Problemset 7

1. Exercise 7.1 from Greene.

2. Exercise 7.2 from Greene.

3. Greene 6.25. Hint: With the full regression, the residuals are $e'e$. With $x_k$ removed the residuals are $e'_{(-1)}e_{(-1)}$. Then

$$e'_{(-1)}e_{(-1)} = e'e + \hat{\beta}_k^2 (x_k'M_{(-1)}x_k)$$

where $M_{(-1)}$ is the matrix projecting onto the space orthogonal to all regressors except $x_k$.

4. The data generating process (DGP) of concern is:

$$y = \alpha + \beta x + \varepsilon, \quad \varepsilon \sim N(0, 2), \quad \alpha = 1, \quad \beta = 3$$

and $x$ is the variable available in PS5dat.txt on the problemset web page. There are 30 observations. Generate 10,000 datasets (*Hint: Start with a smaller number until you are sure your program is correct*) according to this DGP. For each dataset, use a $t-test$ to test the hypothesis:

- $H_0 : \beta = 3$
- $H_1 : \beta = 3.5$
- $H_2 : \beta = 4$
- $H_3 : \beta = 5$
- $H_4 : \beta = 6$
- $H_5 : \beta = 7$
- $H_6 : \beta = 10$

(a) For each hypothesis, what is the percentage of datasets for which you reject the hypothesis?

(b) What does this tell you about the size and power of the $t-test$?