This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike License. Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.
Section D

Two Sample t-test, Two Choices
The “traditional” t-test assumes equal variances in the two groups
  - This can be formally tested with another hypothesis test!
  - But why not just compare observed values of $s_1$ to $s_2$?

There is a slight modification to allow for unequal variances—this modification adjusts the degrees of freedom for the test, using slightly different SE computation (the formula I give you)

If you want to be truly “safe” (desert island choice of t-test)
  - More conservative to use test that allows for unequal variance

Makes little to no difference in large sample
FYI: Equal Variances Assumption

- Actually, the following occurs:
  - If underlying population level standard deviations are equal:
    ▶ Both approaches give valid confidence intervals but intervals by approach assuming unequal standard deviations slightly wider (and p-values slightly larger)
  - If underlying population level standard deviations are not equal:
    ▶ The approach assuming equal variances does not give valid confidence intervals and can severely under-cover the goal of 95%
Unequal SD Approach: Diet Type/ Weight Loss Example

- **Command syntax:**
  
  \[
  \texttt{ttesti } n_1 \bar{x}_1 \, s_1 \, n_2 \bar{x}_2 \, s_2, \text{ unequal}
  \]

  \[
  . \texttt{ttesti 64 -5.7 8.6 68 -1.8 3.9, unequal}
  \]

  **Two-sample t test with unequal variances**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>64</td>
<td>-5.7</td>
<td>1.075</td>
<td>8.6</td>
<td>-7.848216 to -3.551784</td>
</tr>
<tr>
<td>y</td>
<td>68</td>
<td>-1.8</td>
<td>.4729445</td>
<td>3.9</td>
<td>-2.744001 to -1.8559989</td>
</tr>
<tr>
<td>combined</td>
<td>132</td>
<td>-3.690909</td>
<td>.5978226</td>
<td>6.868458</td>
<td>-4.873545 to -2.508273</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-3.9</td>
<td>1.174437</td>
<td></td>
<td>-6.234436 to -1.565564</td>
</tr>
</tbody>
</table>

  \[
  \text{diff} = \text{mean}(x) - \text{mean}(y) \quad t = -3.3207
  \]

  \[
  \text{Satterthwaite's degrees of freedom} = 86.6941
  \]

  \[
  \text{Ha: } \text{diff} < 0 \quad \text{Pr}(T < t) = 0.0007 \quad \text{Pr}(|T| > |t|) = 0.0013 \quad \text{Pr}(T > t) = 0.9993
  \]

  \[
  \text{Ha: } \text{diff} \neq 0
  \]

  \[
  \text{Ha: } \text{diff} > 0
  \]
**Equal SD Approach: Diet Type/ Weight Loss Example**

- **Command syntax:**
  - `ttesti n_1 \bar{x}_1 s_1 n_2 \bar{x}_2 s_2`

  . `ttesti 64 -5.7 8.6 68 -1.8 3.9`

  **Two-sample t test with equal variances**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[  95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>64</td>
<td>-5.7</td>
<td>1.075</td>
<td>8.6</td>
<td>-7.848216 -3.551784</td>
</tr>
<tr>
<td>y</td>
<td>68</td>
<td>-1.8</td>
<td>.4729445</td>
<td>3.9</td>
<td>-2.744001 - .8559989</td>
</tr>
<tr>
<td>combined</td>
<td>132</td>
<td>-3.690909</td>
<td>.5978226</td>
<td>6.868458</td>
<td>-4.873545 -2.508273</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-3.9</td>
<td>1.151038</td>
<td></td>
<td>-6.177191 -1.622809</td>
</tr>
</tbody>
</table>

  `diff = \text{mean}(x) - \text{mean}(y)`

  Ho: `diff = 0`

  $t = -3.3882$

  degrees of freedom = 130

  Ha: `diff < 0`
  Ha: `diff != 0`
  Ha: `diff > 0`

  Pr(T < t) = 0.0005  Pr(|T| > |t|) = 0.0009  Pr(T > t) = 0.9995
Unequal SD Approach: LDL/ Treatment Example

- **Command syntax:**
  
  \texttt{ttesti n_1 \bar{x}_1 s_1 n_2 \bar{x}_2 s_2, unequal}

  \texttt{. ttesti 11 -1.41 .55 12 -.32 .65, unequal}

Two-sample t test with unequal variances

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[ 95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>11</td>
<td>-1.41</td>
<td>0.1658312</td>
<td>0.55</td>
<td>-1.779495 -1.040505</td>
</tr>
<tr>
<td>y</td>
<td>12</td>
<td>-0.32</td>
<td>0.1876388</td>
<td>0.65</td>
<td>-0.7329903 0.0929903</td>
</tr>
<tr>
<td>combined</td>
<td>23</td>
<td>-0.8413043</td>
<td>0.1692296</td>
<td>0.8115967</td>
<td>-1.192265 -0.4903436</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-1.09</td>
<td>0.2504163</td>
<td>-1.61095</td>
<td>-0.5690505</td>
</tr>
</tbody>
</table>

\texttt{diff = mean(x) - mean(y)}

\texttt{t = -4.3528}

\texttt{Ho: diff = 0}

\texttt{Satterthwaite's degrees of freedom = 20.8813}

\texttt{Ha: diff < 0}

\texttt{Pr(T < t) = 0.0001}

\texttt{Ha: diff != 0}

\texttt{Pr(|T| > |t|) = 0.0003}

\texttt{Ha: diff > 0}

\texttt{Pr(T > t) = 0.9999}
Equal SD Approach: LDL/Treatment Example

- **Command syntax:**
  
  \[ \texttt{ttesti } n_1 \overline{x}_1 s_1 \quad n_2 \overline{x}_2 s_2, \text{ unequal} \]

  . ttesti 11 -1.41 .55 12 - .32 .65

  Two-sample t test with equal variances

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>11</td>
<td>-1.41</td>
<td>.1658312</td>
<td>.55</td>
<td>-1.779495 -1.040505</td>
</tr>
<tr>
<td>y</td>
<td>12</td>
<td>-.32</td>
<td>.1876388</td>
<td>.65</td>
<td>-.7329903 .0929903</td>
</tr>
<tr>
<td>combined</td>
<td>23</td>
<td>-.8413043</td>
<td>.1692296</td>
<td>.8115967</td>
<td>-1.192265 - .4903436</td>
</tr>
</tbody>
</table>

  |     | diff | -1.09 | .2523107 | -1.614709 | - .5652911 |

  \[ \text{diff} = \text{mean}(x) - \text{mean}(y) \]
  
  \[ t = -4.3201\]

  Ho: diff = 0
degrees of freedom = 21

  Ha: diff < 0
  Pr(T < t) = 0.0002
  Ha: diff ≠ 0
  Pr(|T| > |t|) = 0.0003
  Ha: diff > 0
  Pr(T > t) = 0.9998