Chapter 13
Hypothesis Tests:
Two Related Samples

PSY 295 - Oswald

Outline
• Related samples?
• Difference scores?
• An example
• t tests on difference scores
• Advantages and disadvantages
• Effect sizes
• Review questions

Related Samples
• The same participants give us data on two measures
  – Before and after treatment (e.g., drugs, counseling, etc.)
  – Aggressive responses before video and aggressive responses after
• With related samples, an individual’s score on one measure is likely to be with his/her score on another measures.

Related Samples
• Correlation between before and after scores
  – between before and after
  – Causes a change in the statistic we can use
• Sometimes called matched samples or
  – We will cover
  – Matched samples are where you pair people up based on similar characteristics (e.g., you look at differences between 2 training conditions, matching individuals’ training scores on equivalent age and experience).

Difference Scores
• Calculate difference between first and second score
  – e.g. Difference = Before – After
  or Difference = X₁ – X₂
• Base subsequent analysis on difference scores
  – Once you calculate the Difference, then don’t use the Before and After data

An Example
• Therapy for rape victims
  – Foa, Rothbaum, Riggs, & Murdock (1991)
• One group received Supportive Counseling
• Measured post-traumatic stress disorder (PTSD) symptoms before and after therapy
Therapy for PTSD

<table>
<thead>
<tr>
<th>Case#</th>
<th>Before</th>
<th>After</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

Mean | 23.89 | 15.67 | 8.22 |
St. Dev. | 4.20 | 4.24 | 3.60 |

Results

• The Supportive Counseling group reduces the number of symptoms.
  • Was this difference statistically significant?
  • Remember that Before and After scores are paired.
    - See raw data
    - $r = .64$

Results--cont.

• If no change, mean of differences should be 0.
  - So, test the obtained mean of difference scores against the null hypothesis $H_0: \mu_D = 0$.
  - Use same test as in Chapter 12.
• We don’t know $\sigma$, so use $s$ and solve for $t$.

$t$ test

$t = \frac{\bar{D} - \mu}{s_D / \sqrt{n}} = \frac{8.22}{3.6 / \sqrt{9}} = \frac{8.22}{1.2} = 6.85$

$df = n - 1 = 9 - 1 = 8$

$t$ test--cont.

• With 8 df, $t_{.025} = \pm 2.306$
• We calculated $t = 6.85$
• Since 6.85 > 2.306, reject $H_0$
• Conclude that the mean number of symptoms after therapy was greater than the mean number before therapy.
• Supportive counseling tends to work.

Advantages of Related Samples

• Eliminate subject-to-subject variability
• Control for extraneous variables
• Need only $n$ subjects
Disadvantages of Related Samples

- Effects
- Carry-over effects
- Change may just be a function of
- Subjects no longer
- Sometimes (e.g., different training programs, where the same knowledge is learned in each condition)

Effect Size Again

- We could simply report the average difference in the raw means.
  - Average Difference = 8.22
- But as with the difference in means for independent samples, the average difference in scores across repeated measures have to us—Is 8.22 large?
- As we did before could "scale" the difference by the size of to see whether the observed difference is a meaningful difference (or not)

Effect Size, cont.

\[ d = \frac{\mu_1 - \mu_2}{\sigma} = \frac{\mu_{\text{Before}} - \mu_{\text{After}}}{\sigma_{\text{Before}}} \]

\[ = \frac{23.89 - 15.67}{4.20} = \frac{8.22}{4.20} = 1.96 \]

SPSS

- Next slide shows SPSS Printout
  - Similar printout from other software
  - Results match ours

Paired Samples Statistics

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>POST</td>
<td>15.6667</td>
<td>4.2426</td>
<td>1.4142</td>
</tr>
<tr>
<td>PRE</td>
<td>23.8909</td>
<td>4.1966</td>
<td>1.3989</td>
</tr>
</tbody>
</table>

Paired Samples Correlations

<table>
<thead>
<tr>
<th>N</th>
<th>Correlation</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>.637</td>
<td>.065</td>
</tr>
</tbody>
</table>

Paired Samples Test

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std. Error Mean</td>
</tr>
<tr>
<td>POST - PRE</td>
<td>8.2222</td>
</tr>
</tbody>
</table>
Review Questions
(These are NOT the only questions to study for)

• Why do we say that the two sets of measures are not independent?
• What are other names for “related samples?”
• How do we calculate difference scores?
  – What happens if we subtract before from after instead of after from before?

Review Questions--cont.
(These are NOT the only questions to study for)

• Why do we usually test $H_0: \mu_D = 0$?
• Why do we have 8 $df$ in our sample when we have 18 observations?
• What are the advantages and disadvantages of related samples?
• What do effect sizes tell you in this case?
• How would you calculate the confidence interval that SPSS produced?