CHAPTER OUTLINE

12-1 MULTIPLE LINEAR REGRESSION MODEL
   12-1.1 Introduction
   12-1.2 Least Squares Estimation of the Parameters
   12-1.3 Matrix Approach to Multiple Linear Regression
   12-1.4 Properties of the Least Squares Estimators

12-2 HYPOTHESIS TESTS IN MULTIPLE LINEAR REGRESSION
   12-2.1 Test for Significance of Regression
   12-2.2 Tests on Individual Regression Coefficients and Subsets of Coefficients

12-3 CONFIDENCE INTERVALS IN MULTIPLE LINEAR REGRESSION

12-3.1 Confidence Intervals on Individual Regression Coefficients
12-3.2 Confidence Interval on the Mean Response

12-4 PREDICTION OF NEW OBSERVATIONS

12-5 MODEL ADEQUACY CHECKING
   12-5.1 Residual Analysis
   12-5.2 Influential Observations

12-6 ASPECTS OF MULTIPLE REGRESSION MODELING
   12-6.1 Polynomial Regression Models
   12-6.2 Categorical Regressors and Indicator Variables
   12-6.3 Selection of Variables and Model Building
   12-6.4 Multicollinearity
12-1 Multiple Linear Regression Models

12-1.1 Introduction

- Many applications of regression analysis involve situations in which there are more than one regressor variable.

- A regression model that contains more than one regressor variable is called a **multiple regression model**.
12-1.1 Introduction

- For example, suppose that the effective life of a cutting tool depends on the cutting speed and the tool angle. A possible multiple regression model could be

\[ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \]

where

\( Y \) – tool life

\( x_1 \) – cutting speed

\( x_2 \) – tool angle
12-1 Multiple Linear Regression Models

12-1.1 Introduction

**Figure 12-1** (a) The regression plane for the model $E(Y) = 50 + 10x_1 + 7x_2$. (b) The contour plot
12-1 Multiple Linear Regression Models

12-1.1 Introduction

In general, the dependent variable or response $Y$ may be related to $k$ independent or regressor variables. The model

$$ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon $$

(12-2)

is called a multiple linear regression model with $k$ regressor variables. The parameters $\beta_j, j = 0, 1, \ldots, k$, are called the regression coefficients. This model describes a hyperplane in the $k$-dimensional space of the regressor variables $\{x_j\}$. The parameter $\beta_j$ represents the expected change in response $Y$ per unit change in $x_j$ when all the remaining regressors $x_i (i \neq j)$ are held constant.
12-1 Multiple Linear Regression Models

12-1.1 Introduction

Figure 12-2 (a) Three-dimensional plot of the regression model $E(Y) = 50 + 10x_1 + 7x_2 + 5x_1x_2$.

(b) The contour plot
12-1 Multiple Linear Regression Models

12-1.1 Introduction

**Figure 12-3** (a) 3-D plot of the regression model

\[ E(Y) = 800 + 10x_1 + 7x_2 - 8.5x_1^2 - 5x_2^2 + 4x_1x_2. \]

(b) The contour plot
12-1 Multiple Linear Regression Models

12-1.2 Least Squares Estimation of the Parameters

The method of least squares may be used to estimate the regression coefficients in the multiple regression model, Equation 12-2. Suppose that \( n > k \) observations are available, and let \( x_{ij} \) denote the \( i \)th observation or level of variable \( x_j \). The observations are

\[
(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i), \quad i = 1, 2, \ldots, n \quad \text{and} \quad n > k
\]

It is customary to present the data for multiple regression in a table such as Table 12-1.

<table>
<thead>
<tr>
<th>( y )</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>\ldots</th>
<th>( x_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_1 )</td>
<td>( x_{11} )</td>
<td>( x_{12} )</td>
<td>\ldots</td>
<td>( x_{1k} )</td>
</tr>
<tr>
<td>( y_2 )</td>
<td>( x_{21} )</td>
<td>( x_{22} )</td>
<td>\ldots</td>
<td>( x_{2k} )</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\ldots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( y_n )</td>
<td>( x_{n1} )</td>
<td>( x_{n2} )</td>
<td>\ldots</td>
<td>( x_{nk} )</td>
</tr>
</tbody>
</table>
12-1 Multiple Linear Regression Models

12-1.2 Least Squares Estimation of the Parameters

- The least squares function is given by

\[
L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2
\]

- The least squares estimates must satisfy

\[
\frac{\partial L}{\partial \beta_0} \bigg|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = 0
\]
\[
= -2 \sum_{i=1}^{n} \left( y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0
\]

and

\[
\frac{\partial L}{\partial \beta_j} \bigg|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = 0
\]
\[
= -2 \sum_{i=1}^{n} \left( y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) x_{ij} = 0 \quad j = 1, 2, \ldots, k
\]
12-1 Multiple Linear Regression Models

12-1.2 Least Squares Estimation of the Parameters

• The least squares normal Equations are

\[
\begin{align*}
    n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} &= \sum_{i=1}^{n} y_i \\
    \hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1} x_{ik} &= \sum_{i=1}^{n} x_{i1} y_i \\
    \vdots & \quad \vdots & \quad \vdots & \quad \vdots & \\
    \hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ik} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 &= \sum_{i=1}^{n} x_{ik} y_i
\end{align*}
\]

• The solution to the normal Equations are the least squares estimators of the regression coefficients.
Suppose the model relating the regressors to the response is

\[ y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \quad i = 1, 2, \ldots, n \]

In matrix notation this model can be written as

\[ y = X\beta + \epsilon \]
12-1 Multiple Linear Regression Models

12-1.3 Matrix Approach to Multiple Linear Regression

\[ y = X\beta + \epsilon \]

where

\[ y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad x = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ and } \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \]
12-1 Multiple Linear Regression Models

12-1.3 Matrix Approach to Multiple Linear Regression

We wish to find the vector of least squares estimators that minimizes:

\[ L = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon'\epsilon = (y - X\beta)'(y - X\beta) \]

The resulting least squares estimate is

\[ \hat{\beta} = (X'X)^{-1} X'y \]  \hspace{1cm} (12-13)
12-1 Multiple Linear Regression Models

12-1.3 Matrix Approach to Multiple Linear Regression

The fitted regression model is

\[ \hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \quad i = 1, 2, \ldots, n \]  (12-14)

In matrix notation, the fitted model is

\[ \hat{y} = X\hat{\beta} \]

The difference between the observation \( y_i \) and the fitted value \( \hat{y}_i \) is a residual, say, \( e_i = y_i - \hat{y}_i \). The \((n \times 1)\) vector of residuals is denoted by

\[ e = y - \hat{y} \]  (12-15)
12-1 Multiple Linear Regression Models

Example 12-2

Table 12-2  Wire Bond Data for Example 11-1

<table>
<thead>
<tr>
<th>Observation Number</th>
<th>Pull Strength $y$</th>
<th>Wire Length $x_1$</th>
<th>Die Height $x_2$</th>
<th>Observation Number</th>
<th>Pull Strength $y$</th>
<th>Wire Length $x_1$</th>
<th>Die Height $x_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.95</td>
<td>2</td>
<td>50</td>
<td>14</td>
<td>11.66</td>
<td>2</td>
<td>360</td>
</tr>
<tr>
<td>2</td>
<td>24.45</td>
<td>8</td>
<td>110</td>
<td>15</td>
<td>21.65</td>
<td>4</td>
<td>205</td>
</tr>
<tr>
<td>3</td>
<td>31.75</td>
<td>11</td>
<td>120</td>
<td>16</td>
<td>17.89</td>
<td>4</td>
<td>400</td>
</tr>
<tr>
<td>4</td>
<td>35.00</td>
<td>10</td>
<td>550</td>
<td>17</td>
<td>69.00</td>
<td>20</td>
<td>600</td>
</tr>
<tr>
<td>5</td>
<td>25.02</td>
<td>8</td>
<td>295</td>
<td>18</td>
<td>10.30</td>
<td>1</td>
<td>585</td>
</tr>
<tr>
<td>6</td>
<td>16.86</td>
<td>4</td>
<td>200</td>
<td>19</td>
<td>34.93</td>
<td>10</td>
<td>540</td>
</tr>
<tr>
<td>7</td>
<td>14.38</td>
<td>2</td>
<td>375</td>
<td>20</td>
<td>46.59</td>
<td>15</td>
<td>250</td>
</tr>
<tr>
<td>8</td>
<td>9.60</td>
<td>2</td>
<td>52</td>
<td>21</td>
<td>44.88</td>
<td>15</td>
<td>290</td>
</tr>
<tr>
<td>9</td>
<td>24.35</td>
<td>9</td>
<td>100</td>
<td>22</td>
<td>54.12</td>
<td>16</td>
<td>510</td>
</tr>
<tr>
<td>10</td>
<td>27.50</td>
<td>8</td>
<td>300</td>
<td>23</td>
<td>56.63</td>
<td>17</td>
<td>590</td>
</tr>
<tr>
<td>11</td>
<td>17.08</td>
<td>4</td>
<td>412</td>
<td>24</td>
<td>22.13</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>37.00</td>
<td>11</td>
<td>400</td>
<td>25</td>
<td>21.15</td>
<td>5</td>
<td>400</td>
</tr>
<tr>
<td>13</td>
<td>41.95</td>
<td>12</td>
<td>500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
12-1 Multiple Linear Regression Models

Figure 12-4 Matrix of scatter plots (from Minitab) for the wire bond pull strength data in Table 12-2.
### Example 12-2

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>50</td>
<td>9.95</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>110</td>
<td>24.45</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>120</td>
<td>31.75</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>550</td>
<td>35.00</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>295</td>
<td>25.02</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>200</td>
<td>16.86</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>375</td>
<td>14.38</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>52</td>
<td>9.60</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>100</td>
<td>24.35</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>300</td>
<td>27.50</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>412</td>
<td>17.08</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>400</td>
<td>37.00</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>500</td>
<td>41.95</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>360</td>
<td>11.66</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>205</td>
<td>21.65</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>400</td>
<td>17.89</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>600</td>
<td>69.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>585</td>
<td>10.30</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>540</td>
<td>34.93</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>250</td>
<td>46.59</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>290</td>
<td>44.88</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>510</td>
<td>54.12</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>590</td>
<td>56.63</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>100</td>
<td>22.13</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>400</td>
<td>21.15</td>
</tr>
</tbody>
</table>

\[ X = \begin{bmatrix} 1 & 2 & 50 \\ 1 & 8 & 110 \\ 1 & 11 & 120 \\ 1 & 10 & 550 \\ 1 & 8 & 295 \\ 1 & 4 & 200 \\ 1 & 2 & 375 \\ 1 & 2 & 52 \\ 1 & 9 & 100 \\ 1 & 8 & 300 \\ 1 & 4 & 412 \\ 1 & 11 & 400 \\ 1 & 12 & 500 \\ 1 & 2 & 360 \\ 1 & 4 & 205 \\ 1 & 4 & 400 \\ 1 & 20 & 600 \\ 1 & 1 & 585 \\ 1 & 10 & 540 \\ 1 & 15 & 250 \\ 1 & 15 & 290 \\ 1 & 16 & 510 \\ 1 & 17 & 590 \\ 1 & 6 & 100 \\ 1 & 5 & 400 \end{bmatrix}, \ y = \begin{bmatrix} 9.95 \\ 24.45 \\ 31.75 \\ 35.00 \\ 25.02 \\ 16.86 \\ 14.38 \\ 9.60 \\ 24.35 \\ 27.50 \\ 17.08 \\ 37.00 \\ 41.95 \\ 11.66 \\ 21.65 \\ 17.89 \\ 69.00 \\ 10.30 \\ 34.93 \\ 46.59 \\ 44.88 \\ 54.12 \\ 56.63 \\ 22.13 \\ 21.15 \end{bmatrix} \]
12-1 Multiple Linear Regression Models

Example 12-2

The \( X'X \) matrix is

\[
X'X = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
2 & 8 & \cdots & 5 \\
50 & 110 & \cdots & 400
\end{bmatrix}
\begin{bmatrix}
1 & 2 & 50 \\
1 & 8 & 110 \\
\vdots & \vdots & \vdots \\
1 & 5 & 400
\end{bmatrix}
= \begin{bmatrix}
25 & 206 & 8,294 \\
206 & 2,396 & 77,177 \\
8,294 & 77,177 & 3,531,848
\end{bmatrix}
\]

and the \( X'y \) vector is

\[
X'y = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
2 & 8 & \cdots & 5 \\
50 & 110 & \cdots & 400
\end{bmatrix}
\begin{bmatrix}
9.95 \\
24.45 \\
\vdots \\
21.15
\end{bmatrix}
= \begin{bmatrix}
725.82 \\
8,008.37 \\
274,811.31
\end{bmatrix}
\]

The least squares estimates are found from Equation 12-13 as

\[
\hat{\beta} = (X'X)^{-1}X'y
\]
Example 12-2

or

\[
\begin{bmatrix}
\hat{\beta}_0 \\
\hat{\beta}_1 \\
\hat{\beta}_2
\end{bmatrix}
= \begin{bmatrix}
25 & 206 & 8,294 \\
206 & 2,396 & 77,177 \\
8,294 & 77,177 & 3,531,848
\end{bmatrix}^{-1}
\begin{bmatrix}
725.82 \\
8,008.37 \\
274,811.31
\end{bmatrix}
\]

\[
= \begin{bmatrix}
0.214653 & -0.007491 & -0.000340 \\
-0.007491 & 0.001671 & -0.000019 \\
-0.000340 & -0.000019 & +0.0000015
\end{bmatrix}
\begin{bmatrix}
725.82 \\
8,008.47 \\
274,811.31
\end{bmatrix}
= \begin{bmatrix}
2.26379143 \\
2.74426964 \\
0.01252781
\end{bmatrix}
\]

Therefore, the fitted regression model with the regression coefficients rounded to five decimal places is

\[
\hat{y} = 2.26379 + 2.74427x_1 + 0.01253x_2
\]

This is identical to the results obtained in Example 12-1.
12-1 Multiple Linear Regression Models

Example 12-2

This regression model can be used to predict values of pull strength for various values of wire length \(x_1\) and die height \(x_2\). We can also obtain the fitted values \(\hat{y}_i\) by substituting each observation \((x_{i1}, x_{i2})\), \(i = 1, 2, \ldots, n\), into the equation. For example, the first observation has \(x_{11} = 2\) and \(x_{12} = 50\), and the fitted value is

\[
\hat{y}_1 = 2.26379 + 2.74427x_{11} + 0.01253x_{12} \\
= 2.26379 + 2.74427(2) + 0.01253(50) \\
= 8.38
\]

The corresponding observed value is \(y_1 = 9.95\). The residual corresponding to the first observation is

\[
e_1 = y_1 - \hat{y}_1 \\
= 9.95 - 8.38 \\
= 1.57
\]

Table 12-3 displays all 25 fitted values \(\hat{y}_i\) and the corresponding residuals. The fitted values and residuals are calculated to the same accuracy as the original data.
12-1 Multiple Linear Regression Models

Example 12-2

<table>
<thead>
<tr>
<th>Observation Number</th>
<th>$y_t$</th>
<th>$\hat{y}_t$</th>
<th>$e_t = y_t - \hat{y}_t$</th>
<th>Observation Number</th>
<th>$y_t$</th>
<th>$\hat{y}_t$</th>
<th>$e_t = y_t - \hat{y}_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.95</td>
<td>8.38</td>
<td>1.57</td>
<td>14</td>
<td>11.66</td>
<td>12.26</td>
<td>-0.60</td>
</tr>
<tr>
<td>2</td>
<td>24.45</td>
<td>25.60</td>
<td>-1.15</td>
<td>15</td>
<td>21.65</td>
<td>15.81</td>
<td>5.84</td>
</tr>
<tr>
<td>3</td>
<td>31.75</td>
<td>33.95</td>
<td>-2.20</td>
<td>16</td>
<td>17.89</td>
<td>18.25</td>
<td>-0.36</td>
</tr>
<tr>
<td>4</td>
<td>35.00</td>
<td>36.60</td>
<td>-1.60</td>
<td>17</td>
<td>69.00</td>
<td>64.67</td>
<td>4.33</td>
</tr>
<tr>
<td>5</td>
<td>25.02</td>
<td>27.91</td>
<td>-2.89</td>
<td>18</td>
<td>10.30</td>
<td>12.34</td>
<td>-2.04</td>
</tr>
<tr>
<td>6</td>
<td>16.86</td>
<td>15.75</td>
<td>1.11</td>
<td>19</td>
<td>34.93</td>
<td>36.47</td>
<td>-1.54</td>
</tr>
<tr>
<td>7</td>
<td>14.38</td>
<td>12.45</td>
<td>1.93</td>
<td>20</td>
<td>46.59</td>
<td>46.56</td>
<td>-0.03</td>
</tr>
<tr>
<td>8</td>
<td>9.60</td>
<td>8.40</td>
<td>1.20</td>
<td>21</td>
<td>44.88</td>
<td>47.06</td>
<td>-2.18</td>
</tr>
<tr>
<td>9</td>
<td>24.35</td>
<td>28.21</td>
<td>-3.86</td>
<td>22</td>
<td>54.12</td>
<td>52.56</td>
<td>1.56</td>
</tr>
<tr>
<td>10</td>
<td>27.50</td>
<td>27.98</td>
<td>-0.48</td>
<td>23</td>
<td>56.63</td>
<td>56.31</td>
<td>0.32</td>
</tr>
<tr>
<td>11</td>
<td>17.08</td>
<td>18.40</td>
<td>-1.32</td>
<td>24</td>
<td>22.13</td>
<td>19.98</td>
<td>2.15</td>
</tr>
<tr>
<td>12</td>
<td>37.00</td>
<td>37.46</td>
<td>-0.46</td>
<td>25</td>
<td>21.15</td>
<td>21.00</td>
<td>0.15</td>
</tr>
<tr>
<td>13</td>
<td>41.95</td>
<td>41.46</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 12-4  Minitab Multiple Regression Output for the Wire Bond Pull Strength Data

Regression Analysis: Strength versus Length, Height

The regression equation is
Strength = 2.26 + 2.74 Length + 0.0125 Height

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.264</td>
<td>1.060</td>
<td>2.14</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>2.74427</td>
<td>0.09352</td>
<td>29.34</td>
<td>0.000</td>
<td>1.2</td>
</tr>
<tr>
<td>Height</td>
<td>0.012528</td>
<td>0.002798</td>
<td>4.48</td>
<td>0.000</td>
<td>1.2</td>
</tr>
</tbody>
</table>

S = 2.288  R-Sq = 98.1%  R-Sq (adj) = 97.9%
PRESS = 156.163  R-Sq (pred) = 97.44%

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>5990.8</td>
<td>2995.4</td>
<td>572.17</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>22</td>
<td>115.2</td>
<td>5.2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>6105.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source | DF | Seq SS  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>1</td>
<td>5885.9</td>
</tr>
<tr>
<td>Height</td>
<td>1</td>
<td>104.9</td>
</tr>
</tbody>
</table>

Predicted Values for New Observations

<table>
<thead>
<tr>
<th>New Obs</th>
<th>Fit</th>
<th>SE Fit</th>
<th>95.0% CI</th>
<th>95.0% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.663</td>
<td>0.482</td>
<td>(26.663, 28.663)</td>
<td>(22.814, 32.512)</td>
</tr>
</tbody>
</table>

Values of Predictors for New Observations

<table>
<thead>
<tr>
<th>New Obs</th>
<th>Length</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.00</td>
<td>275</td>
</tr>
</tbody>
</table>
12-1 Multiple Linear Regression Models

**Estimating \( \sigma^2 \)**

An unbiased estimator of \( \sigma^2 \) is

\[
\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p} = \frac{SS_E}{n-p} \tag{12-16}
\]
12-1 Multiple Linear Regression Models

12-1.4 Properties of the Least Squares Estimators

Unbiased estimators:

\[ E(\hat{\beta}) = E[(X'X)^{-1}X'Y] \]
\[ = E[(X'X)^{-1}X'(X\beta + \epsilon)] \]
\[ = E[(X'X)^{-1}X'X\beta + (X'X)^{-1}X'\epsilon] \]
\[ = \beta \]

Covariance Matrix:

\[ C = (X'X)^{-1} = \begin{bmatrix} C_{00} & C_{01} & C_{02} \\ C_{10} & C_{11} & C_{12} \\ C_{20} & C_{21} & C_{22} \end{bmatrix} \]
12-1 Multiple Linear Regression Models

12-1.4 Properties of the Least Squares Estimators

Individual variances and covariances:

\[ V(\hat{\beta}_j) = \sigma^2 C_{jj}, \quad j = 0, 1, 2 \]
\[ \text{cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 C_{ij}, \quad i \neq j \]

In general,

\[ \text{cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 C \]
12-2 Hypothesis Tests in Multiple Linear Regression

12-2.1 Test for Significance of Regression

The appropriate hypotheses are

\[ H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0 \]
\[ H_1: \beta_j \neq 0 \quad \text{for at least one } j \]  \hspace{1cm} (12-17)

The test statistic is

\[ F_0 = \frac{SS_R/k}{SS_E/(n - p)} = \frac{MS_R}{MS_E} \]  \hspace{1cm} (12-18)
### 12-2 Hypothesis Tests in Multiple Linear Regression

#### 12-2.1 Test for Significance of Regression

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>$F_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>$SS_R$</td>
<td>$k$</td>
<td>$MS_R$</td>
<td>$MS_R/MS_E$</td>
</tr>
<tr>
<td>Error or residual</td>
<td>$SS_E$</td>
<td>$n - p$</td>
<td>$MS_E$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$SS_T$</td>
<td>$n - 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12-9: Analysis of Variance for Testing Significance of Regression in Multiple Regression.
12-2 Hypothesis Tests in Multiple Linear Regression

Example 12-3

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>$f_0$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>5990.7712</td>
<td>2</td>
<td>2995.3856</td>
<td>572.17</td>
<td>1.08E-19</td>
</tr>
<tr>
<td>Error or residual</td>
<td>115.1735</td>
<td>22</td>
<td>5.2352</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6105.9447</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis of variance is shown in Table 12-10. To test $H_0: \beta_1 = \beta_2 = 0$, we calculate the statistic

$$f_0 = \frac{MS_R}{MS_E} = \frac{2995.3856}{5.2352} = 572.17$$

Since $f_0 > f_{0.05,2,22} = 3.44$ (or since the $P$-value is considerably smaller than $\alpha = 0.05$), we reject the null hypothesis and conclude that pull strength is linearly related to either wire length or die height, or both. However, we note that this does not necessarily imply that the
R² and Adjusted R²

The coefficient of multiple determination

\[ R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \]

- For the wire bond pull strength data, we find that \( R^2 = \frac{SS_R}{SS_T} = 5990.7712/6105.9447 = 0.9811. \)
- Thus, the model accounts for about 98% of the variability in the pull strength response.
12-2 Hypothesis Tests in Multiple Linear Regression

R² and Adjusted R²

The adjusted R² is

\[ R_{adj}^2 = 1 - \frac{SS_E/(n - p)}{SS_T/(n - 1)} \]  

(12-22)

- The adjusted R² statistic penalizes the analyst for adding terms to the model.
- It can help guard against overfitting (including regressors that are not really useful)
12-2 Hypothesis Tests in Multiple Linear Regression

12-2.2 Tests on Individual Regression Coefficients and Subsets of Coefficients

- Reject $H_0$ if $|t_0| > t_{\alpha/2, n-p}$.
- This is called a **partial** or **marginal test**.
Example 12-4

Consider the wire bond pull strength data, and suppose that we want to test the hypothesis that the regression coefficient for $x_2$ (die height) is zero. The hypotheses are

$$H_0: \beta_2 = 0$$
$$H_1: \beta_2 \neq 0$$

The main diagonal element of the $(X'X)^{-1}$ matrix corresponding to $\hat{\beta}_2$ is $C_{22} = 0.0000015$, so the $t$-statistic in Equation 12-24 is

$$t_0 = \frac{\hat{\beta}_2}{\sqrt{\hat{\sigma}^2 C_{22}}} = \frac{0.01253}{\sqrt{(5.2352)(0.0000015)}} = 4.4767$$

Note that we have used the estimate of $\sigma^2$ reported to four decimal places in Table 12-10. Since $t_{0.025,22} = 2.074$, we reject $H_0: \beta_2 = 0$ and conclude that the variable $x_2$ (die height) contributes significantly to the model. We could also have used a $P$-value to draw conclusions.
12-2 Hypothesis Tests in Multiple Linear Regression

Example 12-4

The $P$-value for $t_0 = 4.4767$ is $P = 0.0002$, so with $\alpha = 0.05$ we would reject the null hypothesis. Note that this test measures the marginal or partial contribution of $x_2$ given that $x_1$ is in the model. That is, the $t$-test measures the contribution of adding the variable $x_2 = \text{die height}$ to a model that already contains $x_1 = \text{wire length}$. Table 12-4 shows the value of the $t$-test computed by Minitab. The Minitab $t$-test statistic is reported to two decimal places. Note that the computer produces a $t$-test for each regression coefficient in the model. These $t$-tests indicate that both regressors contribute to the model.
R commands and outputs

> dat=read.table("http://www.stat.ucla.edu/~hqxu/stat105/data/table12_2.txt", h=T)
> g=lm(Strength~Length+Height, dat)
> summary(g)

    Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.263791   1.060066   2.136 0.044099 *
Length      2.744270   0.093524  29.343  < 2e-16 ***
Height      0.012528   0.002798   4.477 0.000188 ***

Residual standard error: 2.288 on 22 degrees of freedom
Multiple R-Squared: 0.9811,     Adjusted R-squared: 0.9794
F-statistic: 572.2 on 2 and 22 DF,  p-value: < 2.2e-16