Chapter 9: Testing the Difference Between Two Means, Two Proportions, and Two Variances

Diana Pell

Section 9.2: Testing the Difference Between Two Means of Independent Samples

Note: Samples are independent samples when they are not related.

Assumptions for the $t$ Test for Two Independent Means When $\sigma_1$ and $\sigma_2$ Are Unknown

1. The samples are random samples.
2. The sample data are independent of one another.
3. When the sample sizes are less than 30, the populations must be normally or approximately normally distributed.

**Formula for the $t$ Test for Testing the Difference Between Two Means, Independent Samples**

\[
t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}
\]

where the degrees of freedom are equal to the smaller of $n_1 - 1$ or $n_2 - 1$. 

Assumptions for the $t$ Test for Two Independent Means When $\sigma_1$ and $\sigma_2$ Are Unknown

1. The samples are random samples.
2. The sample data are independent of one another.
3. When the sample sizes are less than 30, the populations must be normally or approximately normally distributed.
Exercise 1. A researcher wishes to see if the average weights of newborn male infants are different from the average weights of newborn female infants. She selects a random sample of 10 male infants and finds the mean weight is 7 pounds 11 ounces and the standard deviation of the sample is 8 ounces. She selects a random sample of 8 female infants and finds that the mean weight is 7 pounds 4 ounces and the standard deviation of the sample is 5 ounces. Can it be concluded at $\alpha = 0.05$ that the mean weight of the males is different from the mean weight of the females? Assume that the variables are normally distributed.
Exercise 2. Find the 95% confidence interval for the data in Exercise 1.

Confidence Intervals for the Difference of Two Means: Independent Samples

Variances assumed to be unequal:

\[
(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

d.f. = smaller value of \(n_1 - 1\) or \(n_2 - 1\)

Section 9.3: Testing the Difference Between Two Means: Dependent Samples

Note: Samples are considered to be dependent samples when the subjects are paired or matched in some way.
When the samples are dependent, a special t test for dependent means is used. This test employs the difference in values of the matched pairs. The hypotheses are as follows:

<table>
<thead>
<tr>
<th>Two-tailed</th>
<th>Left-tailed</th>
<th>Right-tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0: \mu_D = 0$</td>
<td>$H_0: \mu_D = 0$</td>
<td>$H_0: \mu_D = 0$</td>
</tr>
<tr>
<td>$H_1: \mu_D \neq 0$</td>
<td>$H_1: \mu_D &lt; 0$</td>
<td>$H_1: \mu_D &gt; 0$</td>
</tr>
</tbody>
</table>

Assumptions for the t Test for Two Means When the Samples Are Dependent
1. The sample or samples are random.
2. The sample data are dependent.
3. When the sample size or sample sizes are less than 30, the population or populations must be normally or approximately normally distributed.

Formulas for the t Test for Dependent Samples

\[ t = \frac{\bar{D} - \mu_D}{s_D / \sqrt{n}} \]

with d.f. = $n - 1$ and where

\[ \bar{D} = \frac{\sum D}{n} \text{ and } s_D = \sqrt{\frac{n\sum D^2 - (\sum D)^2}{n(n-1)}} \]
Exercise 3. A random sample of nine local banks shows their deposits (in billions of dollars) 3 years ago and their deposits (in billions of dollars) today. At $\alpha = 0.05$, can it be concluded that the average in deposits for the banks is greater today than it was 3 years ago? Use $\alpha = 0.05$. Assume the variable is normally distributed.

<table>
<thead>
<tr>
<th>Bank</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 years ago</td>
<td>11.42</td>
<td>8.41</td>
<td>3.98</td>
<td>7.37</td>
<td>2.28</td>
<td>1.10</td>
<td>1.00</td>
<td>0.9</td>
<td>1.35</td>
</tr>
<tr>
<td>Today</td>
<td>16.69</td>
<td>9.44</td>
<td>6.53</td>
<td>5.58</td>
<td>2.92</td>
<td>1.88</td>
<td>1.73</td>
<td>1.5</td>
<td>1.22</td>
</tr>
</tbody>
</table>
**Exercise 4.** A dietitian wishes to see if a person’s cholesterol level will change if the diet is supplemented by a certain mineral. Six randomly selected subjects were pretested, and then they took the mineral supplement for a 6-week period. The results are shown in the table. (Cholesterol level is measured in milligrams per deciliter.) Can it be concluded that the cholesterol level has been changed at $\alpha = 0.10$? Assume the variable is approximately normally distributed.

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before ($X_1$)</td>
<td>210</td>
<td>235</td>
<td>208</td>
<td>190</td>
<td>172</td>
<td>244</td>
</tr>
<tr>
<td>After ($X_2$)</td>
<td>190</td>
<td>170</td>
<td>210</td>
<td>188</td>
<td>173</td>
<td>229</td>
</tr>
</tbody>
</table>
Exercise 5. Find the 90% confidence interval for the data in Exercise 4.