The Multiple Regression Model

Chapter 5

Prepared by Vera Tabakova, East Carolina University

Chapter 5: The Multiple Regression Model

- 5.1 Model Specification and Data
- 5.2 Estimating the Parameters of the Multiple Regression Model
- 5.3 Sampling Properties of the Least Squares Estimator
- 5.4 Interval Estimation
- 5.5 Hypothesis Testing for a Single Coefficient
- 5.6 Measuring Goodness-of-Fit

5.1.1 The Economic Model

$$S = \beta_1 + \beta_2 P + \beta_3 A \tag{5.1}$$

- β_2 = the change in monthly sales S (\$1000) when the price index P is increased by one unit (\$1), and advertising expenditure A is held constant

$$= \left.\frac{\Delta S}{\Delta P}\right|_{(A \text{ held constant})} = \frac{\partial S}{\partial P}$$

- β_3 = the change in monthly sales S (\$1000) when advertising expenditure A is increased by one unit (\$1000), and the price index P is held constant

$$= \left.\frac{\Delta S}{\Delta A}\right|_{(P \text{ held constant})} = \frac{\partial S}{\partial A}$$

5.1.2 The Econometric Model

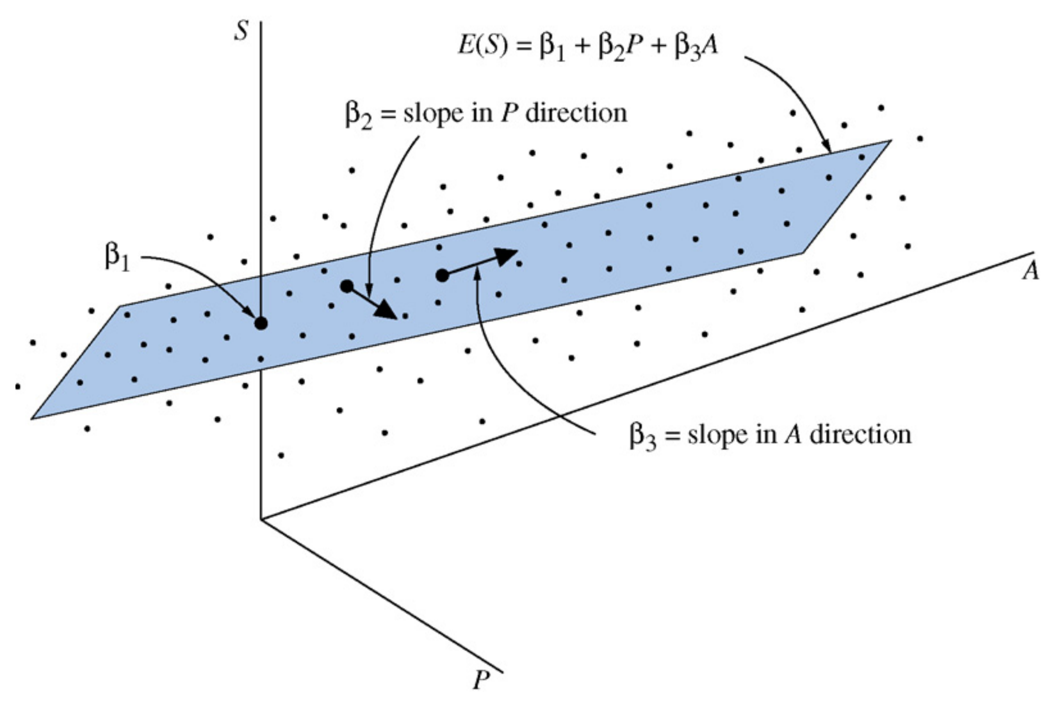

Figure 5.1 The multiple regression plane

5.1.2 The Econometric Model

Table 5.1 **Observations on Monthly Sales, Price, and Advertising in Big Andy's Burger Barn**

City	Sales (S) \$1000 units	Price (P) \$1 units	Advertising (A) \$1000 units
1	73.2	5.69	1.3
2	71.8	6.49	2.9
3	62.4	5.63	0.8
4	67.4	6.22	0.7
5	89.3	5.02	1.5
.	.	.	.
.	.	.	.
.	.	.	.
73	75.4	5.71	0.7
74	81.3	5.45	2.0
75	75.0	6.05	2.2
	Summary statistics		
Sample mean	77.37	5.69	1.84
Median	76.50	5.69	1.80
Maximum	91.20	6.49	3.10
Minimum	62.40	4.83	0.50
Std. Dev.	6.4885	0.5184	0.8317

5.1.2 The Econometric Model

$$S_i = E(S_i) + e_i = \beta_1 + \beta_2 P_i + \beta_3 A_i + e_i \quad (5.2)$$

- The introduction of the error term, and assumptions about its probability distribution, turn the economic model into the **econometric model** in (5.2).

5.1.2a The General Model

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_K x_{iK} + e_i \qquad (5.3)$$

$$\beta_k = \left.\frac{\Delta E(y)}{\Delta x_k}\right|_{\text{other } x\text{'s held constant}} = \frac{\partial E(y)}{\partial x_k}$$

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e \qquad (5.4)$$

5.1.2b The Assumptions of the Model

1. $E(e_i) = 0$

- Each random error has a probability distribution with zero mean. Some errors will be positive, some will be negative; over a large number of observations they will average out to zero.

5.1.2b The Assumptions of the Model

2. $\text{var}(e_i) = \sigma^2$

- Each random error has a probability distribution with variance σ^2. The variance σ^2 is an unknown parameter and it measures the uncertainty in the statistical model. It is the same for each observation, so that for no observations will the model uncertainty be more, or less, nor is it directly related to any economic variable. Errors with this property are said to be **homoskedastic**.

5.1.2b The Assumptions of the Model

3. $\text{cov}(e_i, e_j) = 0$

- The covariance between the two random errors corresponding to any two different observations is zero. The size of an error for one observation has no bearing on the likely size of an error for another observation. Thus, any pair of errors is uncorrelated.

5.1.2b The Assumptions of the Model

4. $e_i \sim N\left(0, \sigma^2\right)$

- We will sometimes further assume that the random errors have normal probability distributions.

5.1.2b The Assumptions of the Model

The statistical properties of y_i follow from the properties of e_i.

1. $$E(y_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$$

- The expected (average) value of y_i depends on the values of the explanatory variables and the unknown parameters. It is equivalent to $E(e_i) = 0$. This assumption says that the average value of y_i changes for each observation and is given by the **regression function** $E(y_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$.

5.1.2b The Assumptions of the Model

2. $\text{var}(y_i) = \text{var}(e_i) = \sigma^2$

- The variance of the probability distribution of y_i does not change with each observation. Some observations on y_i are not more likely to be further from the regression function than others.

5.1.2b The Assumptions of the Model

3. $$\operatorname{cov}(y_i, y_j) = \operatorname{cov}(e_i, e_j) = 0$$

- Any two observations on the dependent variable are uncorrelated. For example, if one observation is above $E(y_i)$, a subsequent observation is not more or less likely to be above $E(y_i)$.

5.1.2b The Assumptions of the Model

4. $y_i \sim N\left[(\beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}), \sigma^2\right]$

- We sometimes will assume that the values of y_i are normally distributed about their mean. This is equivalent to assuming that $e_i \sim N\left(0, \sigma^2\right)$.

5.1.2b The Assumptions of the Model

Assumptions of the Multiple Regression Model

MR1. $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i, \; i = 1, \ldots, N$

MR2. $E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \Leftrightarrow E(e_i) = 0$

MR3. $\operatorname{var}(y_i) = \operatorname{var}(e_i) = \sigma^2$

MR4. $\operatorname{cov}(y_i, y_j) = \operatorname{cov}(e_i, e_j) = 0$

MR5. The values of each x_{tk} are not random and are not exact linear functions of the other explanatory variables

MR6. $y_i \sim N\left[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2\right] \Leftrightarrow e_i \sim N(0, \sigma^2)$

5.2 Estimating the Parameters of the Multiple Regression Model

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e \qquad (5.4)$$

$$\begin{aligned} S(\beta_1, \beta_2, \beta_3) &= \sum_{i=1}^{N} \left(y_i - E(y_i) \right)^2 \\ &= \sum_{i=1}^{N} \left(y_i - \beta_1 - \beta_2 x_{i2} - \beta_3 x_{i3} \right)^2 \end{aligned} \qquad (5.5)$$

5.2.2 Least Squares Estimates Using Hamburger Chain Data

Table 5.2 Least Squares Estimates for Sales Equation for Big Andy's Burger Barn

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	118.9136	6.3516	18.7217	0.0000
P	−7.9079	1.0960	−7.2152	0.0000
A	1.8626	0.6832	2.7263	0.0080

$R^2 = 0.4483$ $SSE = 1718.943$ $\hat{\sigma} = 4.8861$ $\hat{\sigma}_y = 6.48854.$

5.2.2 Least Squares Estimates Using Hamburger Chain Data

$$E(y_i) = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3}$$

$$\hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3}$$

$$= 118.91 - 7.908 x_{i2} + 1.863 x_{i3}$$

$$\hat{S}_i = 118.91 - 7.908 P_i + 1.863 A_i \qquad (5.6)$$

$$\widehat{SALES} = 118.91 - 7.908\, PRICE + 1.863\, ADVERT$$

5.2.2 Least Squares Estimates Using Hamburger Chain Data

Suppose we are interested in predicting sales revenue for a price of $5.50 and an advertising expenditure of $1,200.

This prediction is given by

$$\hat{S} = 118.91 - 7.908\,PRICE + 1.863\,ADVERT$$

$$= 118.914 - 7.9079 \times 5.5 + 1.8626 \times 1.2$$

$$= 77.656$$

5.2.2 Least Squares Estimates Using Hamburger Chain Data

Remark: Estimated regression models describe the relationship between the economic variables for values *similar* to those found in the sample data. Extrapolating the results to extreme values is generally not a good idea. Predicting the value of the dependent variable for values of the explanatory variables far from the sample values invites disaster.

5.2.3 Estimation of the Error Variance σ^2

$$\sigma^2 = \mathrm{var}(e_i) = E(e_i^2)$$

$$\hat{e}_i = y_i - y_i = y_i - \left(b_1 + b_2 x_{i2} + b_3 x_{i3}\right)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} \hat{e}_i^2}{N - K} \qquad (5.7)$$

5.2.3 Estimation of the Error Variance σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{75} \hat{e}_i^2}{N-K} = \frac{1718.943}{75-3} = 23.874$$

$$SSE = \sum_{i=1}^{N} \hat{e}_i^2 = 1718.943$$

$$\hat{\sigma} = \sqrt{23.874} = 4.8861$$

5.3 Sampling Properties of the Least Squares Estimator

The Gauss-Markov Theorem: For the multiple regression model, if assumptions MR1-MR5 listed at the beginning of the Chapter hold, then the least squares estimators are the Best Linear Unbiased Estimators (BLUE) of the parameters.

5.3.1 The Variances and Covariances of the Least Squares Estimators

$$\operatorname{var}(b_2) = \frac{\sigma^2}{(1 - r_{23}^2)\sum_{i=1}^{N}(x_{i2} - \bar{x}_2)^2} \tag{5.8}$$

$$r_{23} = \frac{\sum (x_{i2} - \bar{x}_2)(x_{i3} - \bar{x}_3)}{\sqrt{\sum (x_{i2} - \bar{x}_2)^2 \sum (x_{i3} - \bar{x}_3)^2}} \tag{5.9}$$

5.3.1 The Variances and Covariances of the Least Squares Estimators

1. Larger error variances σ^2 lead to larger variances of the least squares estimators.

2. Larger sample sizes N imply smaller variances of the least squares estimators.

3. More variation in an explanatory variable around its mean, leads to a smaller variance of the least squares estimator.

4. A larger correlation between x_2 and x_3 leads to a larger variance of b_2.

5.3.1 The Variances and Covariances of the Least Squares Estimators

- The covariance matrix for K=3 is

$$\operatorname{cov}(b_1, b_2, b_3) = \begin{bmatrix} \operatorname{var}(b_1) & \operatorname{cov}(b_1, b_2) & \operatorname{cov}(b_1, b_3) \\ \operatorname{cov}(b_1, b_2) & \operatorname{var}(b_2) & \operatorname{cov}(b_2, b_3) \\ \operatorname{cov}(b_1, b_3) & \operatorname{cov}(b_2, b_3) & \operatorname{var}(b_3) \end{bmatrix}$$

- The estimated variances and covariances in the example are

$$\widehat{\operatorname{cov}}(b_1, b_2, b_3) = \begin{bmatrix} 40.343 & -6.795 & -.7484 \\ -6.795 & 1.201 & -.0197 \\ -.7484 & -.0197 & .4668 \end{bmatrix} \qquad (5.10)$$

5.3.1 The Variances and Covariances of the Least Squares Estimators

- Therefore, we have

$$\widehat{\operatorname{var}}\left(b_1\right) = 40.343 \qquad \widehat{\operatorname{cov}}\left(b_1, b_2\right) = -6.795$$

$$\widehat{\operatorname{var}}\left(b_2\right) = 1.201 \qquad \widehat{\operatorname{cov}}\left(b_1, b_3\right) = -.7484$$

$$\widehat{\operatorname{var}}\left(b_3\right) = .4668 \qquad \widehat{\operatorname{cov}}\left(b_2, b_3\right) = -.0197$$

5.3.1 The Variances and Covariances of the Least Squares Estimators

Table 5.3 Covariance Matrix for Coefficient Estimates

	C	P	A
C	40.3433	−6.7951	−0.7484
P	−6.7951	1.2012	−0.0197
A	−0.7484	−0.0197	0.4668

5.3.1 The Variances and Covariances of the Least Squares Estimators

- The standard errors are

$$\text{se}(b_1) = \sqrt{\widehat{\text{var}}(b_1)} = \sqrt{40.343} = 6.352$$

$$\text{se}(b_2) = \sqrt{\widehat{\text{var}}(b_2)} = \sqrt{1.201} = 1.096$$

$$\text{se}(b_3) = \sqrt{\widehat{\text{var}}(b_3)} = \sqrt{.4668} = .6832$$

5.3.2 The Properties of the Least Squares Estimators Assuming Normally Distributed Errors

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_K x_{iK} + e_i$$

$$y_i \sim N\left[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2\right] \Leftrightarrow e_i \sim N(0, \sigma^2)$$

$$b_k \sim N\left[\beta_k, \operatorname{var}\left(b_k\right)\right]$$

5.3.2 The Properties of the Least Squares Estimators Assuming Normally Distributed Errors

$$z = \frac{b_k - \beta_k}{\sqrt{\operatorname{var}\left(b_k\right)}} \sim N\left(0,1\right), \text{ for } k = 1, 2, \ldots, K \qquad (5.11)$$

$$t = \frac{b_k - \beta_k}{\sqrt{\widehat{\operatorname{var}}\left(b_k\right)}} = \frac{b_k - \beta_k}{\operatorname{se}(b_k)} \sim t_{(N-K)} \qquad (5.12)$$

5.4 Interval Estimation

$$P(-t_c < t_{(72)} < t_c) = .95 \tag{5.13}$$

$$P\left(-1.993 \le \frac{b_2 - \beta_2}{\text{se}(b_2)} \le 1.993\right) = .95 \tag{5.14}$$

$$P\left[b_2 - 1.993 \times \text{se}(b_2) \;\le\; \beta_2 \;\le\; b_2 + 1.993 \times \text{se}(b_2)\right] = .95$$

$$\left[b_2 - 1.993 \times \text{se}(b_2),\; b_2 + 1.993 \times \text{se}(b_2)\right] \tag{5.15}$$

5.4 Interval Estimation

- A 95% interval estimate for β_2 based on our sample is given by

$$(-10.092,\ -5.724)$$

- A 95% interval estimate for β_3 based on our sample is given by

$$(1.8626 - 1.993 \times .6832,\ 1.8626 + 1.993 \times .6832) = (.501, 3.224)$$

- The general expression for a $100(1-\alpha)\%$ confidence interval is

$$[b_k - t_{(1-\alpha/2,\, N-K)} \times \mathrm{se}(b_k),\ b_k + t_{(1-\alpha/2,\, N-K)} \times \mathrm{se}(b_k))$$

5.5 Hypothesis Testing for a Single Coefficient

STEP-BY-STEP PROCEDURE FOR TESTING HYPOTHESES

1. Determine the null and alternative hypotheses.
2. Specify the test statistic and its distribution if the null hypothesis is true.
3. Select α and determine the rejection region.
4. Calculate the sample value of the test statistic and, if desired, the *p*-value.
5. State your conclusion.

5.5.1 Testing the Significance of a Single Coefficient

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0$$

$$t = \frac{b_k}{\mathrm{se}(b_k)} \sim t_{(N-K)}$$

- For a test with level of significance α

$$t_c = t_{(1-\alpha/2,\, N-K)} \text{ and } -t_c = t_{(\alpha/2,\, N-K)}$$

5.5.1 Testing the Significance of a Single Coefficient

- **Big Andy's Burger Barn example**

1. The null and alternative hypotheses are: $H_0 : \beta_2 = 0$ and $H_1 : \beta_2 \neq 0$
2. The test statistic, if the null hypothesis is true, is $t = b_2 \big/ \mathrm{se}\left(b_2\right) \sim t_{(N-K)}$
3. Using a 5% significance level (α=.05), and 72 degrees of freedom, the critical values that lead to a probability of 0.025 in each tail of the distribution are

$$t_{(.975,\,72)} = 1.993 \text{ and } t_{(.025,\,72)} = -1.993$$

5.5.1 Testing the Significance of a Single Coefficient

4. The computed value of the t-statistic is $t = \dfrac{-7.908}{1.096} = -7.215$

   the p-value in this case can be found as

$$P\left(t_{(72)} > 7.215\right) + P\left(t_{(72)} < -7.215\right) = 2 \times (2.2 \times 10^{-10}) = .000$$

5. Since $-7.215 < -1.993$, we reject $H_0 : \beta_2 = 0$ and conclude that there is evidence from the data to suggest sales revenue depends on price. Using the p-value to perform the test, we reject H_0 because $.000 < .05$.

5.5.1 Testing the Significance of a Single Coefficient

- **Testing whether sales revenue is related to advertising expenditure**

1. $H_0 : \beta_3 = 0$ and $H_1 : \beta_3 \neq 0$

2. The test statistic, if the null hypothesis is true, is $t = b_3 / \mathrm{se}(b_3) \sim t_{(N-K)}$

3. Using a 5% significance level, we reject the null hypothesis if $t \geq 1.993$ or $t \leq -1.993$. In terms of the *p*-value, we reject H_0 if $p \leq .05$.

5.5.1 Testing the Significance of a Single Coefficient

- **Testing whether sales revenue is related to advertising expenditure**

4. The value of the test statistic is $t = \dfrac{1.8626}{.6832} = 2.726,$

   the *p*-value in given by $P\left(t_{(72)} > 2.726\right) + P\left(t_{(72)} < -2.726\right) = 2 \times .004 = .008$

5. Because $2.726 > 1.993,$ we reject the null hypothesis; the data support the conjecture that revenue is related to advertising expenditure. Using the *p*-value we reject H_0 because $.008 < .05.$

5.5.2 One-Tailed Hypothesis Testing for a Single Coefficient

- **5.5.2a Testing for elastic demand**

  We wish to know if

- $\beta_2 \geq 1$: a decrease in price leads to a decrease in sales revenue (demand is price inelastic), or
- $\beta_2 < 1$: a decrease in price leads to an increase in sales revenue (demand is price elastic)

5.5.2 One-Tailed Hypothesis Testing for a Single Coefficient

1. $H_0: \beta_2 \geq 0$ (demand is unit elastic or inelastic)

   $H_1: \beta_2 < 0$ (demand is elastic)

2. To create a test statistic we assume that $H_0: \beta_2 = 0$ is true and use

   $t = b_2 / \text{se}(b_2) \sim t_{(N-K)}$

3. At a 5% significance level, we reject H_0 if $t \leq -1.666$ or if the $p-\text{value} < .05$

5.5.2 One-Tailed Hypothesis Testing for a Single Coefficient

4. The value of the test statistic is $t = \dfrac{b_2}{\text{se}(b_2)} = \dfrac{-7.908}{1.096} = -7.215$

   The corresponding *p*-value is $P(t_{(72)} < -7.215) = .000$

5. Since $-7.215 < -1.666$ we reject $H_0 : \beta_2 \geq 0$. Since $.000 < .05$, the same conclusion is reached using the *p*-value.

5.5.2 One-Tailed Hypothesis Testing for a Single Coefficient

5.5.2b Testing Advertising Effectiveness

1. $H_0: \beta_3 \le 1$ and $H_1: \beta_3 > 1$

2. To create a test statistic we assume that $H_0: \ \beta_3 = 1$ is true and use

$$t = \frac{b_3 - 1}{\text{se}(b_3)} \sim t_{(N-K)}$$

3. At a 5% significance level, we reject H_0 if $t \ge -1.666$ or if the $p-$value $\le$ .05

5.5.2 One-Tailed Hypothesis Testing for a Single Coefficient

5.5.2b Testing Advertising Effectiveness

1. The value of the test statistic is $t = \dfrac{b_3 - \beta_3}{\text{se}\left(b_3\right)} = \dfrac{1.8626 - 1}{.6832} = 1.263$

   The corresponding *p*-value is $P(t_{(72)} > 1.263) = .105$

5. Since 1.263<1.666 we do not reject H_0. Since .105>.05, the same conclusion is reached using the *p*-value.

5.6 Measuring Goodness-of-Fit

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{N}\left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2}$$

$$= 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^{N}\hat{e}_i^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \qquad (5.16)$$

5.6 Measuring Goodness-of-Fit

$$\hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3} + \cdots + b_k x_{iK}$$

$$\hat{\sigma}_y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} = \sqrt{\frac{SST}{N-1}}$$

$$SST = (N-1)\hat{\sigma}_y^2$$

5.6 Measuring Goodness-of-Fit

- For Big Andy's Burger Barn we find that

$$SST = 74 \times 6.48854^2 = 3115.485$$

$$SSE = 1718.943$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \hat{e}_i^2}{\sum_{i=1}^{N} \left(y_i - \bar{y}\right)^2} = 1 - \frac{1718.943}{3115.485} = .448$$

5.6 Measuring Goodness-of-Fit

- An alternative measure of goodness-of-fit called the adjusted-R^2, is usually reported by regression programs and it is computed as

$$\bar{R}^2 = 1 - \frac{SSE/(N-K)}{SST/(N-1)}$$

5.6 Measuring Goodness-of-Fit

- If the model does not contain an intercept parameter, then the measure R^2 given in (5.16) is no longer appropriate. The reason it is no longer appropriate is that, without an intercept term in the model,

$$\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2 \neq \sum_{i=1}^{N}\left(\hat{y}_i-\bar{y}\right)^2+\sum_{i=1}^{N} e_i^2$$

$$SST \neq SSR + SSE$$

5.6.1 Reporting the Regression Results

$$\widehat{SALES} = 118.9 - 7.908\,PRICE + 1.8626\,ADVERT \qquad R^2 = .448 \qquad (5.17)$$

$$\text{(se)} \qquad (6.35) \quad (1.096) \qquad\qquad (.6832)$$

- From this summary we can read off the estimated effects of changes in the explanatory variables on the dependent variable and we can predict values of the dependent variable for given values of the explanatory variables. For the construction of an interval estimate we need the least squares estimate, its standard error, and a critical value from the *t*-distribution.

Keywords

- BLU estimator
- covariance matrix of least squares estimator
- critical value
- error variance estimate
- error variance estimator
- goodness of fit
- interval estimate
- least squares estimates
- least squares estimation
- least squares estimators
- multiple regression model
- one-tailed test
- *p*-value
- regression coefficients
- standard errors
- sum of squared errors
- sum of squares of regression
- testing significance
- total sum of squares
- two-tailed test

Chapter 5 Appendices

- Appendix 5A Derivation of the least squares estimators

Appendix 5A
Derivation of the least squares estimators

$$S(\beta_1,\beta_2,\beta_3)=\sum_{i=1}^{N}(y_i-\beta_1-\beta_2x_{i2}-\beta_3x_{i3})^2$$

$$\frac{\partial S}{\partial\beta_1}=2N\beta_1+2\beta_2\sum x_{i2}+2\beta_3\sum x_{i3}-2\sum y_i$$

$$\frac{\partial S}{\partial\beta_2}=2\beta_1\sum x_{i2}+2\beta_2\sum x_{i2}^2+2\beta_3\sum x_{i2}x_{i3}-2\sum x_{i2}y_i$$

$$\frac{\partial S}{\partial\beta_3}=2\beta_1\sum x_{i3}+2\beta_2\sum x_{i2}x_{i3}+2\beta_3\sum x_{i3}^2-2\sum x_{i3}y_i$$

Appendix 5A
Derivation of the least squares estimators

$$Nb_1 + \sum x_{i2}b_2 + \sum x_{i3}b_3 = \sum y_i$$

$$\sum x_{i2}b_1 + \sum x_{i2}^2 b_2 + \sum x_{i2}x_{i3}b_3 = \sum x_{i2}y_i \qquad (5A.1)$$

$$\sum x_{i3}b_1 + \sum x_{i2}x_{i3}b_2 + \sum x_{i3}^2 b_3 = \sum x_{i3}y_i$$

$$\text{let } y_i^* = y_i - \bar{y}, \quad x_{i2}^* = x_{i2} - \bar{x}_2, \quad x_{i3}^* = x_{i3} - \bar{x}_3$$

Appendix 5A
Derivation of the least squares estimators

$$b_1 = \bar{y} - b_2\bar{x}_2 - b_3\bar{x}_3$$

$$b_2 = \frac{\left(\sum y_i^* x_{i2}^*\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum y_i^* x_{i3}^*\right)\left(\sum x_{i2}^* x_{i3}^*\right)}{\left(\sum x_{i2}^{*2}\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum x_{i2}^* x_{i3}^*\right)^2}$$

$$b_3 = \frac{\left(\sum y_i^* x_{i3}^*\right)\left(\sum x_{i2}^{*2}\right) - \left(\sum y_i^* x_{i2}^*\right)\left(\sum x_{i3}^* x_{i2}^*\right)}{\left(\sum x_{i2}^{*2}\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum x_{i2}^* x_{i3}^*\right)^2}$$