Extra Practice Questions

Chapter 6: 6.1-6.5, 6.9

1. By equation (6.15) in the text, we know

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2).$$

Thus, that values of $\bar{R}^2$ are 0.175, 0.189, and 0.193 for columns (1)–(3).

2. (a) Workers with college degrees earn \$5.46/hour more, on average, than workers with only high school degrees.
   (b) Men earn \$2.64/hour more, on average, than women.

3. (a) On average, a worker earns \$0.29/hour more for each year he ages.
   (b) Sally's earnings prediction is $4.40 + 5.48 \times 1 - 2.62 \times 1 + 0.29 \times 29 = 15.67$ dollars per hour. Betsy's earnings prediction is $4.40 + 5.48 \times 1 - 2.62 \times 1 + 0.29 \times 34 = 17.12$ dollars per hour. The difference is 1.45

4. (a) Workers in the Northeast earn \$0.69 more per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the Northeast earn \$0.60 more per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the South earn \$0.27 less than workers in the West.
   (b) The regressor *West* is omitted to avoid perfect multicollinearity. If *West* is included, then the intercept can be written as a perfect linear function of the four regional regressors. Because of perfect multicollinearity, the OLS estimator cannot be computed.
   (c) The expected difference in earnings between Juanita and Jennifer is $-0.27 - 0.6 = -0.87$.

5. (a) \$23,400 (recall that *Price* is measured in \$1000s).
   (b) In this case $\Delta BDR = 1$ and $\Delta Hsize = 100$. The resulting expected change in price is $23.4 + 0.156 \times 100 = 39.0$ thousand dollars or \$39,000.
   (c) The loss is \$48,800.
   (d) From the text $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2)$, so $R^2 = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2)$, thus, $R^2 = 0.727$.

9. For omitted variable bias to occur, two conditions must be true: X_1 (the included regressor) is correlated with the omitted variable, and the omitted variable is a determinant of the dependent variable. Since X_1 and X_2 are uncorrelated, the estimator of β_1 does not suffer from omitted variable bias.

Chapter 7: 7.1, 7.2 (a), 7.3-7.5, 7.7

1.

Regressor	(1)	(2)	(3)
College (X_1)	5.46** (0.21)	5.48** (0.21)	5.44** (0.21)
Female (X_2)	–2.64** (0.20)	–2.62** (0.20)	–2.62** (0.20)
Age (X_3)		0.29** (0.04)	0.29** (0.04)
Ntheast (X_4)			0.69* (0.30)
Midwest (X_5)			0.60* (0.28)
South (X_6)			–0.27 (0.26)
Intercept	12.69** (0.14)	4.40** (1.05)	3.75** (1.06)

(a) The t-statistic is $5.46/0.21 = 26.0 > 1.96$, so the coefficient is statistically significant at the 5% level. The 95% confidence interval is $5.46 \pm 1.96 \times 0.21$.

(b) t-statistic is $-2.64/0.20 = -13.2$, and $13.2 > 1.96$, so the coefficient is statistically significant at the 5% level. The 95% confidence interval is $-2.64 \pm 1.96 \times 0.20$.

3. (a) Yes, age is an important determinant of earnings. Using a t-test, the t-statistic is $\frac{0.29}{0.04} = 7.25$, with a p-value of 4.2×10^{-13}, implying that the coefficient on age is statistically significant at the 1% level. The 95% confidence interval is $0.29 \pm 1.96 \times 0.04$.

(b) $\Delta Age \times [0.29 \pm 1.96 \times 0.04] = 5 \times [0.29 \pm 1.96 \times 0.04] = 1.45 \pm 1.96 \times 0.20 = \$1.06$ to $\$1.84$

4. (a) The F-statistic testing the coefficients on the regional regressors are zero is 6.10. The 1% critical value (from the $F_{3,\infty}$ distribution) is 3.78. Because $6.10 > 3.78$, the regional effects are significant at the 1% level.

(b) The expected difference between Juanita and Molly is $(X_{6,\text{Juanita}} - X_{6,\text{Molly}}) \times \beta_6 = \beta_6$. Thus a 95% confidence interval is $-0.27 \pm 1.96 \times 0.26$.

(c) The expected difference between Juanita and Jennifer is $(X_{5,\text{Juanita}} - X_{5,\text{Jennifer}}) \times \beta_5 + (X_{6,\text{Juanita}} - X_{6,\text{Jennifer}}) \times \beta_6 = -\beta_5 + \beta_6$. A 95% confidence interval could be contructed using the general methods discussed in Section 7.3. In this case, an easy way to do this is to omit *Midwest* from the regression and replace it with $X_5 = West$. In this new regression the coefficient on *South* measures the difference in wages between the *South* and the *Midwest*, and a 95% confidence interval can be computed directly.

5. The t-statistic for the difference in the college coefficients is $t = (\hat{\beta}_{college,1998} - \hat{\beta}_{college,1992})/SE(\hat{\beta}_{college,1998} - \hat{\beta}_{college,1992})$. Because $\hat{\beta}_{college,1998}$ and $\hat{\beta}_{college,1992}$ are computed from independent samples, they are independent, which means that $\text{cov}(\hat{\beta}_{college,1998}, \hat{\beta}_{college,1992}) = 0$ Thus, $\text{var}(\hat{\beta}_{college,1998} - \hat{\beta}_{college,1992}) = \text{var}(\hat{\beta}_{college,1998}) + \text{var}(\hat{\beta}_{college,1998})$. This implies that $SE(\hat{\beta}_{college,1998} - \hat{\beta}_{college,1992}) = (0.21^2 + 0.20^2)^{\frac{1}{2}}$. Thus, $t^{act} = \frac{5.48-5.29}{(0.21^2+0.20^2)^{\frac{1}{2}}} = 0.6552$. There is no significant change since the calculated t-statistic is less than 1.96, the 5% critical value.

7. (a) The t-statistic is $\frac{0.485}{2.61} = 0.186 < 1.96$. Therefore, the coefficient on BDR is not statistically significantly different from zero.

(b) The coefficient on *BDR* measures the *partial effect* of the number of bedrooms holding house size (*Hsize*) constant. Yet, the typical 5-bedroom house is much larger than the typical 2-bedroom house. Thus, the results in (a) says little about the conventional wisdom.

(c) The 99% confidence interval for effect of lot size on price is $2000 \times [.002 \pm 2.58 \times .00048]$ or 1.52 to 6.48 (in thousands of dollars).

(d) Choosing the scale of the variables should be done to make the regression results easy to read and to interpret. If the lot size were measured in thousands of square feet, the estimate coefficient would be 2 instead of 0.002.

(e) The 10% critical value from the $F_{2,\infty}$ distribution is 2.30. Because $0.08 < 2.30$, the coefficients are not jointly significant at the 10% level.