Chapter 4
Linear Regression with One Regressor

■ Solutions to Exercises

1. (a) The predicted average test score is

$$\widehat{TestScore} = 520.4 - 5.82 \times 22 = 392.36$$

(b) The predicted change in the classroom average test score is

$$\Delta\widehat{TestScore} = (-5.82 \times 19) - (-5.82 \times 23) = 23.28$$

(c) Using the formula for $\hat{\beta}_0$ in Equation (4.8), we know the sample average of the test scores across the 100 classrooms is

$$\overline{TestScore} = \hat{\beta}_0 + \hat{\beta}_1 \times \overline{CS} = 520.4 - 5.82 \times 21.4 = 395.85.$$

(d) Use the formula for the standard error of the regression (SER) in Equation (4.19) to get the sum of squared residuals:

$$SSR = (n-2)SER^2 = (100-2) \times 11.5^2 = 12961.$$

Use the formula for R^2 in Equation (4.16) to get the total sum of squares:

$$TSS = \frac{SSR}{1-R^2} = \frac{12961}{1-0.08^2} = 13044.$$

The sample variance is $s_Y^2 = \frac{TSS}{n-1} = \frac{13044}{99} = 131.8.$ Thus, standard deviation is $s_Y = \sqrt{s_Y^2} = 11.5.$

2. The sample size $n = 200$. The estimated regression equation is

$$\widehat{Weight} = (2.15) - 99.41 + (0.31)\ 3.94 Height, \qquad R^2 = 0.81, \quad \text{SER} = 10.2.$$

(a) Substituting $Height = 70,\ 65,$ and 74 inches into the equation, the predicted weights are 176.39, 156.69, and 192.15 pounds.

(b) $\Delta\widehat{Weight} = 3.94 \times \Delta Height = 3.94 \times 1.5 = 5.91.$

(c) We have the following relations: $1\,in = 2.54\,cm$ and $1\,lb = 0.4536\,kg$. Suppose the regression equation in the centimeter-kilogram space is

$$\widehat{Weight} = \hat{\gamma}_0 + \hat{\gamma}_1\, Height.$$

The coefficients are $\hat{\gamma}_0 = -99.41 \times 0.4536 = -45.092\,kg$; $\hat{\gamma}_1 = 3.94 \times \frac{0.4536}{2.54} = 0.7036\,kg$ per cm. The R^2 is unit free, so it remains at $R^2 = 0.81$. The standard error of the regression is $SER = 10.2 \times 0.4536 = 4.6267\,kg$.

3. (a) The coefficient 9.6 shows the marginal effect of *Age* on *AWE*; that is, *AWE* is expected to increase by $9.6 for each additional year of age. 696.7 is the intercept of the regression line. It determines the overall level of the line.
(b) *SER* is in the same units as the dependent variable (*Y*, or *AWE* in this example). Thus *SER* is measures in dollars per week.
(c) R^2 is unit free.
(d) (i) $696.7 + 9.6 \times 25 = \$936.7;$
(ii) $696.7 + 9.6 \times 45 = \$1{,}128.7$
(e) No. The oldest worker in the sample is 65 years old. 99 years is far outside the range of the sample data.
(f) No. The distribution of earning is positively skewed and has kurtosis larger than the normal.
(g) $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X},$ so that $\overline{Y} = \hat{\beta}_0 + \hat{\beta}_1 \overline{X}.$ Thus the sample mean of *AWE* is $696.7 + 9.6 \times 41.6 =$ $1,096.06.

4. (a) $(R - R_f) = \beta(R_m - R_f) + u,$ so that $\text{var}(R - R_f) = \beta^2 \times \text{var}(R_m - R_f) + \text{var}(u) + 2\beta \times \text{cov}(u, R_m - R_f).$
But $\text{cov}(u, R_m - R_f) = 0,$ thus $\text{var}(R - R_f) = \beta^2 \times \text{var}(R_m - R_f) + \text{var}(u).$ With $\beta > 1$, $\text{var}(R - R_f) >$ $\text{var}(R_m - R_f)$, follows because $\text{var}(u) \geq 0$.
(b) Yes. Using the expression in (a)
$\text{var}(R - R_f) - \text{var}(R_m - R_f) = (\beta^2 - 1) \times \text{var}(R_m - R_f) + \text{var}(u),$ which will be positive if
$\text{var}(u) > (1 - \beta^2) \times \text{var}(R_m - R_f).$
(c) $R_m - R_f = 7.3\% - 3.5\% = 3.8\%.$ Thus, the predicted returns are
$\hat{R} = R_f + \hat{\beta}(R_m - R_f) = 3.5\% + \hat{\beta} \times 3.8\%$
Kellog: $3.5\% + 0.24 \times 3.8\% = 4.4\%$
Waste Management: $3.5\% + 0.38 \times 3.8\% = 4.9\%$
Sprint: $3.5\% + 0.59 \times 3.8\% = 5.7\%$
Walmart: $3.5\% + 0.89 \times 3.8\% = 6.9\%$
Barnes and Noble: $3.5\% + 1.03 \times 3.8\% = 7.4\%$
Best Buy: $3.5\% + 1.8 \times 3.8\% = 10.3\%$
Microsoft: $3.5\% + 1.83 \times 3.8\% = 10.5\%$

5. (a) u_i represents factors other than time that influence the student's performance on the exam including amount of time studying, aptitude for the material, and so forth. Some students will have studied more than average, other less; some students will have higher than average aptitude for the subject, others lower, and so forth.
(b) Because of random assignment u_i is independent of X_i. Since u_i represents deviations from average $E(u_i) = 0$. Because u and X are independent $E(u_i|X_i) = E(u_i) = 0$.
(c) (2) is satisfied if this year's class is typical of other classes, that is, students in this year's class can be viewed as random draws from the population of students that enroll in the class. (3) is satisfied because $0 \leq Y_i \leq 100$ and X_i can take on only two values (90 and 120).
(d) (i) $49 + 0.24 \times 90 = 70.6;$ $49 + 0.24 \times 120 = 77.8;$ $49 + 0.24 \times 150 = 85.0$
(ii) $0.24 \times 10 = 2.4.$

6. Using $E(u_i|X_i) = 0,$ we have

$$E(Y_i|X_i) = E(\beta_0 + \beta_1 X_i + u_i|X_i) = \beta_0 + \beta_1 E(X_i|X_i) + E(u_i|X_i) = \beta_0 + \beta_1 X_i.$$

7. The expectation of $\hat{\beta}_0$ is obtained by taking expectations of both sides of Equation (4.8):

$$E(\hat{\beta}_0) = E(\bar{Y} - \hat{\beta}_1\bar{X}) = E\left[\left(\beta_0 + \beta_1\bar{X} + \frac{1}{n}\sum_{i=1}^{n} u_i\right) - \hat{\beta}_1\bar{X}\right]$$

$$= \beta_0 + E(\beta_1 - \hat{\beta}_1)\bar{X} + \frac{1}{n}\sum_{i=1}^{n} E(u_i|X_i) = \beta_0,$$

where the third equality in the above equation has used the facts that $\hat{\beta}_1$ is unbiased so $E(\beta_1 - \hat{\beta}_1) = 0$ and $E(u_i|X_i) = 0.$

8. The only change is that the mean of $\hat{\beta}_0$ is now $\beta_0 + 2$. An easy way to see this is this is to write the regression model as

$$Y_i = (\beta_0 + 2) + \beta_1 X_i + (u_i - 2).$$

The new regression error is $(u_i - 2)$ and the new intercept is $(\beta_0 + 2)$. All of the assumptions of Key Concept 4.3 hold for this regression model.

9. (a) With $\hat{\beta}_1 = 0, \hat{\beta}_0 = \bar{Y},$ and $\hat{Y}_i = \hat{\beta}_0 = \bar{Y}.$ Thus ESS = 0 and $R^2 = 0$.

(b) If $R^2 = 0$, then ESS = 0, so that $\hat{Y}_i = \bar{Y}$ for all i. But $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i,$ so that $\hat{Y}_i = \bar{Y}$ for all i, which implies that $\hat{\beta}_1 = 0,$ or that X_i is constant for all i. If X_i is constant for all i, then $\sum_{i-1}^{n}(X_i - \bar{X})^2 = 0$ and $\hat{\beta}_1$ is undefined (see equation (4.7)).

10. (a) $E(u_i|X = 0) = 0$ and $E(u_i|X = 1) = 0$. (X_i, u_i) are i.i.d. so that (X_i, Y_i) are i.i.d. (because Y_i is a function of X_i and u_i). X_i is bounded and so has finite fourth moment; the fourth moment is non-zero because $\Pr(X_i = 0)$ and $\Pr(X_i = 1)$ are both non-zero. Following calculations like those exercise 2.13, u_i also has nonzero finite fourth moment.

(b) $\text{var}(X_i) = 0.2 \times (1 - 0.2) = 0.16$ and $\mu_X = 0.2.$ Also

$$\text{var}[(X_i - \mu_X)u_i] = E[(X_i - \mu_X)u_i]^2$$

$$= E[(X_i - \mu_X)u_i|X_i = 0]^2 \times \Pr(X_i = 0) + E[(X_i - \mu_X)u_i|X_i = 1]^2 \times \Pr(X_i = 1)$$

where the first equality follows because $E[(X_i - \mu_X)u_i] = 0$, and the second equality follows from the law of iterated expectations.

$$E[(X_i - \mu_X)u_i|X_i = 0]^2 = 0.2^2 \times 1, \text{and } E[(X_i - \mu_X)u_i|X_i = 1]^2 = (1 - 0.2)^2 \times 4.$$

Putting these results together

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{(0.2^2 \times 1 \times 0.8) + ((1 - 0.2)^2 \times 4 \times 0.2)}{0.16^2} = \frac{1}{n}21.25$$

11. (a) The least squares objective function is $\sum_{i=1}^{n}(Y_i - b_1X_i)^2$. Differentiating with respect to b_1 yields $\frac{\partial \sum_{i=1}^{n}(Y_i - b_1X_i)^2}{\partial b_1} = -2\sum_{i=1}^{n} X_i(Y_i - b_1X_i)$. Setting this zero, and solving for the least squares estimator yields $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_iY_i}{\sum_{i=1}^{n} X_i^2}$.

(b) Following the same steps in (a) yields $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i(Y_i - 4)}{\sum_{i=1}^{n} X_i^2}$

12. (a) Write

$$ESS = \sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2 = \sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_i - \overline{Y})^2 = \sum_{i=1}^{n}[\hat{\beta}_1(X_i - \overline{X})]^2$$

$$= \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{\left[\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})\right]^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

So that

$$R^2 = \frac{ESS}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} = \frac{\left[\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})\right]^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}$$

$$= \left[\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}\right]^2$$

$$= \left[\frac{s_{XY}}{s_X s_Y}\right]^2 = r_{XY}^2$$

(b) This follows from part (a) because $r_{XY} = r_{YX}$.