Intro to Econometric Theory
Heinz School, Carnegie Mellon University
90-906, Spring 2003-4

Solution #3

Consider the hospital cost data described on the datasets page of the course website. Suppose we are interested in understanding the costs of hospital care in California. We are going to model these data with the following CNLRM:

$$\text{cost}_i = \beta_1 + \beta_2 \text{days}_i + \beta_3 \text{visits}_i + \epsilon_i$$

Analysis of these data appears in the file hwk3.lst available on the homework page of the course website. The analysis was generated with SAS from the program hwk3.sas also available on the course website.

1. Interpret β_2.

   β_2 is the effect, on hospital costs, of a one unit increase in inpatient days. That is, it is the marginal cost of inpatient days.

2. Interpret $\hat{\beta}_2$.

   $\hat{\beta}_2$ is an *estimator* of the effect, on hospital costs, of a one unit increase in inpatient days. That is, it is an estimator of the marginal cost of inpatient days.

   In our case, it is the OLS estimator of the marginal cost of an inpatient day.

3. Interpret 207.82

   207.82 is our *OLS estimate* of the effect, on hospital costs, of a one unit increase in inpatient days. That is, it is our OLS estimate of the marginal cost of inpatient days.

4. Form a 95% confidence interval for the marginal cost of an outpatient visit.

   We are asked for a confidence interval for β_3. Notice, since N is so large, the t and normal tables are basically the same.

$$\begin{aligned}
\beta_2 &= \hat{\beta}_{2,OLS} \pm 1.96\sqrt{\hat{V}(\hat{\beta}_{2,OLS})} \\
&= 1477 \pm 1.96(12.5) \\
&= 1477 \pm 25
\end{aligned}$$

5. Test, at the 1% confidence level, the hypothesis that the marginal cost of an inpatient day is one fifth the marginal cost of an outpatient visit.

   The null hypothesis here is: $H_0: \ \beta_2 = 0.2\beta_3$ **or** $H_0: \ \beta_2 - 0.2\beta_3 = 0$.

$$
\begin{aligned}
\text{t}-\text{stat} &= \frac{\hat{\beta}_{2,OLS} - 0.2\hat{\beta}_{3,OLS} - 0}{\sqrt{\begin{pmatrix} 0 & 1 & -0.2 \end{pmatrix} \hat{V}(\hat{\beta}_{OLS}) \begin{pmatrix} 0 \\ 1 \\ -0.2 \end{pmatrix}}} \\
&= \frac{-87.6}{\sqrt{24.13 + 0.04(156) + 2(0.2)(-44)}} \\
&= \frac{-87.6}{\sqrt{12.77}} \\
&= -24.5
\end{aligned}
$$

   Clearly, 24.5 is high enough that the null hypothesis will be rejected at any reasonable significance level.

6. What is $X'X$? (hint: first prove that $X'X = (N-1)\hat{V}(X) + N\overline{X}\,\overline{X}'$)

   We have argued in class that $V(x) = E(xx') - E(x)(E(x))'$. **This is the same as** $E(x-Ex)(x-Ex)' = E(xx') - E(x)(E(x))'$, **or** $E(xx') = E(x-Ex)(x-Ex)' + E(x)(E(x))'$. **The sample analogue of this (which is proved by following the exact same steps) is:**

$$
\frac{1}{N}\sum X^i X^{i'} = \frac{1}{N}\sum (X^i - \overline{X^i})(X^i - \overline{X^i})' + \overline{X^i}\,\overline{X^{i'}}
$$

   Multiplying the equation by N gives the result.

   Now, calculating:

$$
\begin{aligned}
X'X &= (N-1)\hat{V}(x) + N\overline{x}\,\overline{x} \\
&= 7534 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 584022051 & 165336858 \\ 0 & 165336858 & 90425970 \end{bmatrix} + 7535 \begin{bmatrix} 1 \\ 20785 \\ 9853 \end{bmatrix} \begin{bmatrix} 1 & 20785 & 9853 \end{bmatrix} \\
&= 7534 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 584022051 & 165336858 \\ 0 & 165336858 & 90425970 \end{bmatrix} + 7535 \begin{bmatrix} 1 & 20785 & 9853 \\ 20785 & 432016225 & 204794605 \\ 9853 & 204794605 & 97081609 \end{bmatrix} \\
&= \begin{bmatrix} 7535 & 1.57 \times 10^8 & 7.42 \times 10^7 \\ 1.57 \times 10^8 & 7.66 \times 10^12 & 2.79 \times 10^12 \\ 7.42 \times 10^7 & 2.79 \times 10^12 & 1.41 \times 10^12 \end{bmatrix}
\end{aligned}
$$

7. When there is a constant in the model, what do the first row and the first column of $X'X$ contain?

   The 1,1 position contains N and the other elements of the first row/column contain the sum of each of the corresponding Xs.