<table>
<thead>
<tr>
<th>Parameter</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference between means, independent samples, unpooled $\mu_1 - \mu_2$</td>
<td>$(\bar{x}_1 - \bar{x}<em>2) \pm z</em>{a/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ or $(\bar{x}_1 - \bar{x}<em>2) \pm t</em>{a/2,df} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, $df = \min(n_1 - 1, n_2 - 1)$</td>
</tr>
<tr>
<td>Difference between means, independent samples, pooled variance $\mu_1 - \mu_2$</td>
<td>$(\bar{x}_1 - \bar{x}<em>2) \pm z</em>{a/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ or $(\bar{x}_1 - \bar{x}<em>2) \pm t</em>{a/2,df} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, $df = n_1 + n_2 - 2$</td>
</tr>
</tbody>
</table>

Often reasonable to assume the two populations have equal population standard deviations, or equivalently, equal population variances.

Estimate of this variance based on the combined or “pooled” data is called the pooled variance. The square root of the pooled variance is called the pooled standard deviation:

Pooled standard deviation $s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$
Example: Male and Female Sleep Times

Question: How much difference is there between how long female and male students slept the previous night?

Data: The 83 female and 65 male responses from students in an intro stat class.

Task: Make a 95% CI for the difference between the two population mean sleep hours for females versus males.

Female

\[ n_F = 83; \quad \bar{x}_F = 7.02; \quad s_F = 1.75 \]

Male

\[ n_M = 65; \quad \bar{x}_M = 6.55; \quad s_M = 1.68 \]

Two sample standard deviations are very similar so we will assume equal population variances.

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

\[ = \sqrt{\frac{(83-1)(1.75)^2 + (65-1)(1.68)^2}{83 + 65 - 2}} \]

\[ = \sqrt{2.957} = 1.72 \]

Weighted average of the two sample variances

Pooled s.e.(\(\bar{x}_1 - \bar{x}_2\)) = \(s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\)

\[ = 1.72\sqrt{\frac{1}{83} + \frac{1}{65}} = 0.285 \]

\((\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2, df}\cdot s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad df = n_1 + n_2 - 2\)
\[(7.02 - 6.55) \pm t_{0.025,146} \times 0.285\]
\[\Rightarrow 0.47 \pm 1.98 \times 0.285\]
\[\Rightarrow 0.47 \pm 0.564\]
\[\Rightarrow (-0.094, 1.034)\]

95% confidence interval contains 0 so cannot rule out that the population means may be equal.

**Pooled or Unpooled Variances**

- **If sample sizes are equal**, the pooled and unpooled standard errors are equal.

  If sample standard deviations are similar, assumption of equal population variance may be reasonable and the pooled procedure could be used.

- **If sample sizes are very different**, pooled test can be quite misleading unless sample standard deviations are similar. If the smaller standard deviation accompanies the larger sample size, not recommended to use the pooled procedure.

- **If sample sizes are very different**, the standard deviations are similar, and the larger sample size produced the larger standard deviation, the pooled procedure is acceptable because it will be conservative.