Chapter 7 dealt with the comparison of two independent samples with respect to a quantitative variable Y. In the present chapter we consider the comparison of two samples which are not independent but are paired. In a paired design, the observations \((Y_1, Y_2)\) occur in pairs; the observational units in a pair are linked in some way, so that they have more in common with each other than with members of another pair.

The following is an example of a paired design.

Certain types of nerve cells have the ability to regenerate a part of the cell that has been amputated. In an early study of this process, measurements were made on the nerves in the spinal cord in rhesus monkeys. Nerves emanating from the left side of the cord were cut, while nerves from the right side were kept intact. During the regeneration process, the content of creatine phosphate (CP) was measured in the left and the right portion of the spinal cord. Table 9.1 shows the data for the right (control) side \((Y_1)\), and for the left (regenerating) side \((Y_2)\).

The units of measurement are mg CP per 100 gm tissue.$^1$

<table>
<thead>
<tr>
<th>ANIMAL</th>
<th>RIGHT SIDE (CONTROL)</th>
<th>LEFT SIDE (REGENERATING)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.3</td>
<td>11.5</td>
</tr>
<tr>
<td>2</td>
<td>4.8</td>
<td>3.6</td>
</tr>
<tr>
<td>3</td>
<td>10.9</td>
<td>12.5</td>
</tr>
<tr>
<td>4</td>
<td>14.2</td>
<td>6.3</td>
</tr>
<tr>
<td>5</td>
<td>16.3</td>
<td>15.2</td>
</tr>
<tr>
<td>6</td>
<td>9.9</td>
<td>8.1</td>
</tr>
<tr>
<td>7</td>
<td>29.2</td>
<td>16.6</td>
</tr>
<tr>
<td>8</td>
<td>22.4</td>
<td>13.1</td>
</tr>
</tbody>
</table>

Mean 15.50 10.86
SD 7.61 4.49

In Example 9.1, the observational units (sides of the spinal cord) occur in pairs, and the members of a pair are linked by virtue of being in the same animal. A suitable analysis of the data should take advantage of this pairing.

In Section 9.2 we show how to analyze paired data using methods based on Student's \(t\) distribution. In Section 9.4 we describe a nonparametric test for paired data. Sections 9.3, 9.5, and 9.6 contain more examples and discussion of the paired design.

In this section we discuss the use of Student's \(t\) distribution to obtain tests and confidence intervals for paired data.

**STANDARD ERROR OF \((\bar{Y}_1 - \bar{Y}_2)\)**

To compare the means of two paired samples, we need to calculate a standard error for the difference between the means, that is,

\[
SE(\bar{Y}_1 - \bar{Y}_2)
\]
To calculate this SE, we make a simple shift of viewpoint: instead of considering \( Y_1 \) and \( Y_2 \) separately, we consider the difference \( d \), defined as
\[
   d = Y_1 - Y_2
\]

Let us denote the mean of the \( d \)'s as \( \bar{d} \). The quantity \( \bar{d} \) is related to the individual sample means as follows:
\[
   \bar{d} = \bar{y}_1 - \bar{y}_2
\]

The relationship between population means is analogous:
\[
   \mu_d = \mu_1 - \mu_2
\]

Thus, we may say that the mean of the difference is equal to the difference of the means. Because of this simple relationship, a comparison of the means can be carried out by concentrating entirely on the \( d \)'s.

A standard error for \( \bar{d} \) is easy to calculate. Because \( \bar{d} \) is just the mean of a single sample, we can apply the SE formula of Chapter 6 to obtain the following formula:
\[
   \text{SE}(\bar{y}_1 - \bar{y}_2) = \text{SE}_{\bar{d}} = \frac{s_d}{\sqrt{n_d}}
\]

where \( s_d \) is the standard deviation of the \( d \)'s and \( n_d \) is the number of \( d \)'s. The following example illustrates the calculation.

Table 9.2 shows the spinal cord CP data of Example 9.1 and the differences \( d \).

<table>
<thead>
<tr>
<th>ANIMAL</th>
<th>CONTROL SIDE</th>
<th>REGENERATING SIDE</th>
<th>( d = Y_1 - Y_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.3</td>
<td>11.5</td>
<td>4.8</td>
</tr>
<tr>
<td>2</td>
<td>4.8</td>
<td>3.6</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>10.9</td>
<td>12.5</td>
<td>-1.6</td>
</tr>
<tr>
<td>4</td>
<td>14.2</td>
<td>6.3</td>
<td>7.9</td>
</tr>
<tr>
<td>5</td>
<td>16.3</td>
<td>15.2</td>
<td>1.1</td>
</tr>
<tr>
<td>6</td>
<td>9.9</td>
<td>8.1</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>29.2</td>
<td>16.6</td>
<td>12.6</td>
</tr>
<tr>
<td>8</td>
<td>22.4</td>
<td>13.1</td>
<td>9.3</td>
</tr>
</tbody>
</table>

Mean 15.50 10.86 4.64
SD 7.61 4.49 4.89

Note that the mean of the difference is indeed the difference of the means:
\[
   \bar{d} = 4.64 = 15.50 - 10.86
\]
We calculate the standard error of this difference as follows:
\[ s_d = 4.8858 \]
\[ n_d = 8 \]
\[ SE(\bar{y}_1 - \bar{y}_2) = \frac{4.8858}{\sqrt{8}} = 1.727 \]

**CONFIDENCE INTERVAL AND TEST OF HYPOTHESIS**

The standard error described above is the basis for the paired-sample t method of analysis, which can take the form of a confidence interval or a test of hypothesis. A 95% confidence interval for \((\mu_1 - \mu_2)\) is constructed as
\[ (\bar{y}_1 - \bar{y}_2) \pm t_{0.05} \ SE(\bar{y}_1 - \bar{y}_2) \]
that is,
\[ \bar{d} \pm t_{0.05} \ SE_d \]
where the constant \(t_{0.05}\) is determined from Student's t distribution with
\[ df = n_d - 1 \]

Intervals with other confidence coefficients (such as 90%, 99%, etc.) are constructed analogously (using \(t_{1.00}, t_{0.01}, \text{etc.}\)). The following example illustrates the confidence interval.

**EXAMPLE 8.3**
**NERVE CELL REGENERATION**

For the spinal cord CP data, we have \(df = 8 - 1 = 7\). From Table 4 we find that \(t_{0.05} = 2.365\); thus, the 95% confidence interval for \((\mu_1 - \mu_2)\) is
\[ 15.50 - 10.86 \pm (2.365)(1.727) \]
\[ 4.64 \pm 4.08 \]
\[ .6 < \mu_1 - \mu_2 < 8.7 \text{ mg/100 gm} \]

We can also compare the two means with a t test. To test the null hypothesis
\[ H_0: \ \mu_1 = \mu_2 \]
we use the test statistic
\[ t_s = \frac{\bar{y}_1 - \bar{y}_2}{SE(\bar{y}_1 - \bar{y}_2)} \]
that is,
\[ t_s = \frac{\bar{d}}{SE_d} \]
Critical values are obtained from Student's \( t \) distribution (Table 4) with \( df = n_d - 1 \). The following example illustrates the \( t \) test.

For the spinal cord CP data, let us formulate the null hypothesis and nondirectional alternative:

\[
H_0: \text{ Mean CP is the same in regenerating and in normal tissue} \\
H_A: \text{ Mean CP is different in regenerating and in normal tissue}
\]

or, in symbols,

\[
H_0: \mu_1 = \mu_2 \\
H_A: \mu_1 \neq \mu_2
\]

Let us test \( H_0 \) against \( H_A \) at significance level \( \alpha = .05 \). The test statistic is

\[
t_s = \frac{4.64}{1.727} = 2.69
\]

From Table 4, \( t_{.05} = 2.365 \) and \( t_{.02} = 2.998 \). We reject \( H_0 \) and find that there is sufficient evidence (.02 < \( P < .05 \)) to conclude that mean CP is lower in regenerating than in normal tissue.  

The confidence interval and \( t \) test formulas for paired data are directly analogous to those given in Chapter 7 for analysis of independent samples. The formula for a 95% confidence interval is

\[
(\bar{y}_1 - \bar{y}_2) \pm t_{.05} SE(\bar{y}_1 - \bar{y}_2)
\]

and for the \( t \) statistic is

\[
t_s = \frac{\bar{y}_1 - \bar{y}_2}{SE(\bar{y}_1 - \bar{y}_2)}
\]

Notice that these formulas appear identical to those given in Chapter 7. However, the paired-sample analysis is different from the independent-samples analysis in two respects: the SE is calculated differently and the number of degrees of freedom is different.

RESULT OF IGNORING PAIRING

Suppose that a study is conducted using a paired design, but that the pairing is ignored in the analysis of the data. Such an analysis is not valid because it assumes that the samples are independent when in fact they are not. The incorrect analysis can be grossly inadequate, as the following example illustrates.
For the spinal cord CP data, the two sample SDs are $s_1 = 7.609$ and $s_2 = 4.488$. If we proceed as if the samples were independent and apply the SE formula of Chapter 7, we obtain

$$SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= \sqrt{\frac{(7.609)^2}{8} + \frac{(4.488)^2}{8}}$$

$$= 3.123$$

This SE is much larger than the value ($SE = 1.727$) that we calculated using the pairing.

Continuing to proceed as if the samples were independent, we calculate $df = n_1 + n_2 - 2 = 14$. If we conduct a $t$ test, our test statistic is

$$t = \frac{4.64}{3.123} = 1.49$$

which does not approach significance, even at $\alpha = .10$. When we used a paired-sample analysis of these data (Example 9.4), we found evidence of a treatment effect; the present analysis, ignoring the pairing, fails to discover this evidence.

To further compare the paired and unpaired analysis, let us consider the 95% confidence interval for $(\mu_1 - \mu_2)$. For the unpaired analysis, we find $t_{0.05} = 2.145$ with $df = 14$; this yields the confidence interval

$$4.64 \pm (2.145)(3.123)$$

$$4.64 \pm 6.70$$

The paired analysis (Example 9.3) yielded the narrower interval

$$4.64 \pm (2.365)(1.727)$$

$$4.64 \pm 4.08$$

The paired-sample interval is narrower because it uses a smaller SE; this effect is slightly offset by a larger value of $t_{0.05} (2.365$ vs $2.145)$.

Why is the paired-sample SE so much smaller than the independent-samples SE calculated from the same data ($SE = 1.7$ vs $SE = 3.1$)? Table 9.1 reveals the reason. The data show that there is large variation from one animal to the next. For instance, animal 2 has very low CP values (on both the right and left sides) and animal 7 has very high values. The independent-samples SE formula incorporates all of this variation (expressed through $s_1$ and $s_2$); in the paired-sample approach, interanimal variation in CP level has no influence on the calculations because only the $d's$ are used. By using each animal as its own control, the experimenter has greatly increased the precision of the experiment. But if the pairing is ignored in the analysis, the extra precision is wasted.