Additional Tutorial 5

Binary Variables

Problem 1 (Wooldridge, Problem 7.1, Page 255)

Using the data in SLEEP75.RAW (see also Problem 3.3), we obtain the estimated equation

\[
\hat{\text{sleep}} = 3,840.83 - .163 \cdot \text{totwrk} - 11.71 \cdot \text{educ} - 8.70 \cdot \text{age} + .128 \cdot \text{age}^2 + 87.75 \cdot \text{male} \\
\]

\[
n = 706, R^2 = .123, \bar{R}^2 = .117.
\]

The variable \( \text{sleep} \) is total minutes per week spent sleeping at night, \( \text{totwrk} \) is total weekly minutes spent working, \( \text{educ} \) and \( \text{age} \) are measured in years, and \( \text{male} \) is a gender dummy.

(i) All other factors being equal, is there evidence that men sleep more than women? How strong is the evidence?

(ii) Is there a statistically significant tradeoff between working and sleeping? What is the estimated tradeoff?

(iii) What other regression do you need to run to test the null hypothesis that, holding other factors fixed, age has no effect on sleeping?

Problem 2 (Wooldridge, Problem 7.3, Page 256)

Using the data in GPA2.RAW, the following equation was estimated:

\[
\hat{\text{sat}} = 1,028.10 + 19.30 \cdot \text{hsize} - 2.19 \cdot \text{hsize}^2 - 45.09 \cdot \text{female} - 169.81 \cdot \text{black} + 62.31 \cdot \text{female} \times \text{black} \\
\]

\[
n = 4,137, R^2 = .0858.
\]

The variable \( \text{sat} \) is the combined SAT score, \( \text{hsize} \) is size of the student’s high school graduating class, in hundreds, \( \text{female} \) is a gender dummy variable, and \( \text{black} \) is a race dummy variable equal to one for blacks and zero otherwise.

(i) Is there strong evidence that \( \text{hsize}^2 \) should be included in the model? From this equation, what is the optimal high school size?

(ii) Holding \( \text{hsize} \) fixed, what is the estimated difference in SAT score between nonblack females and nonblack males? How statistically significant is this estimated difference?

(iii) What is the estimated difference in SAT score between nonblack males and black males? Test the null hypothesis that there is no difference between their scores, against the alternative that there is a difference.

(iv) What is the estimated difference in SAT score between black females and nonblack females? What would you need to do to test whether the difference is statistically significant?
Problem 3 (Wooldridge, Problem 7.4, Page 256-257)

An equation explaining chief executive officer salary is

\[
\log(\text{salary}) = 4.59 + 0.257 \log(\text{sales}) + 0.011 \text{roe} + 0.158 \text{finance} + 0.181 \text{consprod} - 0.283 \text{utility}
\]

\[
R^2 = 0.357
\]

The data used are in CEOSAL1.RAW, where \text{finance}, \text{consprod}, and \text{utility} are binary variables indicating the financial, consumer products, and utilities industries. The omitted industry is transportation.

(i) Compute the approximate percentage difference in estimated salary between the utility and transportation industries, holding \text{sales} and \text{roe} fixed. Is the difference statistically significant at the 1\% level?

(ii) Use equation (7.10) to obtain the exact percentage difference in estimated salary between the utility and transportation industries and compare this with the answer obtained in part (i).

(iii) What is the approximate percentage difference in estimated salary between the consumer products and finance industries? Write an equation that would allow you to test whether the difference is statistically significant.