CHAPTER 5

Exercise Solutions

EXERCISE 5.1

(a) $\bar{y} = 1, \quad \bar{x}_2 = 0, \quad \bar{x}_3 = 0$

x_{i2}^*	x_{i3}^*	y_i^*
0	1	0
1	−2	1
2	1	2
−2	0	−2
1	−1	−1
−2	−1	−2
0	1	1
−1	1	0
1	0	1

(b) $\sum y_i^* x_{i2}^* = 13, \qquad \sum x_{i2}^{*2} = 16, \qquad \sum y_i^* x_{i3}^* = 4, \qquad \sum x_{i3}^{*2} = 10$

(c)
$$b_2 = \frac{\left(\sum y_i^* x_{i2}^*\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum y_i^* x_{i3}^*\right)\left(\sum x_{i2}^* x_{i3}^*\right)}{\left(\sum x_{i2}^{*2}\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum x_{i2}^* x_{i3}^*\right)^2} = \frac{13 \times 10 - 4 \times 0}{16 \times 10 - 0^2} = 0.8125$$

$$b_3 = \frac{\left(\sum y_i^* x_{i3}^*\right)\left(\sum x_{i2}^{*2}\right) - \left(\sum y_i^* x_{i2}^*\right)\left(\sum x_{i2}^* x_{i3}^*\right)}{\left(\sum x_{i2}^{*2}\right)\left(\sum x_{i3}^{*2}\right) - \left(\sum x_{i2}^* x_{i3}^*\right)^2} = \frac{4 \times 16 - 13 \times 0}{16 \times 10 - 0^2} = 0.4$$

$$b_1 = \bar{y} - b_2 \bar{x}_2 - b_3 \bar{x}_3 = 1$$

(d) $\hat{e} = \left(-0.4,\ 0.9875,\ -0.025,\ -0.375,\ -1.4125,\ 0.025,\ 0.6,\ 0.4125,\ 0.1875\right)$

(e)
$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N - K} = \frac{3.8375}{9 - 3} = 0.6396$$

(f)
$$r_{23} = \frac{\sum (x_{i2} - \bar{x}_2)(x_{i3} - \bar{x}_3)}{\sqrt{\sum (x_{i2} - \bar{x}_2)^2 \sum (x_{i3} - \bar{x}_3)^2}} = \frac{\sum x_{i2}^* x_{i3}^*}{\sqrt{\sum x_{i2}^{*\,2} \sum x_{i3}^{*\,2}}} = 0$$

(g)
$$\text{se}(b_2) = \sqrt{\widehat{\text{var}(b_2)}} = \sqrt{\frac{\hat{\sigma}^2}{\sum (x_{i2} - \bar{x}_2)^2 (1 - r_{23}^2)}} = \sqrt{\frac{0.6396}{16}} = 0.1999$$

(h) $SSE = \sum \hat{e}_i^2 = 3.8375 \qquad SST = \sum (y_i - \bar{y})^2 = 16,$

$$SSR = SST - SSE = 12.1625 \qquad R^2 = \frac{SSR}{SST} = \frac{12.1625}{16} = 0.7602$$

EXERCISE 5.2

(a) A 95% confidence interval for β_2 is

$$b_2 \pm t_{(0.975,6)}\,\text{se}(b_2) = 0.8125 \pm 2.447 \times 0.1999 = (0.3233, 1.3017)$$

(b) The null and alternative hypotheses are

$$H_0 : \beta_2 = 1, \quad H_1 : \beta_2 \neq 1$$

The calculated t-value is

$$t = \frac{b_2 - 1}{\text{se}(b_2)} = \frac{0.8125 - 1}{0.1999} = -0.9377$$

At a 5% significance level, we reject H_0 if $|t| > t_{(0.975,\,6)} = 2.447$. Since $|-0.9377| < 2.447$, we do not reject H_0.

EXERCISE 5.3

(a) (i) The t-statistic for b_1 is $\dfrac{b_1}{\text{se}(b_1)} = \dfrac{0.0091}{0.0191} = 0.476$.

(ii) The standard error for b_2 is $\text{se}(b_2) = \dfrac{0.0276}{6.6086} = 0.00418$.

(iii) The estimate for β_3 is $b_3 = 0.0002 \times (-6.9624) = -0.0014$.

(iv) To compute R^2, we need *SSE* and *SST*. From the output, $SSE = 5.752896$. To find *SST*, we use the result

$$\hat{\sigma}_y = \sqrt{\frac{SST}{N-1}} = 0.0633$$

which gives $SST = 1518 \times (0.0633)^2 = 6.08246$.

Thus, $R^2 = 1 - \dfrac{SSE}{SST} = 1 - \dfrac{5.75290}{6.08246} = 0.054$

(v) The estimated error standard deviation is $\hat{\sigma} = \sqrt{\dfrac{SSE}{(N-K)}} = \sqrt{\dfrac{5.752896}{1519-4}} = 0.061622$

(b) The value $b_2 = 0.0276$ implies that if $\ln(TOTEXP)$ increases by 1 unit the alcohol share will increase by 0.0276. The change in the alcohol share from a 1-unit change in total expenditure depends on the level of total expenditure. Specifically, $d(WALC)/d(TOTEXP) = 0.0276/TOTEXP$. A 1% increase in total expenditure leads to a 0.000276 increase in the alcohol share of expenditure.

The value $b_3 = -0.0014$ suggests that if the age of the household head increases by 1 year the share of alcohol expenditure of that household decreases by 0.0014.

The value $b_4 = -0.0133$ suggests that if the household has one more child the share of the alcohol expenditure decreases by 0.0133.

(c) A 95% confidence interval for β_3 is

$$b_3 \pm t_{0.975,1515)}\,\text{se}(b_3) = -0.0014 \pm 1.96 \times 0.0002 = (-0.0018, -0.0010)$$

This interval tells us that, if the age of the household head increases by 1 year, the share of the alcohol expenditure is estimated to decrease by an amount between 0.0018 and 0.001.

Exercise 5.3 (Continued)

(d) The null and alternative hypotheses are $H_0: \beta_4 = 0,\ H_1: \beta_4 \neq 0$.

The calculated *t*-value is $t = \dfrac{b_4}{\text{se}(b_4)} = -4.075$

At a 5% significance level, we reject H_0 if $|t| > t_{(0.975,\,1515)} = 1.96$. Since $|-4.075| > 1.96$, we reject H_0 and conclude that the number of children in the household influences the budget proportion on alcohol. Having an additional child is likely to lead to a smaller budget share for alcohol because of the non-alcohol expenditure demands of that child. Also, perhaps households with more children prefer to drink less, believing that drinking may be a bad example for their children.

EXERCISE 5.4

(a) The regression results are:

$$\widehat{WTRANS} = -0.0315 + 0.0414\ln\left(TOTEXP\right) - 0.0001AGE - 0.0130\,NK \qquad R^2 = 0.0247$$
$$\text{(se)} \qquad (0.0322)\ (0.0071) \qquad\qquad (0.0004) \qquad (0.0055)$$

(b) The value $b_2 = 0.0414$ suggests that as $\ln\left(TOTEXP\right)$ increases by 1 unit the budget proportion for transport increases by 0.0414. Alternatively, one can say that a 10% increase in total expenditure will increase the budget proportion for transportation by 0.004. (See Section A.4.8 of Appendix A.) The positive sign of b_2 is according to our expectation because as households become richer they tend to use more luxurious forms of transport and the proportion of the budget for transport increases.

The value $b_3 = -0.0001$ implies that as the age of the head of the household increases by 1 year the budget share for transport decreases by 0.0001. The expected sign for b_3 is not clear. For a given level of total expenditure and a given number of children, it is difficult to predict the effect of age on transport share.

The value $b_4 = -0.0130$ implies that an additional child decreases the budget share for transport by 0.013. The negative sign means that adding children to a household increases expenditure on other items (such as food and clothing) more than it does on transportation. Alternatively, having more children may lead a household to turn to cheaper forms of transport.

(c) The *p*-value for testing $H_0 : \beta_3 = 0$ against the alternative $H_1 : \beta_3 \neq 0$ where β_3 is the coefficient of *AGE* is 0.869, suggesting that *AGE* could be excluded from the equation. Similar tests for the coefficients of the other two variables yield *p*-values less than 0.05.

(d) The proportion of variation in the budget proportion allocated to transport explained by this equation is 0.0247.

(e) For a one-child household:

$$\begin{aligned}\widehat{WTRANS}_0 &= -0.0315 + 0.0414\ln(TOTEXP_0) - 0.0001AGE_0 - 0.013NK_0 \\ &= -0.0315 + 0.0414 \times \ln(98.7) - 0.0001 \times 36 - 0.013 \times 1 \\ &= 0.1420\end{aligned}$$

For a two-child household:

$$\begin{aligned}\widehat{WTRANS}_0 &= -0.0315 + 0.0414\ln(TOTEXP_0) - 0.0001AGE_0 - 0.013NK_0 \\ &= -0.0315 + 0.0414 \times \ln(98.7) - 0.0001 \times 36 - 0.013 \times 2 \\ &= 0.1290\end{aligned}$$

EXERCISE 5.5

(a) The estimated equation is

$$\begin{array}{llllll} \widehat{VALUE} = & 28.4067 - & 0.1834CRIME - & 22.8109NITOX + & 6.3715ROOMS - & 0.0478AGE \\ \text{(se)} & (5.3659) & (0.0365) & (4.1607) & (0.3924) & (0.0141) \end{array}$$

$$\begin{array}{llll} -1.3353DIST + & 0.2723ACCESS - & 0.0126TAX - & 1.1768PTRATIO \\ (0.2001) & (0.0723) & (0.0038) & (0.1394) \end{array}$$

The estimated equation suggests that as the per capita crime rate increases by 1 unit the home value decreases by \$183.4. The higher the level of air pollution the lower the value of the home; a one unit increase in the nitric oxide concentration leads to a decline in value of \$22,811. Increasing the average number of rooms leads to an increase in the home value; an increase in one room leads to an increase of \$6,372. An increase in the proportion of owner-occupied units built prior to 1940 leads to a decline in the home value. The further the weighted distances to the five Boston employment centers the lower the home value by \$1,335 for every unit of weighted distance. The higher the tax rate per \$10,000 the lower the home value. Finally, the higher the pupil-teacher ratio, the lower the home value.

(b) A 95% confidence interval for the coefficient of *CRIME* is

$$b_2 \pm t_{(0.975,497)}\text{se}(b_2) = -0.1834 \pm 1.965 \times 0.0365 = (-0.255, -0.112).$$

A 95% confidence interval for the coefficient of *ACCESS* is

$$b_7 \pm t_{(0.975,497)}\text{se}(b_7) = 0.2723 \pm 1.965 \times 0.0723 = (0.130, 0.414)$$

(c) We want to test $H_0 : \beta_{room} = 7$ against $H_1 : \beta_{room} \neq 7$. The value of the *t* statistic is

$$t = \frac{b_{rooms} - 7}{\text{se}(b_{rooms})} = \frac{6.3715 - 7}{0.3924} = -1.6017$$

At $\alpha = 0.05$, we reject H_0 if the absolute calculated *t* is greater than 1.965. Since $|-1.6017| < 1.965$, we do not reject H_0. The data is consistent with the hypothesis that increasing the number of rooms by one increases the value of a house by \$7000.

(d) We want to test $H_0 : \beta_{ptratio} \geq -1$ against $H_1 : \beta_{ptratio} < -1$. The value of the *t* statistic is

$$t = \frac{-1.1768 + 1}{0.1394} = -1.2683$$

At a significance level of $\alpha = 0.05$, we reject H_0 if the calculated *t* is less than the critical value $t_{(0.05,497)} = -1.648$. Since $-1.2683 > -1.648$, we do not reject H_0. We cannot conclude that reducing the pupil-teacher ratio by 10 will increase the value of a house by more than \$10,000.

EXERCISE 5.6

The EViews output for verifying the answers to Exercise 5.1 is given below.

Method: Least Squares
Dependent Variable: Y
Method: Least Squares
Included observations: 9

	Coefficient	Std. Error	t-Statistic	Prob.
X1	1.000000	0.266580	3.751221	0.0095
X2	0.812500	0.199935	4.063823	0.0066
X3	0.400000	0.252900	1.581654	0.1648

R-squared	0.760156	Mean dependent var	1.000000
Adjusted R-squared	0.680208	S.D. dependent var	1.414214
S.E. of regression	0.799740	Akaike info criterion	2.652140
Sum squared resid	3.837500	Schwarz criterion	2.717882
Log likelihood	-8.934631	Hannan-Quinn criter.	1.728217

EXERCISE 5.7

(a) Estimates, standard errors and *p*-values for each of the coefficients in each of the estimated share equations are given in the following table.

Explanatory Variables		Dependent Variable					
		Food	Fuel	Clothing	Alcohol	Transport	Other
Constant	Estimate	0.8798	0.3179	−0.2816	0.0149	−0.0191	0.0881
	Std Error	0.0512	0.0265	0.0510	0.0370	0.0572	0.0536
	p-value	0.0000	0.0000	0.0000	0.6878	0.7382	0.1006
ln(*TOTEXP*)	Estimate	−0.1477	−0.0560	0.0929	0.0327	0.0321	0.0459
	Std Error	0.0113	0.0058	0.0112	0.0082	0.0126	0.0118
	p-value	0.0000	0.0000	0.0000	0.0001	0.0111	0.0001
AGE	Estimate	0.00227	0.00044	−0.00056	−0.00220	0.00077	−0.00071
	Std Error	0.00055	0.00029	0.00055	0.00040	0.00062	0.00058
	p-value	0.0000	0.1245	0.3062	0.0000	0.2167	0.2242
NK	Estimate	0.0397	0.0062	−0.0048	−0.0148	−0.0123	−0.0139
	Std Error	0.0084	0.0044	0.0084	0.0061	0.0094	0.0088
	p-value	0.0000	0.1587	0.5658	0.0152	0.1921	0.1157

An increase in total expenditure leads to decreases in the budget shares allocated to food and fuel and increases in the budget shares of the commodity groups clothing, alcohol, transport and other. Households with an older household head devote a higher proportion of their budget to food, fuel and transport and a lower proportion to clothing, alcohol and other. Having more children means a higher proportion spent on food and fuel and lower proportions spent on the other commodities.

The coefficients of $\ln(TOTEXP)$ are significantly different from zero for all commodity groups. At a 5% significance level, age has a significant effect on the shares of food and alcohol, but its impact on the other budget shares is measured less precisely. Significance tests for the coefficients of the number of children yield a similar result. NK has an impact on the food and alcohol shares, but we can be less certain about the effect on the other groups. To summarize, $\ln(TOTEXP)$ has a clear impact in all equations, but the effect of *AGE* and *NK* is only significant in the food and alcohol equations.

Exercise 5.7 (continued)

(b) The *t*-values and *p*-values for testing $H_0:\beta_2 \le 0$ against $H_1:\beta_2 > 0$ are reported in the table below. Using a 5% level of significance, the critical value for each test is $t_{(0.95,496)} = 1.648$.

	t-value	*p*-value	decision
WFOOD	−13.083	1.0000	Do not reject H_0
WFUEL	−9.569	1.0000	Do not reject H_0
WCLOTH	8.266	0.0000	Reject H_0
WALC	4.012	0.0000	Reject H_0
WTRANS	2.548	0.0056	Reject H_0
WOTHER	3.884	0.0001	Reject H_0

Those commodities which are regarded as necessities ($b_2 < 0$) are food and fuel. The tests suggest the rest are luxuries. While alcohol, transportation and other might be luxuries, it is difficult to see clothing categorized as a luxury. Perhaps a finer classification is necessary to distinguish between basic and luxury clothing.

EXERCISE 5.8

(a) The expected sign for β_2 is negative because, as the number of grams in a given sale increases, the price per gram should decrease, implying a discount for larger sales. We expect β_3 to be positive; the purer the cocaine, the higher the price. The sign for β_4 will depend on how demand and supply are changing over time. For example, a fixed demand and an increasing supply will lead to a fall in price. A fixed supply and increased demand would lead to a rise in price.

(b) The estimated equation is:

$$\widehat{PRICE} = 90.8467 - 0.0600\,QUANT + 0.1162\,QUAL - 2.3546\,TREND \qquad R^2 = 0.5097$$
$$\begin{array}{lllll} \text{(se)} & (8.5803) & (0.0102) & (0.2033) & (1.3861) \\ (t) & (10.588) & (-5.892) & (0.5717) & (-1.6987) \end{array}$$

The estimated values for β_2, β_3 and β_4 are -0.0600, 0.1162 and -2.3546, respectively. They imply that as quantity (number of grams in one sale) increases by 1 unit, the price will go down by 0.0600. Also, as the quality increases by 1 unit the price goes up by 0.1162. As time increases by 1 year, the price decreases by 2.3546. All the signs turn out according to our expectations, with β_4 implying supply has been increasing faster than demand.

(c) The proportion of variation in cocaine price explained by the variation in quantity, quality and time is 0.5097.

(d) For this hypothesis we test $H_0 : \beta_2 \geq 0$ against $H_1 : \beta_2 < 0$. The calculated *t*-value is -5.892. We reject H_0 if the calculated *t* is less than the critical $t_{(0.95,52)} = -1.675$. Since the calculated *t* is less than the critical *t* value, we reject H_0 and conclude that sellers are willing to accept a lower price if they can make sales in larger quantities.

(e) We want to test $H_0 : \beta_3 \leq 0$ against $H_1 : \beta_3 > 0$. The calculated *t*-value is 0.5717. At $\alpha = 0.05$ we reject H_0 if the calculated *t* is greater than 1.675. Since for this case, the calculated *t* is not greater than the critical *t*, we do not reject H_0. We cannot conclude that a premium is paid for better quality cocaine.

(f) The average annual change in the cocaine price is given by the value of $b_4 = -2.3546$. It has a negative sign suggesting that the price decreases over time. A possible reason for a decreasing price is the development of improved technology for producing cocaine, such that suppliers can produce more at the same cost.

EXERCISE 5.9

(a) The coefficients, β_3, β_4 and β_5 are expected to be positive, whereas β_2 should be negative. The dependent variable is the log of per capita consumption of beef. As the price of beef declines, consumption of beef should increase ($\beta_2 < 0$). Also, as the prices of lamb and pork increase, the consumption of beef should increase ($\beta_3 > 0$ and $\beta_4 > 0$). If per capita disposable income increases, the consumption of beef should increase ($\beta_5 > 0$).

(b) The least squares estimates and standard errors are given by

$$\begin{aligned}\widehat{\ln(QB)} &= \underset{(1.6596)}{4.6726} - \underset{(0.1826)}{0.8266}\ln(PB) + \underset{(0.2127)}{0.1997}\ln(PL) \qquad (\text{se}) \\ &\quad + \underset{(0.3837)}{0.4371}\ln(PP) + \underset{(0.2940)}{0.1017}\ln(IN) \qquad R^2 = 0.7609\end{aligned}$$

The estimated coefficients are elasticities, implying that a 1% increase in the price of beef leads to an 0.83% decrease in the quantity of beef consumed. Likewise, 1% changes in the prices of lamb and pork will cause 0.20% and 0.44% changes, respectively, in the quantity of beef consumed. The estimates seem reasonable. They all have the expected signs and the price of beef has the greatest effect on the quantity of beef consumed, a logical result. However, the standard errors of the coefficient estimates for the lamb and pork prices, and for income, are relatively large.

(c) The estimated covariance matrix is given by

$$\widehat{\text{cov}}(b_1,b_2,b_3,b_4,b_5) = \begin{pmatrix} 2.7542 & 0.1087 & -0.0736 & -0.2353 & -0.3314 \\ 0.1087 & 0.0334 & 0.0040 & -0.0107 & -0.0326 \\ -0.0736 & 0.0040 & 0.0453 & -0.0439 & 0.1661 \\ -0.2353 & -0.0107 & -0.0439 & 0.1472 & -0.0303 \\ -0.3314 & -0.0326 & 0.1661 & -0.0303 & 0.0864 \end{pmatrix}$$

and the standard errors, reported in part (b), are the square roots of the diagonal of this matrix. The matrix contains estimates of the variances and covariances of the least squares estimators b_1, b_2, b_3, b_4 and b_5. The variances and covariances are measures of how the least squares estimates will vary and "covary" in repeated samples.

(d) Using $t_{(0.975,12)} = 2.179$, the 95% confidence intervals are as follows.

	β_1	β_2	β_3	β_4	β_5
lower limit	1.0563	–1.2246	–0.2638	–0.3989	–0.5389
upper limit	8.2888	–0.4286	0.6632	1.2732	0.7423

EXERCISE 5.10

(a) The estimated regression models and 95% confidence intervals for β_3 follow.

(i) All houses:

$$\widehat{PRICE} = -41948 + 90.970SQFT - 755.04AGE$$
$$\text{(se)} \quad (6990) \quad (2.403) \quad (140.89)$$

$$b_3 \pm t_{(0.975,1077)}\text{se}(b_3) = -755.04 \pm 1.962 \times 140.89 = (-1031.5, -478.6)$$

(ii) Town houses:

$$\widehat{PRICE} = 90415 + 44.014SQFT - 2621.6AGE$$
$$\text{(se)} \quad (11437) \quad (5.688) \quad (375.4)$$

$$b_3 \pm t_{(0.975,67)}\text{se}(b_3) = -2621.6 \pm 1.996 \times 375.4 = (-3370.8, -1872.3)$$

(iii) French style homes:

$$\widehat{PRICE} = -293804 + 184.21SQFT - 64.713AGE$$
$$\text{(se)} \quad (32988) \quad (10.43) \quad (3473.323)$$

$$b_3 \pm t_{(0.975,94)}\text{se}(b_3) = -64.7 \pm 1.986 \times 3473.3 = (-6961.1, 6831.7)$$

The value of additional square feet is highest for French style homes and lowest for town houses. Town houses depreciate more with age than do French style homes. The larger number of observations used in the regression for all houses has led to standard errors that are much lower than those for the special-category houses. In terms of the confidence intervals for β_3, it also means the confidence interval for all houses is much narrower than those for the special-category houses. The confidence interval for French style homes is very wide and includes both positive and negative ranges. Age does not seem to be an important determinant of the price of French style homes. For town houses it is more important than for the total population of houses.

(b) We wish to test the hypothesis $H_0: \beta_3 = -1000$ against the alternative $H_0: \beta_3 \neq -1000$. The results for each of the three cases follow.

(i) All houses: The critical values are $t_{(0.975,1077)} = 1.962$ and $t_{(0.025,1077)} = -1.962$. We reject H_0 if the calculated *t*-value is such that $t \geq 1.962$ or $t \leq -1.962$. The calculated value is

$$t = \frac{-755.04 - (-1000)}{140.89} = 1.74$$

Since $-1.962 < 1.74 < 1.962$, we do not reject H_0. The data is consistent with the hypothesis that having an older house reduces its price by \$1000 per year for each year of its age.

Exercise 5.10(b) (continued)

(b) (ii) Town houses: The critical values are $t_{(0.975,67)} = 1.996$ and $t_{(0.025,67)} = -1.996$. We reject H_0 if the calculated *t*-value is such that $t \geq 1.996$ or $t \leq -1.996$. The calculated value is

$$t = \frac{-2621.6 - (-1000)}{375.4} = -4.32$$

Since $t = -4.32 < -1.996$, we reject H_0. The data is not consistent with the hypothesis that having an older house reduces its price by \$1000 per year for each year of its age.

(iii) French style homes: The critical values are $t_{(0.975,94)} = 1.986$ and $t_{(0.025,94)} = -1.986$. We reject H_0 if the calculated *t*-value is such that $t \geq 1.986$ or $t \leq -1.986$. The calculated value is

$$t = \frac{-64.713 - (-1000)}{3473.3} = 0.27$$

Since $-1.986 < 0.27 < 1.986$, we do not reject H_0. The data is consistent with the hypothesis that having an older house reduces its price by \$1000 per year for each year of its age.

EXERCISE 5.11

(a) The estimated regression model is

$$\widehat{VOTE} = 52.44 + 0.6488\,GROWTH - 0.1862\,INFLATION$$
$$\text{(se)} \quad (1.49) \quad (0.1675) \qquad\qquad (0.4320)$$

The hypothesis test results on the significance of the coefficients are:

$H_0:\beta_2 = 0 \quad H_1:\beta_2 > 0$ $\quad$ *p*-value = 0.0003 $\quad$ significant at 10% level

$H_0:\beta_3 = 0 \quad H_1:\beta_3 < 0$ $\quad$ *p*-value = 0.335 $\quad$ not significant at 10% level

One-tail tests were used because more growth is considered favorable, and more inflation is considered not favorable, for re-election of the incumbent party.

(b) The predicted percentage vote for the incumbent party when $INFLATION = 4$ and $GROWTH = -4$ is

$$\widehat{VOTE_0} = 52.44 + 0.6488\times(-4) - 0.1862\times 4 = 49.104$$

(c) Ignoring the error term, the incumbent party will get the majority of the vote when

$$\beta_1 + \beta_2 GROWTH + \beta_3 INFLATION > 50$$

Assuming $\beta_1 = b_1$ and $\beta_3 = b_3$, and given $INFLATION = 4$ and $GROWTH = -4$, the incumbent party will get the majority of the vote when

$$52.44 - 4\beta_2 - 0.1862\times 4 \geq 50$$

which is equivalent to

$$\beta_2 < (-50 - 0.186221\times 4 + 52.44357)/4 = 0.42467$$

To ensure accuracy, we have included more decimal places in this calculation than are in the reported equation.

For testing $H_0:\beta_2 \leq 0.42467$ against the alternative $H_1:\beta_2 > 0.42467$, the *t*-value is

$$t = \frac{0.64876 - 0.42467}{0.16746} = 1.338$$

For a 5% significance level, the critical value is $t_{(0.95,\,28)} = 1.701$. The rejection region is $t \geq 1.701$. Thus, H_0 is not rejected. The incumbent party might still get elected if $GROWTH = -4\%$.

EXERCISE 5.12

(a) The estimated equation is

$$\widehat{TIME} = 19.9166 + 0.36923 DEPART + 1.3353 REDS + 2.7548 TRAINS$$
$$\text{(se)} \quad (1.2548) \ (0.01553) \qquad (0.1390) \qquad (0.3038)$$

Interpretations of each of the coefficients are:

β_1: The estimated time it takes Bill to get to work when he leaves Carnegie at 6:30AM and encounters no red lights and no trains is 19.92 minutes.

β_2: If Bill leaves later than 6:30AM, his traveling time increases by 3.7 minutes for every 10 minutes that his departure time is later than 6:30AM (assuming the number of red lights and trains are constant).

β_3: Each red light increases traveling time by 1.34 minutes.

β_4: Each train increases traveling time by 2.75 minutes.

(b) The 95% confidence intervals for the coefficients are:

β_1: $b_1 \pm t_{(0.975,227)}\text{se}(b_1) = 19.9166 \pm 1.970 \times 1.2548 = (17.44, 22.39)$

β_2: $b_2 \pm t_{(0.975,227)}\text{se}(b_2) = 0.36923 \pm 1.970 \times 0.01553 = (0.339, 0.400)$

β_3: $b_3 \pm t_{(0.975,227)}\text{se}(b_3) = 1.3353 \pm 1.970 \times 0.1390 = (1.06, 1.61)$

β_4: $b_4 \pm t_{(0.975,227)}\text{se}(b_4) = 2.7548 \pm 1.970 \times 0.3038 = (2.16, 3.35)$

In the context of driving time, these intervals are relatively narrow ones. We have obtained precise estimates of each of the coefficients.

(c) The hypotheses are $H_0:\beta_3 \ge 2$ and $H_1:\beta_3 < 2$. The critical value is $t_{(0.05,227)} = -1.652$. We reject H_0 when the calculated *t*-value is less than -1.652. This *t*-value is

$$t = \frac{1.3353 - 2}{0.1390} = -4.78$$

Since $-4.78 < -1.652$, we reject H_0. We conclude that the delay from each red light is less than 2 minutes.

(d) The hypotheses are $H_0:\beta_4 = 3$ and $H_1:\beta_4 \ne 3$. The critical values are $t_{(0.05,227)} = -1.652$ and $t_{(0.95,227)} = 1.652$. We reject H_0 when the calculated *t*-value is such that $t < -1.652$ or $t > 1.652$. This *t*-value is

$$t = \frac{2.7548 - 3}{0.3038} = 0.807$$

Since $-1.652 < 0.807 < 1.652$, we do not reject H_0. The data are consistent with the hypothesis that each train delays Bill by 3 minutes.

Exercise 5.12 (continued)

(e) Delaying the departure time by 30 minutes, increases travel time by $30\beta_2$. Thus, the null hypothesis is $H_0: 30\beta_2 \geq 10$, or $H_0: \beta_2 \geq 1/3$, and the alternative is $H_1: \beta_2 < 1/3$. We reject H_0 if $t \leq t_{(0.05,227)} = -1.652$, where the calculated *t*-value is

$$t = \frac{0.36923 - 0.33333}{0.01553} = 2.31$$

Since $2.31 > -1.652$, we do not reject H_0. The data are consistent with the hypothesis that delaying departure time by 30 minutes increases travel time by at least 10 minutes.

(f) If we assume that β_2, β_3 and β_4 are all non-negative, then the minimum time it takes Bill to travel to work is β_1. Thus, the hypotheses are $H_0: \beta_1 \leq 20$ and $H_1: \beta_1 > 20$. We reject H_0 if $t \geq t_{(0.95,227)} = 1.652$, where the calculated *t*-value is

$$t = \frac{19.9166 - 20}{1.2548} = -0.066$$

Since $-0.066 < 1.652$, we do not reject H_0. The data support the null hypothesis that the minimum travel time is less than or equal to 20 minutes. It was necessary to assume that β_2, β_3 and β_4 are all positive or zero, otherwise increasing one of the other variables will lower the travel time and the hypothesis would need to be framed in terms of more coefficients than β_1.

EXERCISE 5.13

(a) The coefficient estimates, standard errors, *t*-values and *p*-values are in the following table.

Dependent Variable: ln(*PROD*)

	Coeff	Std. Error	*t*-value	*p*-value
C	-1.5468	0.2557	-6.0503	0.0000
ln(*AREA*)	0.3617	0.0640	5.6550	0.0000
ln(*LABOR*)	0.4328	0.0669	6.4718	0.0000
ln(*FERT*)	0.2095	0.0383	5.4750	0.0000

All estimates have elasticity interpretations. For example, a 1% increase in labor will lead to a 0.4328% increase in rice output. A 1% increase in fertilizer will lead to a 0.2095% increase in rice output. All *p*-values are less than 0.0001 implying all estimates are significantly different from zero at conventional significance levels.

(b) The null and alternative hypotheses are $H_0:\beta_2 = 0.5$ and $H_1:\beta_2 \neq 0.5$. The 1% critical values are $t_{(0.995,348)} = 2.59$ and $t_{(0.005,348)} = -2.59$. Thus, the rejection region is $t \geq 2.59$ or $t \leq -2.59$. The calculated value of the test statistic is

$$t = \frac{0.3617 - 0.5}{0.064} = -2.16$$

Since $-2.59 < -2.16 < 2.59$, we do not reject H_0. The data are compatible with the hypothesis that the elasticity of production with respect to land is 0.5.

(c) A 95% interval estimate of the elasticity of production with respect to fertilizer is given by

$$b_4 \pm t_{(0.975,348)} \times \text{se}(b_4) = 0.2095 \pm 1.967 \times 0.03826 = (0.134,\ 0.285)$$

This relatively narrow interval implies the fertilizer elasticity has been precisely measured.

(d) This hypothesis test is a test of $H_0:\beta_3 \leq 0.3$ against $H_1:\beta_3 > 0.3$. The rejection region is $t \geq t_{(0.95,348)} = 1.649$. The calculated value of the test statistic is

$$t = \frac{0.433 - 0.3}{0.067} = 1.99$$

We reject H_0 because $1.99 > 1.649$. There is evidence to conclude that the elasticity of production with respect to labor is greater than 0.3. Reversing the hypotheses and testing $H_0:\beta_3 \geq 0.3$ against $H_1:\beta_3 < 0.3$, leads to a rejection region of $t \leq -1.649$. The calculated *t*-value is $t = 1.99$. The null hypothesis is not rejected because $1.99 > -1.649$.

EXERCISE 5.14

(a) The predicted logarithm of rice production for $AREA = 1$, $LABOR = 50$, and $FERT = 100$ is

$$\widehat{\ln(PROD)} = -1.5468 + 0.3617 \times \ln(1) + 0.4328 \times \ln(50) + 0.2095 \times \ln(100) = 1.111318$$

Using the corrected predictor given on page 95 of the text, the corresponding prediction of rice production is

$$\widehat{PROD} = \exp\left(\widehat{\ln(PROD)}\right) \times \exp\left(\hat{\sigma}^2/2\right) = \exp(1.111318) \times \exp\left(0.341419^2/2\right)$$

$$= 3.22071$$

(b) The marginal principle says apply more fertilizer if

$$\frac{\partial PROD}{\partial FERT} > \frac{PRICE_{FERT}}{PRICE_{RICE}} = 0.004$$

Since

$$\frac{\partial PROD}{\partial FERT} = \beta_4 \times \frac{PROD}{FERT},$$

the above marginal principle can be stated as: Apply more fertilizer if

$$\beta_4 \times \frac{PROD}{FERT} > 0.004$$

Substituting $PROD = 3.22071$ and $FERT = 100$, this inequality is the same as

$$\beta_4 > 0.004 \times \frac{100}{3.22071} = 0.1242$$

(c) The hypotheses are $H_0 : \beta_4 \le 0.1242$ and $H_1 : \beta_4 > 0.1242$. The 5% critical value is $t_{(0.95,348)} = 1.649$, leading to a rejection region of $t \ge 1.649$. The calculated *t*-value is

$$t = \frac{0.2095 - 0.1242}{0.0383} = 2.23$$

Since $2.23 > 1.649$, we reject H_0. The farmer should apply more fertilizer. We chose $\beta_4 > 0.1242$ as the alternative hypothesis because the farmer would not want to incur the cost of more fertilizer unless there was strong evidence from the data that it is profitable to do so.

Exercise 5.14 (continued)

(d) The predicted logarithm of rice production for $AREA = 1$, $LABOR = 50$, and $FERT = 1$ is

$$\widehat{\ln(PROD)} = -1.5468 + 0.3617 \times \ln(1) + 0.4328 \times \ln(50) + 0.2095 \times \ln(1) = 0.146525$$

Noting that $\ln(1) = 0$, the estimated variance of the prediction error can be written as

$$\begin{aligned}\widehat{\mathrm{var}(f)} &= \widehat{\mathrm{var}(b_1)} + \left[\ln(50)\right]^2 \times \widehat{\mathrm{var}(b_3)} + 2 \times \left[\ln(50)\right] \times \widehat{\mathrm{cov}(b_1, b_3)} + \hat{\sigma}^2 \\ &= 0.06535875 + 15.303924 \times 0.004473275 + 2 \times 3.912023 \times (-0.01360617) \\ &\qquad + 0.341419^2 \\ &= 0.143929\end{aligned}$$

and

$$\mathrm{se}(f) = \sqrt{\widehat{\mathrm{var}(f)}} = \sqrt{0.143929} = 0.37938$$

A 95% interval estimate for the logarithm of rice production at the specified values is

$$\widehat{\ln(PROD)} \pm t_{(0.975,348)} \mathrm{se}(f) = 0.146525 \pm 1.9668 \times 0.37938 = (-0.59964, 0.89269)$$

The corresponding 95% interval estimate for rice production is

$$\left(\exp(-0.59964), \exp(0.89269)\right) = (0.549, 2.442)$$

EXERCISE 5.15

(a) Taking logarithms yields the equation

$$\ln(Y_t) = \beta_1 + \beta_2 \ln(K_t) + \beta_3 \ln(L_t) + \beta_4 \ln(E_t) + \beta_5 \ln(M_t) + e_t$$

where $\beta_1 = \ln(\alpha)$. This form of the production function is linear in the coefficients β_1 , β_2 , β_3 , β_4 and β_5 , and hence is suitable for least squares estimation.

(b) Coefficient estimates and their standard errors are given in the following table.

	Estimated coefficient	Standard error
β_2	0.05607	0.25927
β_3	0.22631	0.44269
β_4	0.04358	0.38989
β_5	0.66962	0.36106

(c) The estimated coefficients show the proportional change in output that results from proportional changes in *K*, *L*, *E* and *M*. All these estimated coefficients have positive signs, and lie between zero and one, as is required for profit maximization to be realistic. Furthermore, they sum to approximately one, indicating that the production function has constant returns to scale. However, from a statistical point of view, all the estimated coefficients are not significantly different from zero; the large standard errors suggest the estimates are not reliable.