scale from 1 to 7, with higher scores indicating more positive attitudes. The attitudes before and after listening to the lecture were as indicated in the second and third columns of the table. Test for a relationship between the raw scores of attitudes toward socialized medicine using a correlated groups *t* test.

<table>
<thead>
<tr>
<th>Individual</th>
<th>Before speech</th>
<th>After speech</th>
<th>Difference (Δ)</th>
<th>( \bar{D} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>6</td>
<td>-3</td>
<td>-1.5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>-2</td>
<td>-1.4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>4</td>
<td>-2</td>
<td>-1.0</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>2.0</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

\( \bar{D} = \frac{-1.6}{8} = -0.2 \)

- Calculated the mean difference.

- **State the research hypothesis.**
  - Does listening to a pro-socialized medicine lecture change an individual's attitude toward socialized medicine?

- **State the statistical hypotheses.**
  - \( H_0 : \mu_D = 0 \)
  - \( H_A : \mu_D \neq 0 \)

- **Set the decision rule.**

  \( \alpha = .05 \)

  \( df = \text{number of difference scores} - 1 = 8 - 1 = 7 \)

  \( t_{\text{crit}} = 2.365 \)

- **Calculate the test statistic.**

  \[ t = \frac{\bar{D} - \mu_D}{s_D} = \frac{-0.2}{0.42} = -4.76 \]

- **Decide if your results are significant.**
  - Reject \( H_0 \), -4.76<-2.365

- **Interpret your results.**
  - After the pro-socialized medicine lecture, individuals' attitudes toward socialized medicine were significantly more positive than before the lecture.
Issues with Repeated Measures Designs

- Order effects.
  - Use counterbalancing in order to eliminate any potential bias in favor of one condition because most subjects happen to experience it first (order effects).
  - Randomly assign half of the subjects to experience the two conditions in a particular order.
- Practice effects.
  - Do not repeat measurement if effects linger.

The t Test for Independent Samples

- Observations in each sample are independent (not from the same population) of each other.
- We want to compare differences between sample means.
  \[ t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_{hyp}}{s_{\bar{X}_1 - \bar{X}_2}} \]

Sampling Distribution of the Difference Between Means

- Imagine two sampling distributions of the mean...
- And then subtracting one from the other...
- If you create a sampling distribution of the difference between the means...
  - Given the null hypothesis, we expect the mean of the sampling distribution of differences, \( \mu_1 - \mu_2 \), to be 0.
  - We must estimate the standard deviation of the sampling distribution of the difference between means.

Pooled Estimate of the Population Variance

- Using the assumption of homogeneity of variance, both \( s_1 \) and \( s_2 \) are estimates of the same population variance.
- If this is so, rather than make two separate estimates, each based on some small sample, it is preferable to combine the information from both samples and make a single pooled estimate of the population variance.
  \[ s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \]

Pooled Estimate of the Population Variance

- The pooled estimate of the population variance becomes the average of both sample variances, once adjusted for their degrees of freedom.
  - Multiplying each sample variance by its degrees of freedom ensures that the contribution of each sample variance is proportionate to its degrees of freedom.
  - You know you have made a mistake in calculating the pooled estimate of the variance if it does not come out between the two estimates.
  - You have also made a mistake if it does not come out closer to the estimate from the larger sample.
- The degrees of freedom for the pooled estimate of the variance equals the sum of the two sample sizes minus two, or \( (n_1 - 1) + (n_2 - 1) \).
Estimating Standard Error of the Difference Between Means

\[ s_{p}^{2} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \]

\[ s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} \]

\[ t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{s_{\bar{x}_1 - \bar{x}_2}} \]

The \( t \) Test for Independent Samples: An Example

• Stereotype Threat

<table>
<thead>
<tr>
<th>Control Subjects</th>
<th>Threat Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 9 12 8</td>
<td>7 8 7 2</td>
</tr>
<tr>
<td>9 13 12 13</td>
<td>6 9 7 10</td>
</tr>
<tr>
<td>13 7 6</td>
<td>5 0 10 8</td>
</tr>
</tbody>
</table>

"Trying to develop the test itself."  "This test is a measure of your academic ability."

• State the research question.
  – Does stereotype threat hinder the performance of those individuals to which it is applied?

• State the statistical hypotheses.

\[ H_0 : \mu_1 - \mu_2 \geq 0 \]
\[ H_1 : \mu_1 - \mu_2 < 0 \]

or

\[ H_0 : \mu_1 \geq \mu_2 \]
\[ H_1 : \mu_1 < \mu_2 \]

• Set the decision rule.

\[ df = (n_1 - 1) + (n_2 - 1) = (11 - 1) + (12 - 1) = 21 \]
\[ t_{crit} = -1.721 \]

• Calculate the test statistic.

\[ s^2_p = \frac{12(621) - (79)^2}{12(11)} = 9.18 \]
\[ s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}} \]

\[ t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{s_{\bar{x}_1 - \bar{x}_2}} \]

\[ \bar{x}_1 = \frac{79}{12} = 6.58 \]
\[ \bar{x}_2 = \frac{106}{11} = 9.64 \]

\[ s^2_p = \frac{11(122) - (106)^2}{11(10)} = 10.05 \]
\[ s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}} \]

\[ s^2_{\bar{x}_1 - \bar{x}_2} = \frac{(n_1 - 1)s^2_1 + (n_2 - 1)s^2_2}{(n_1 - 1) + (n_2 - 1)} \]

\[ s^2_{\bar{x}_1 - \bar{x}_2} = \frac{(12 - 1)9.18 + (11 - 1)10.05}{(12 - 1) + (11 - 1)} = 9.59 \]
The t Test for Independent Samples: An Example

• Calculate the test statistic.

\[ t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}} \]

\[ s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \]

\[ \bar{X}_1 = 6.58 \quad \bar{X}_2 = 9.64 \]

\[ s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{9.59^2}{12} + \frac{9.59^2}{11}} = 1.29 \]

\[ t = \frac{6.58 - 9.64}{1.29} = -2.37 \]

The t Test for Independent Samples: An Example

• Decide if your result is significant.

– Reject \( H_0 \), \( -2.37 < -1.721 \)

• Interpret your results.

– Stereotype threat significantly reduced performance of those to whom it was applied.

Assumptions

1) The observations within each sample must be independent.
2) The two populations from which the samples are selected must be normal.
3) The two populations from which the samples are selected must have equal variances.
   - This is also known as homogeneity of variance, and there are two methods for testing that we have equal variances:
     a) informal method – simply compare sample variances
     b) Levene’s test – We’ll see this on the SPSS output
4) Random Assignment
   To make causal claims
5) Random Sampling
   To make generalizations to the target population

Which test?

• Each of the following studies requires a t test for one or more population means. Specify whether the appropriate t test is for one sample or two independent samples.

– College students are randomly assigned to undergo either behavioral therapy or Gestalt therapy. After 20 therapeutic sessions, each student earns a score on a mental health questionnaire.

– One hundred college freshmen are randomly assigned to sophomore roommates having either similar or dissimilar vocational goals. At the end of their freshman year, the GPAs of these 100 freshmen are to be analyzed on the basis of the previous distinction.

– According to the U.S. Department of Health and Human Services, the average 16-year-old male can do 23 push-ups. A physical education instructor finds that in his school district, 30 randomly selected 16-year-old males can do an average of 28 push-ups.

For next week

• Read Russ Lenth’s paper on effective sample-size determination.