Answers for Homework #5

Zheng Tian

7.7

1. The t-statistic is $0.485/2.61 = 0.186 < 1.96$. Therefore, the coefficient on BDR is not statistically significantly different from zero.
2. 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.
3. The 99% confidence interval for effect of lot size on price is $2000 \times [0.002 \pm 2.58 \times 0.00048]$ or 1.52 to 6.48 (in thousands of dollars).
4. 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.
5. 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.

7.8

1. Using the expressions for R^2 and $\bar{R}^2$, algebra shows that

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

   - Column 1: $R_2 = 1 - \frac{420-1-1}{420-1}(1-0.049) = 0.051$
   - Column 2: $R^2 = 0.0427$
   - Column 3: $R_2 = 0.775$
   - Column 4: $R_2 = 0.629$
   - Column 5: $R_2 = 0.775$

2. $H_0 : \beta_3 = \beta_4 = 0$ vs. $\beta_3 \neq 0$ or $\beta_4 \neq 0$
   - Unrestricted regression (Column 5): $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$, $R^2_{\text{unrestricted}} = 0.775$.
   - Restricted regression (Column 2): $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, $R^2_{\text{restricted}} = 0.427$
   - The homoskedasticity-only F statistic is

$$\begin{aligned} F &= \frac{(R^2_{\text{unrestricted}} - R^2_{\text{restricted}})/q}{(1-R^2_{\text{unrestricted}})/(n-k-1)} \\ &= \frac{(0.775-0.427)/2}{(1-0.775)/(420-4-1)} = 322.22 \end{aligned}$$

The 5% critical value of the $F(2, 415)$ distribution is 3.017, which is less than the F-statistic. Thus, we reject the null hypothesis at the 5% level.

3. The confidence interval is $-1.01 \pm 2.58 \times 0.27$.

7.9 1. Estimate

$$Y = \beta_0 + \gamma X_1 + \beta_2(X_1 + X_2) + u$$

and test whether $\gamma = 0$.

2. Estimate

$$Y = \beta_0 + \gamma X_1 + \beta_2(X_2 - aX_1) + u$$

and test whether $\gamma = 0$.

3. Estimate

$$Y - X_1 = \beta_0 + \gamma X_1 + \beta_2(X_2 - X_1) + u$$

and test whether $\gamma = 0$.

7.11 1. Treatment (assignment to small classes) was not randomly assigned in the population (the continuing and newly-enrolled students) because of the difference in the proportion of treated continuing and newly-enrolled students. Thus, the treatment indicator X_1 is correlated with X_2. If newly-enrolled students perform systematically differently on standardized tests than continuing students (perhaps because of adjustment to a new school), then this becomes part of the error term u. This leads to correlation between X_1 and u, so that $E(u|X_1) \neq 0$. Because $E(u|X_1) \neq 0$, the $\hat{\beta}_1$ is biased and inconsistent.

2. Because treatment was randomly assigned conditional on enrollment status (continuing or newly-enrolled), $E(u|X_1, X_2)$ will not depend on X_1. This means that the assumption of conditional mean independence is satisfied, and $\hat{\beta}_1$ is unbiased and consistent. However, because X_2 was not randomly assigned (newly-enrolled students may, on average, have attributes other than being newly enrolled that affect test scores), $E(u|X_1, X_2)$ may depend of X_2, so $\hat{\beta}_2$ that may be biased and inconsistent.