Chapter 6. Inference about $\mu_1 - \mu_2$

1. Large Populations ($n_1 \geq 30$, $n_2 \geq 30$)
   Two Independent Samples

   **EXAMPLE 6.1**

   A survey of credit card holders revealed that Americans carried an average credit card balance of $3900 in 1995 and $3300 in 1994 (U.S. News & World Report, January 1, 1996). Suppose that these averages are based on random samples of 400 credit card holders in 1995 and 450 credit card holders in 1994 and that the population standard deviations of the balances were $880 in 1995 and $810 in 1994.

   (a) What is the point estimate of $\mu_1 - \mu_2$?
   (b) Construct a 95% confidence interval for the difference between the mean credit card balances for all credit card holders in 1995 and 1994.

   **SOLUTION** Refer to all credit card holders in 1995 as population 1 and all credit card holders in 1994 as population 2. The respective samples, then, are samples 1 and 2. Let $\mu_1$ and $\mu_2$ be the mean credit card balances for populations 1 and 2, and let $\bar{x}_1$ and $\bar{x}_2$ be the means of the respective samples. From the given information,

   For 1995: $n_1 = 400$, $\bar{x}_1 = $3900, $\sigma_1 = $880
   For 1994: $n_2 = 450$, $\bar{x}_2 = $3300, $\sigma_2 = $810

   (a) The point estimate of $\mu_1 - \mu_2$ is given by the value of $\bar{x}_1 - \bar{x}_2$. Thus,

   Point estimate of $\mu_1 - \mu_2 = $3900 - $3300 = $600

   (b) The confidence level is $1 - \alpha = .95$.

   First, we calculate the standard deviation of $\bar{x}_1 - \bar{x}_2$ as follows.

   $$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(880)^2}{400} + \frac{(810)^2}{450}} = $58.25804665$$

   Next, we find the $z$ value for the 95% confidence level. From the normal distribution table, this value of $z$ is 1.96.
Finally, substituting all the values in the confidence interval formula, we obtain the 95% confidence interval for $\mu_1 - \mu_2$ as

\[
(\bar{x}_1 - \bar{x}_2) \pm z\sigma_{\bar{x}_1-x_2} = (3900 - 3300) \pm 1.96 \cdot 58.25804665
\]

\[
= 600 \pm 114.19 = \$485.81 \text{ to } \$714.19
\]

Thus, with 95% confidence we can state that the difference in the mean credit card balances for all credit card holders in 1995 and 1994 was between $485.81 and $714.19.

EXAMPLE 6.2

In 1996, *Money* magazine conducted an experiment in which one-dollar coins and quarters were left on busy sidewalks in six major U.S. cities (*Money*, October 1996). An observer then recorded the length of time each coin remained on the sidewalk before being picked up by a pedestrian. According to the results of the experiment, the mean length of time one-dollar coins remained on the sidewalks before being picked up by pedestrians was 6.50 minutes and the mean length of time quarters stayed on the sidewalks was 5.75 minutes. Assume that these means are based on samples of 50 one-dollar coins and 45 quarters and that the two sample standard deviations are 1.75 minutes and 1.20 minutes, respectively. Find a 99% confidence interval for the difference between the corresponding population means.

EXAMPLE 6.3

Refer to Example 6.2 about the mean length of times one-dollar coins and quarters remained on the sidewalks before being picked up by pedestrians. Test at the 2.5% significance level if the mean length of time all one-dollar coins will remain on the sidewalks before being picked up is higher than the mean length of time all quarters will stay on the sidewalks.

SOLUTION

From the information given in Example 6.2,

For $1$ coins: $n_1 = 50, \quad \bar{x}_1 = 6.50, \quad \sigma_1 = 1.75 \text{ minutes}$

For quarters: $n_2 = 45, \quad \bar{x}_2 = 5.75, \quad \sigma_2 = 1.20 \text{ minutes}$

Let $\mu_1$ and $\mu_2$ be the mean lengths of time all one-dollar coins and all quarters will remain on the sidewalks before being picked up, respectively.
EXAMPLE 6.4

Refer to Example 6.1 about the mean credit card balances maintained by credit card holders in America in 1995 and 1994. Test at the 1% significance level if the mean credit card balances for all credit card holders in America in 1995 and 1994 were different.

SOLUTION From the information given in Example 6.1,

For 1995:
\[ n_1 = 400, \quad \bar{x}_1 = \$3900, \quad \sigma_1 = \$880 \]

For 1994:
\[ n_2 = 450, \quad \bar{x}_2 = \$3300, \quad \sigma_2 = \$810 \]

Let \( \mu_1 \) and \( \mu_2 \) be the mean credit card balances for all credit card holders in America in 1995 and 1994, respectively.

Step 1. State the null and alternative hypotheses

We are to test if the two population means are different. The two possibilities are

(i) The mean credit card balances for all credit card holders in America in 1995 and 1994 are not different. In other words, \( \mu_1 = \mu_2 \), which can be written as \( \mu_1 - \mu_2 = 0 \).

(ii) The mean credit card balances for all credit card holders in America in 1995 and 1994 are different. That is, \( \mu_1 \neq \mu_2 \), which can be written as \( \mu_1 - \mu_2 \neq 0 \).

Considering these two possibilities, the null and alternative hypotheses are

\[ H_0 : \mu_1 - \mu_2 = 0 \quad \text{(the two population means are not different)} \]
\[ H_1 : \mu_1 - \mu_2 \neq 0 \quad \text{(the two population means are different)} \]

Step 2. Select the distribution to use

Because \( n_1 > 30 \) and \( n_2 > 30 \), both sample sizes are large. Therefore, the sampling distribution of \( \bar{x}_1 - \bar{x}_2 \) is approximately normal, and we use the normal distribution to make the hypothesis test.

Step 3. Determine the rejection and non-rejection regions

The significance level is given to be .01. The \( \neq \) sign in the alternative hypothesis indicates that the test is two-tailed. The area in each tail of the normal distribution curve is \( \alpha/2 = .01/2 = .005 \). The critical values of \( z \) for .005 area in each tail of the normal distribution curve are (approximately) 2.58 and -2.58. These values are shown in the next Figure.
Step 4. Calculate the value of the test statistic

The value of the test statistic \( z \) for \( \bar{x}_1 - \bar{x}_2 \) is computed as follows.

\[
\sigma_{x_1-x_2} = \sqrt{\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}} = \sqrt{\frac{(880)^2}{400} + \frac{(810)^2}{450}} = 58.25804665
\]

\[
z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{x_1-x_2}} = \frac{(3900 - 3300) - 0}{58.25804665} = 10.30
\]

Step 5. Make a decision

Because the value of the test statistic \( z = 10.30 \) falls in the rejection region, we reject the null hypothesis \( H_0 \). Therefore, we conclude that the mean credit card balances for all credit card holders in America in 1995 and 1994 were different. Note that we cannot say for sure that the two population means are different. All we can say is that the evidence from the two samples is very strong that the corresponding population means are different.

2. Large Populations, \( \sigma^2_1 = \sigma^2_2 \) (unknown)

Two Independent Samples

Confidence Interval for \( \mu_1 - \mu_2 \), Independent Samples

\[
(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2} \hat{s}_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}
\]

where

\[
\hat{s}_p = \sqrt{\frac{(n_1 - 1)s^2_1 + (n_2 - 1)s^2_2}{n_1 + n_2 - 2}}
\]

and

\[
df = n_1 + n_2 - 2.
\]
Table 1: Potency Reading for Two Samples

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.2</td>
<td>10.6</td>
</tr>
<tr>
<td>10.5</td>
<td>10.7</td>
</tr>
<tr>
<td>10.3</td>
<td>10.2</td>
</tr>
<tr>
<td>10.8</td>
<td>10.0</td>
</tr>
<tr>
<td>9.8</td>
<td>10.6</td>
</tr>
</tbody>
</table>

$s_p^2$, a weighted average

The quantity $s_p$ in the confidence interval is an estimate of the standard deviation $\sigma$ for the two populations and is formed by combining (pooling) information from the two samples. In fact, $s_p^2$ is a weighted average of the sample variances $s_1^2$ and $s_2^2$. For the special case where the sample sizes are the same ($n_1 = n_2$), the formula for $s_p^2$ reduces to $s_p^2 = (s_1^2 + s_2^2)/2$, the mean of the two sample variances. The degrees of freedom for the confidence interval are a combination of the degrees of freedom for the two samples; that is, $df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$. Recall that we are assuming that the two populations from which we draw the samples have normal distributions with a common variance $\sigma^2$. If the confidence interval presented were valid only when these assumptions were met exactly, the estimation procedure would be of limited use. Fortunately, the confidence coefficient remains relatively stable if both distributions are mound-shaped and the sample sizes are approximately equal. More discussion about these assumptions is presented at the end of this section.

EXAMPLE 6.5

Company officials were concerned about the length of time a particular drug product retained its potency. A random sample, sample 1, of $n_1 = 10$ bottles of the product was drawn from the production line and analyzed for potency. A second sample, sample 2, of $n_2 = 10$ bottles was obtained and stored in a regulated environment for a period of one year.

The readings obtained from each sample are given in Table 1.

Suppose we let $\mu_1$ denote the mean potency for all bottles that might be sampled coming off the production line and $\mu_2$ denote the mean potency for all bottles that may be retained for a period of one year. Estimate $\mu_1 - \mu_2$ by using a 95% confidence interval.
\[
\sum_{j} y_{1j} = 103.7 \quad \sum_{j} y_{2j} = 98.3 \\
\sum_{j} y_{1j}^2 = 1076.31 \quad \sum_{j} y_{2j}^2 = 966.81
\]

Then
\[
\bar{y}_1 = \frac{103.7}{10} = 10.37 \quad \bar{y}_2 = \frac{98.3}{10} = 9.83
\]

\[
s_1^2 = \frac{1}{9} \left[ 1076.31 - \frac{103.7^2}{10} \right] = .105 \quad s_2^2 = \frac{1}{9} \left[ 966.81 - \frac{98.3^2}{10} \right] = .058.
\]

The estimate of the common standard deviation \(\sigma\) is

\[
s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{9(.105) + 9(.058)}{18}} = .285.
\]

The \(t\)-value based on \(df = n_1 + n_2 - 2 = 18\) and \(\alpha = .025\) is 2.101. A 95\% confidence interval for the difference in mean potencies is

\[
(10.37 - 9.83) \pm 2.101(.285)\sqrt{1/10 + 1/10} \text{ or } .54 \pm .268.
\]

We estimate that the difference in mean potencies for the bottles from the production line and those stored for 1 year, \(\mu_1 - \mu_2\), lies in the interval .272 to .808.

**EXAMPLE 6.6**

A study was conducted to determine whether persons in suburban district 1 have a different mean income from those in district 2. A random sample of 20 homeowners was taken in district 1. Although 20 homeowners were to be interviewed in district 2 also, 1 person refused to provide the information requested, even though the researcher promised to keep the interview confidential. So only 19 observations were obtained from district 2. The data, recorded in thousands of dollars, produced sample means and variances as shown in Table 2. Use these data to construct a 95\% confidence interval for \((\mu_1 - \mu_2)\).

**SOLUTION** Histograms plotted for the two samples suggest that the two populations are mound-shaped (near normal). Also, the sample variances are very similar. The difference in the sample means is

\[
\bar{y}_1 - \bar{y}_2 = 18.27 - 16.78 = 1.49.
\]
<table>
<thead>
<tr>
<th>Sample size</th>
<th>District 1</th>
<th>District 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Sample mean</td>
<td>18.27</td>
<td>16.78</td>
</tr>
<tr>
<td>Sample variance</td>
<td>8.74</td>
<td>6.58</td>
</tr>
</tbody>
</table>

Table 2:

The estimate of the common standard deviation $\sigma$ is

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

$$= \sqrt{\frac{19(8.74) + 18(6.58)}{20 + 19 - 2}} = 2.77.$$

The $t$-value for $a = \alpha/2 = .025$ and $df = 20 + 19 - 2 = 37$ is not listed in Table 4 of the Appendix, but taking the labeled value for the nearest $df$ ($df = 40$), we have $t = 2.021$. A 95% confidence interval for the difference in mean incomes for the two districts is of the form

$$\bar{y}_1 - \bar{y}_2 \pm \frac{t_{\alpha/2}s_p}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Substituting into the formula we obtain

$$1.49 \pm 2.021(2.77)\sqrt{\frac{1}{20} + \frac{1}{19}}$$

or

$$1.49 \pm 1.79.$$  

Thus, we estimate the difference in mean incomes to lie somewhere in the interval from $-0.30$ to $3.28$. If we multiply these limits by $\$1,000$, the confidence interval for the difference in mean incomes is $-\$300$ to $\$3,280$. Since this interval includes both positive and negative values for $\mu_1 - \mu_2$, we are unable to determine whether the mean income for district 1 is larger or smaller than the mean income for district 2.

We can also test a hypothesis about the difference between two population means. As with any test procedure, we begin by specifying a research hypothesis for the difference in population means. Thus, we might, for example, specify that the difference $\mu_1 - \mu_2$ is greater than some value $D_0$. (Note: $D_0$ will often be 0.) The entire test procedure is summarized here.

\[\text{A statistical Test for } \mu_1 - \mu_2, \text{ Independent Samples}\]

$H_0$: $\mu_1 - \mu_2 = D_0$ ($D_0$ is specified)
\[ H_0: \]

1. \( \mu_1 - \mu_2 > D_0 \)
2. \( \mu_1 - \mu_2 < D_0 \)
3. \( \mu_1 - \mu_2 \neq D_0 \)

T.S. :

\[ t = \frac{\bar{y}_1 - \bar{y}_2 - D_0}{s_p \sqrt{1/n_1 + 1/n_2}} \]

R.R. : For a Type I error \( \alpha \) and \( df = n_1 + n_2 - 2 \)

1. reject \( H_0 \) if \( t > t_\alpha \)
2. reject \( H_0 \) if \( t < -t_\alpha \)
3. reject \( H_0 \) if \( |t| > t_{\alpha/2} \)

**EXAMPLE 6.7**

An experiment was conducted to compare the mean number of tapeworms in the stomachs of sheep that had been treated for worms against the mean number in those that were untreated. A sample of 14 worm-infected lambs was randomly divided into 2 groups. Seven were injected with the drug and the remainder were left untreated. After a 6-month period, the lambs were slaughtered and the following worm counts were recorded:

<table>
<thead>
<tr>
<th>Drug-treated sheep</th>
<th>18</th>
<th>43</th>
<th>28</th>
<th>50</th>
<th>16</th>
<th>32</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated sheep</td>
<td>40</td>
<td>54</td>
<td>26</td>
<td>63</td>
<td>21</td>
<td>37</td>
<td>39</td>
</tr>
</tbody>
</table>

a. Test a hypothesis that there is no difference in the mean number of worms between treated and untreated lambs. Assume that the drug cannot increase the number of worms and, hence, use the alternative hypothesis that the mean for treated lambs is less than the mean for untreated lambs. Use \( \alpha = .05 \).

b. Indicate the level of significance for this test.

**SOLUTION**

a. The calculations for the samples of treated and untreated sheep are summarized next.
Under the assumption of equal population variances, the sample variances are combined to form an estimate of the common population standard deviation \( \sigma \). This assumption appears reasonable based on the sample variances.

\[
s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{6(198.62) + 6(215.33)}{12}} = 14.39.
\]

The test procedure for the research hypothesis that the treated sheep will have a mean infestation level (\( \mu_1 \)) less than the mean level (\( \mu_2 \)) for untreated sheep is as follows:

- \( H_0: \mu_1 - \mu_2 = 0 \) (that is, no difference in the mean infestation levels)
- \( H_a: \mu_1 - \mu_2 < 0 \)

T.S. :

\[
t = \frac{\bar{y}_1 - \bar{y}_2}{s_p \sqrt{1/n_1 + 1/n_2}} = \frac{28.57 - 40}{14.39 \sqrt{1/7 + 1/7}} = -1.49
\]

R.R. : For \( \alpha = .05 \), the critical \( t \)-value for a one-tailed test with \( df = n_1 + n_2 - 2 = 12 \) can be obtained from Table 4 in the Appendix, using \( a = .05 \). We will reject \( H_0 \) if \( t < -1.782 \).

Conclusion: Since the observed value of \( t \), -1.49, does not fall in the rejection region, we have insufficient evidence to reject the hypothesis that there is no difference in the mean number of worms in treated and untreated lambs.

b. Using Table 4 in the Appendix with \( t = -1.49 \) and \( df = 12 \), we see the level of significance for this test is in the range \( .05 < p < .10 \).
3. Normal Populations. Dependent Samples (Paired data)

Test Statistic for Two Dependent Samples

\[ t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} \]

where degrees of freedom = \( n - 1 \).

If the number of pairs of data is large (\( n > 30 \)), the number of degrees of freedom will be at least 30, so critical values will be \( z \) scores (Table A-2) instead of \( t \) scores (Table A-3).

EXAMPLE 6.7

Using a reaction timer similar to the one described in the Cooperative Group Activities of Chapter 5, subjects are tested for reaction times with their left and right hands. (Only right-handed subjects were used.) The results (in thousandths of a second) are given in the accompanying table. Use a 0.05 significance level to test the claim that there is a difference between the mean of the right- and left-hand reaction times. If an engineer is designing a fighter-jet cockpit and must locate the ejection-seat activator to be accessible to either the right or the left hand, does it make a difference which hand she chooses?

<table>
<thead>
<tr>
<th>Subject</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right</td>
<td>191</td>
<td>97</td>
<td>116</td>
<td>165</td>
<td>116</td>
<td>129</td>
<td>171</td>
<td>155</td>
<td>112</td>
<td>102</td>
<td>188</td>
<td>158</td>
<td>121</td>
<td>133</td>
</tr>
<tr>
<td>Left</td>
<td>224</td>
<td>171</td>
<td>191</td>
<td>207</td>
<td>196</td>
<td>165</td>
<td>177</td>
<td>165</td>
<td>140</td>
<td>188</td>
<td>155</td>
<td>219</td>
<td>177</td>
<td>174</td>
</tr>
</tbody>
</table>

SOLUTION Using the traditional method of hypothesis testing, we will test the claim that there is a difference between the right- and left-hand reaction times. Because we are dealing with paired data, begin by finding the differences \( d = \text{right} - \text{left} \).

Step 1: If there is a difference, we expect the mean of the \( d \) values to be different from 0. This is expressed in symbolic form as \( \mu_d \neq 0 \).

Step 2: If the original claim is not true, we have \( \mu_d = 0 \).

Step 3: The null hypothesis must contain equality, so we have

\[ H_0 : \mu_d = 0 \quad H_1 : \mu_d \neq 0 \quad (\text{original claim}) \]
Step 4: The significance level is \( \alpha = 0.05 \).

Step 5: Because we are testing a claim about the means of paired dependent data, we use the Student \( t \) distribution.

Step 6: Before finding the value of the test statistic, we must first find the values of \( d \) and \( s_d \). When we evaluate the difference \( d \) for each subject, we find these differences \( d = \text{right} - \text{left} \):

\[-33, -74, -75, -42, -80, -36, -6, -10, -28, -86, 33, -61, -56, -41\]

\[
\bar{d} = \frac{\sum d}{n} = \frac{-595}{14} = -42.5
\]

\[
s_d = \sqrt{n\left(\frac{\sum d^2}{n} - \frac{(\sum d)^2}{n(n-1)}\right)} = \sqrt{\frac{14(39,593) - (-595)^2}{14(14-1)}} = 33.2
\]

With these statistics and the assumption that \( \mu_d = 0 \), we can now find the value of the test statistic.

\[
t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-42.5 - 0}{33.2 / \sqrt{14}} = -4.790.
\]

The critical values of \( t = -2.160 \) and \( t = 2.160 \) are found from Table A-3; use the column for 0.05 (two tails), and use the row with degrees of freedom of \( n - 1 = 13 \). The above figure shows the test statistic, critical values, and critical region.

Step 7: Because the test statistic does fall in the critical region, we reject the null hypothesis of \( \mu_d = 0 \).

Step 8: There is sufficient evidence to support the claim of a difference between the right- and left-hand reaction times. Because there does appear to be
such a difference, an engineer designing a fighter-jet cockpit should locate the ejection-seat activator so that it is readily accessible to the faster hand, which appears to be the right hand with seemingly lower reaction times. (We could require special training for left-handed pilots if a similar test of left-handed pilots shows that their dominant hand is faster.)

Confidence Intervals

The confidence interval estimate of the mean difference $\mu_d$ is as follows:

$$\bar{d} - E < \mu_d < \bar{d} + E$$

where $E = t_{\alpha/2} \frac{s_d}{\sqrt{n}}$ and degrees of freedom = $n - 1$.

EXAMPLE 6.8

Use the sample data from the preceding example to construct a 95% confidence interval estimate of $\mu_d$.

SOLUTION Using the values of $\bar{d} = -42.5$, $s_d = 33.2$, $n = 14$, and $t_{\alpha/2} = 2.160$, we first find the value of the margin of error $E$.

$$E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 2.160 \frac{33.2}{\sqrt{14}} = 19.2$$

The confidence interval can now be found.

$$\bar{d} - E < \mu_d < \bar{d} + E$$

$$-42.5 - 19.2 < \mu_d < -42.5 + 19.2$$

$$-61.7 < \mu_d < -23.3$$

Crest and Dependent Samples

In the late 1950s, Procter & Gamble introduced Crest toothpaste as the first such product with fluoride. To test the effectiveness of Crest in reducing cavities, researchers conducted experiments with several sets of twins. One of the twins in each set was given Crest with fluoride, while the other twin continued to use ordinary toothpaste without fluoride. It was believed that each pair of twins would have similar eating, brushing, and genetic characteristics. Results showed that the twins who used Crest had significantly fewer cavities than those who did not. This use of twins as dependent samples allowed the researchers to control many of the different variables affecting cavities.
4. Choosing Sample Sizes for Inferences about $\mu_1 - \mu_2$

Sections 5.3 and 5.5 were devoted to sample-size calculations to obtain a confidence interval about $\mu$ with a fixed width and specified degree of confidence or to conduct a statistical test concerning $\mu$ with predefined levels for $\alpha$ and $\beta$. Similar calculations can be made for inferences about $\mu_1 - \mu_2$ of width $2E$ based on independent samples is possible by solving the following expression for $n$. We will assume that both samples are of the same size.

$$z_{\alpha/2} \sigma \sqrt{\frac{1}{n} + \frac{1}{n}} = E$$

Note that in this formula $\sigma$ is the common population standard deviation and that we have assumed equal sample sizes.

### Sample Sizes for a 100(1 - $\alpha$)% C. I. for $\mu_1 - \mu_2$ of the Form $\bar{y}_1 - \bar{y}_2 \pm E$, Ind. Samples

$$n = \frac{2z^2 \sigma^2}{E^2}$$

**Note:** If $\sigma$ is unknown, substitute an estimated value to get an approximate sample size.

The sample sizes obtained using this formula are usually approximate because we have to substitute an estimated value of $\sigma$, the common population standard deviation. This estimate will probably be based on an educated guess from information on a previous study or on the range of population values.

Corresponding sample sizes for one- and two-sided tests of $H_0 : \mu_1 - \mu_2 = D_0$ based on specific values of $\alpha$ and $\beta$ are shown here.

### Sample Sizes for Testing $H_0 : \mu_1 - \mu_2 = D_0$, Independent Samples

One-sided test: $n = 2\sigma^2 (z_{\alpha/2} + z_\beta)^2 / \Delta^2$

Two-sided test: $n = 2\sigma^2 (z_{\alpha/2} + z_\beta)^2 / \Delta^2$.

where $n_1 = n_2 = n$ and the probability of a Type II error is to be $\leq \beta$ when the true difference $|\mu_1 - \mu_2| \geq \Delta$.

**Note:** If $\sigma$ is unknown, substitute an estimated value to obtain an approximate sample size. Sample-size calculation can also be done using the formulas shown when $n_1 \neq n_2$. In this situation, we let $n_2$ be some multiple $m$ (e.g., $m = .5$) of $n_1$; then we substitute $(m+1)/m$ for 2 in the sample size formulas. After solving for $n_1$, $n_2 = mn_1$.

Sample sizes for estimating $\mu_d$ and conducting a statistical test for $\mu_d$ based on paired data (differences) are found using the formulas of Chapter 5 for $\mu$. 

13
The only change is that we’re working with a single sample of differences rather than a single sample of $y$ values. For convenience, the appropriate formulas are shown here.

\[
\text{Sample Sizes Required for a } 100(1 - \alpha)\% \text{ C.I. for } \mu_d \text{ of the Form } d \pm E \]

\[
n = \frac{z_{\alpha/2}^2 \sigma_d^2}{E^2}
\]

**Note:** If $\sigma_d$ is unknown, substitute an estimated value to obtain approximate sample size.

\[
\text{Sample Sizes for One- and Two-Sided Tests of } H_0 : \mu_d = D_0
\]

One-sided test: $n = \frac{\sigma_d^2(z_{\alpha} + z_\beta)^2}{\Delta^2}$

Two-sided test: $n = \frac{\sigma_d^2(z_{\alpha/2} + z_\beta)^2}{\Delta^2}$.

where the probability of a Type II error is $\beta$ or less if the true difference $\mu_d \geq \Delta$.

**Note:** If $\sigma_d$ is unknown, substitute an estimated value to obtain approximate sample size.

**EXAMPLE 6.9**

An experiment was done to determine the effect on dairy cattle of a diet supplemented with liquid whey. While no differences were noted in milk production measurements among cattle given a standard diet (7.5 kg of grain plus hay by choice) with water and those on the standard diet and liquid whey only, a considerable difference between the groups was noted in the amount of hay ingested. Suppose that one tests the null hypothesis of no difference in mean hay consumption for the two diet groups of dairy cattle. For a two-tailed test with $\alpha = .05$, determine the approximate number of dairy cattle that should be included in each group if we want $\beta \leq .10$ for $|\mu_1 - \mu_2| \geq .5$. Previous experimentation has shown $\sigma$ to be approximately .8.

**SOLUTION** From the description of the problem, we have $\alpha = .05, \beta \leq .10$ for $\Delta = |\mu_1 - \mu_2| \geq .5$ and $\sigma = .8$. Table 2 in the Appendix gives us $z_{.025} = 1.96$ and $z_{.10} = 1.28$. Substituting into the formula we have

\[
n \approx \frac{\frac{2(.8)^2(1.96 + 1.28)^2}{(.5)^2}}{53.75, \text{ or } 54.}
\]

That is, we need 54 cattle per group to run the desired test.
5. A Little Bit of History

Carl Friedrich Gauss

The normal probability distribution is often referred to as the Gaussian distribution in honor of Carl Gauss, the individual thought to have discovered the idea. However, it was actually Abraham de Moivre who first wrote down the equation of the normal distribution. Gauss was born in Brunswick, Germany, on April 30, 1777. Gauss' mathematical prowess was evident early in his life. At age eight he was able to instantly add the first 100 integers. In 1792, Gauss entered the Brunswick Collegium Corolinum and remained there for three years. In 1795, Gauss entered the University of Göttingen. In 1799, Gauss earned his doctorate. The subject of his dissertation was the Fundamental Theorem of Algebra. In 1809, Gauss published a book on the mathematics of planetary orbits. In this book, he further developed the theory of least-squares regression by analyzing the errors. The analysis of these errors led to the discovery that errors follow a normal distribution. Gauss was considered to be “glacially cold” as a person and had troubled relationships with his family. Gauss died February 23, 1855.

Abraham de Moivre

Abraham de Moivre was born in France on May 26, 1667. He is known as a great contributor to the areas of probability and trigonometry. De Moivre studied for five years at the Protestant academy at Sedan. From 1682 to 1684, he studied logic at Saumur. In 1685, he moved to England. De Moivre was elected a fellow of the Royal Society in 1697. He was part of the commission to settle the dispute between Newton and Leibniz regarding the discoverer of calculus. He published The Doctrine of Chance in 1718. In 1733, he developed the equation that describes the normal curve. Unfortunately, de Moivre had a difficult time being accepted in English society (perhaps due to his accent) and was able to make only a meager living tutoring mathematics. An interesting piece of information regarding de Moivre; he correctly predicted the day of his death, Nov. 27, 1754.

William Sealy Gossett

William Sealy Gossett was born on June 13, 1876, in Canterbury, England. Gossett earned a degree in chemistry from New College in Oxford in 1899. He then got a job as a chemist for the Guinness Brewing Company. Gossett, along with other chemists, was asked to find a way to make the best beer at the cheapest cost. This allowed him to concentrate on statistics. In 1904, Gossett wrote a paper on the brewing of beer that included a discussion of standard errors. In July, 1905, Gossett met with Karl Pearson to learn about the theory of standard errors.
the next few years, he developed his $t$-distribution. The Guinness Brewery did not allow its employees to publish, so Gossett published his research using the pen name Student. Gossett died October 16, 1937.
Choosing an appropriate statistic for inference about $\mu_1 - \mu_2$:

- **Option 2 is *preferred***

**Flowchart**:

1. Are the two samples dependent? **Yes**
   - Paired $t$ test (samples must come from normal populations):
     \[ t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} \]
     where $df = n - 1$.

2. Do $n_1$ and $n_2$ both exceed 30? **Yes** **option 1**
   - $z$ test (normal distribution):
     \[ z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]
     (If $\sigma_1^2$ and $\sigma_2^2$ are unknown, use $s_1^2$ and $s_2^2$ instead.)

3. Are both populations normally distributed? **Yes**
   - $z$ test (normal distribution):
     \[ z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]
     ($s_1^2$ and $s_2^2$ cannot be used to estimate $\sigma_1^2$ and $\sigma_2^2$.)

4. Are $\sigma_1$ and $\sigma_2$ both known? **No**
   - $t$ test (samples must come from normal populations):
     \[ t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]
     where $df$ is the smaller of $n_1 - 1$ and $n_2 - 1$.
     *It is better to use the Welch-Satterthwaite formula for the d.f. which is implemented in SAS.*

5. After applying the $F$ test. What do we conclude about $\sigma_1^2 = \sigma_2^2$? **Reject**
   - Fail to reject $\sigma_1^2 = \sigma_2^2$
1. Do fraternities help or hurt your academic progress at college? To investigate this question, 5 students who joined fraternities in 1998 were randomly selected. It was found that their GPA were as follows:

<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>After</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

(a) Please test the research hypothesis at the significance level 0.05

(1) Please test the research hypothesis at the significance level 0.05
(2) What assumption(s) did you assume in the above test?

2. The Food and Drug Administration (FDA) wants to compare the mean caffeine contents of two brands of 6-oz cola, Pepsi and Shasta. A typical test is the following: 5 cans of each brand were randomly selected. The results (in ounces) are as follows:

Pepsi: 17, 16, 19, 20, 18
Shasta: 22, 20, 18, 19, 21

(a) Please construct a confidence interval for the mean caffeine contents of Pepsi and Shasta.

(b) What is the p-value of your test?

(c) Please construct a 95% confidence interval for the mean delivery time.

1. Paired-Samples, Small Sample Sizes

\[ \bar{X}_d = 0.8, \quad S_d = 0.447, \quad n = 5, \quad \alpha = 0.05 \]

(a) \[ H_0 : \mu_d = 0, \quad H_a : \mu_d \neq 0. \]

\[ t_0 = \frac{\bar{X}_d - \mu_d}{s_d / \sqrt{n}} = \frac{0.8 - 0}{0.447 / \sqrt{5}} = 4.02. \quad ∴ \quad 4.02 > 2.306. \]

We can reject \( H_0 \).

(b) The distribution of the difference is normal.

2. Two independent Samples, small samples sizes

\[ n_1 = n_2 = 5 \]

\[ \bar{X}_1 = 18, \quad \bar{X}_2 = 20, \quad s_1 = 2.5, \quad s_2 = 2.5, \quad s_p = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 2.5 \times 1.58. \]

(a) 95% C.I. for \( \mu_1 - \mu_2 \):

\[ (\bar{X}_1 - \bar{X}_2) - t_{n-2} \cdot s_p / \sqrt{n} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{n-2} \cdot s_p / \sqrt{n} \]

\[ \Rightarrow (4.08, 0.08), (4.02, 0.02), (4.02, 0.02), (4.02, 0.02), (4.02, 0.02). \]

(b) \[ H_0 : \mu_1 = \mu_2, \quad H_a : \mu_1 \neq \mu_2. \]

\[ t_0 = \frac{\bar{X}_1 - \bar{X}_2}{s_p / \sqrt{n}} = -2. \quad ∴ \quad -2 < -2. \]

We can NOT reject \( H_a \).

OR Since 0 is inside the 95% C.I. for \( \mu_1 - \mu_2 \), therefore, we can not reject \( H_0 \) at \( \alpha = 0.05 \).

3. One sample, large sample size, about pop. Mean.

\[ n = 36, \quad \bar{X} = 36, \quad s = 5. \]

(a) \[ H_0 : \mu = 38, \quad H_a : \mu < 38 (\mu_0 = 38). \]

\[ z_0 = \frac{\bar{X} - \mu_0}{s / \sqrt{n}} = \frac{36 - 38}{5 / \sqrt{36}} = -3.6. \quad ∴ \quad -3.6 < -1.645. \]

We can reject \( H_0 \) at \( \alpha = 0.05 \).

(b) p-value = 0

(c) 95% C.I. is \( \bar{X} - z_{0.025} \frac{s}{\sqrt{n}} < \mu < \bar{X} + z_{0.025} \frac{s}{\sqrt{n}} \)

\[ = (33.367, 36.633). \]

AMS 315/576 Test 4

1. Do fraternities help or hurt your academic progress at college? To investigate this question, 5 fraternity members, and 5 non-fraternity members were randomly selected among the class of 1998. It was found that their GPA were as follows:

<table>
<thead>
<tr>
<th>Fraternity members</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-fraternity members</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

(a) Please test the research hypothesis at the significance level 0.05

(b) What assumption(s) did you assume in the above test?

2. A new weight-reducing diet is currently undergoing tests by the Food and Drug Administration (FDA). A typical test is the following: The weights of a random sample of 10 people are recorded before the diet as well as three weeks after the diet. The results (in pounds) are as follows:

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>150</td>
<td>195</td>
<td>188</td>
<td>197</td>
<td>204</td>
</tr>
<tr>
<td>After</td>
<td>143</td>
<td>190</td>
<td>185</td>
<td>101</td>
<td>200</td>
</tr>
</tbody>
</table>

(a) Please construct a confidence interval for the difference between the mean weights before and after the diet is used.

(b) At the significance level 0.05, can you conclude that the diet is effective?

(c) Please construct a 95% confidence interval for the mean delivery time.

1. Two independent Samples, Small Sample Sizes

\[ \bar{X}_1 = 2, \quad \bar{X}_2 = 3, \quad S_1^2 = \frac{1}{2}, \quad S_2^2 = \frac{1}{2}, \quad S_p = \sqrt{\frac{1}{2} + \frac{1}{2} - 0.71} \]

(a) \[ H_0 : \mu_1 = \mu_2, \quad H_a : \mu_1 \neq \mu_2, \quad \alpha = 0.05. \]

\[ t_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{1}{2} + \frac{1}{2}} - (0.71)} = 2.306. \]

We can reject \( H_0 \) if \( t_0 > 2.306 \) or \( t_0 < -2.306. \)

We can NOT reject \( H_0 \).

(b) normal populations, equal population variances

2. Paired Samples, small sample sizes

\[ \bar{X}_d = 3, \quad s_d = 4.18, \quad n = 5. \]

(a) 95% C.I. for \( \mu_d \),

\[ (\bar{X}_d - t_{n-2} \cdot s_d / \sqrt{n}), (\bar{X}_d + t_{n-2} \cdot s_d / \sqrt{n}) \]

Plug in \( t_{n-2} = 2.776, \quad \alpha = 0.025 \)

(b) \[ \bar{X}_d : \mu_d = 0, \quad H_a : \mu_d > 0. \]

\[ t_0 = \frac{\bar{X}_d}{s_d / \sqrt{n}} = \frac{3}{4.18 / \sqrt{5}} = 1.60. \]

We can NOT reject \( H_a \).

Or Since 0.01 value 0.025

3. One sample, population mean, small sample size.

Assume the population is normal \( n = 16, \quad \bar{X} = 35, \quad s = 5. \)

(a) \[ H_0 : \mu = 38, \quad H_a : \mu < 38 (\mu_0 = 38). \]

\[ t_0 = \frac{\bar{X} - \mu_0}{s / \sqrt{n}} = \frac{35 - 38}{5 / \sqrt{16}} = -2.4. \quad ∴ \quad -2.4 < -1.753. \]

We can reject \( H_0 \) at \( \alpha = 0.05 \).

(b) \[ t_{15,0.01} = 2.131, \quad 2.4 < 2.131. \quad ∴ \quad 0.01 < P-value < 0.025 \]

(c) 95% C.I. is

\[ (\bar{X} - t_{15,0.025} \cdot \frac{s}{\sqrt{n}}, \bar{X} + t_{15,0.025} \cdot \frac{s}{\sqrt{n}}) \]

\[ = (32.336, 37.664). \]