HOMEWORK 3

Wittawat Suwattananon
6104641466

1. Using the data in GPA2 on 4,137 college students, the following equation was estimated by OLS:

$$\widehat{colgpa} = 1.392 - .0135\ hsperc + .00148\ sat$$
$$n = 4{,}137,\ R^2 = .273,$$

where *colgpa* is measured on a four-point scale, *hsperc* is the percentile in the high school graduating class (defined so that, for example, $hsperc = 5$ means the *top* 5% of the class), and *sat* is the combined math and verbal scores on the student achievement test.

   i. Why does it make sense for the coefficient on *hsperc* to be negative?

   ii. What is the predicted college GPA when $hsperc = 20$ and $sat = 1{,}050$?

   iii. Suppose that two high school graduates, A and B, graduated in the same percentile from high school, but Student A's SAT score was 140 points higher (about one standard deviation in the sample). What is the predicted difference in college GPA for these two students? Is the difference large?

   iv. Holding *hsperc* fixed, what difference in SAT scores leads to a predicted *colgpa* difference of .50, or one-half of a grade point? Comment on your answer.

(i) The smaller the hsperc the higher rank in class you will be.

(ii) $\widehat{colgpa} = 1.392 - .0135(20) + .00148(1{,}050)$
$= 2.676$

(iii) $\frac{\partial colgpa}{\partial SAT} = 0.00148$, ceteris paribus.

$\therefore \Delta colgpa = 0.00148 \cdot \Delta SAT^{(140)}$
$= 0.2072$

(iv) $0.5 = 0.00148 \cdot \Delta SAT$ The effect is too large.
$\Delta SAT = 337.84$,

---

2. The data in WAGE2 on working men was used to estimate the following equation:

$$\widehat{educ} = 10.36 - .094\ sibs + .131\ meduc + .210\ feduc$$
$$n = 722,\ R^2 = .214,$$

where *educ* is years of schooling, *sibs* is number of siblings, *meduc* is mother's years of schooling, and *feduc* is father's years of schooling.

   i. Does *sibs* have the expected effect? Explain. Holding *meduc* and *feduc* fixed, by how much does *sibs* have to increase to reduce predicted years of education by one year? (A noninteger answer is acceptable here.)

   ii. Discuss the interpretation of the coefficient on *meduc*.

   iii. Suppose that Man A has no siblings, and his mother and father each have 12 years of education. Man B has no siblings, and his mother and father each have 16 years of education. What is the predicted difference in years of education between B and A?

(i) sibs have negative effect to education.

$\frac{\partial educ}{\partial sibs} = -0.094$

$\Delta educ = -0.094 \Delta sibs$

$1 = -0.094 \Delta sibs$

$\Delta sibs = -10.64$

(ii) meduc, a year of mother schooling increases an education by 0.131 year.

$\frac{\partial educ}{\partial meduc} = 0.131$

(iii) $\widehat{educ}_A = 10.36 + 0.131(12) + 0.210(12)$
$= 14.452$

$\widehat{educ}_B = 10.36 + 0.131(16) + 0.210(16)$
$= 15.816$

$\Delta educ_{A,B} = 1.364$

1. Let *kids* denote the number of children ever born to a woman, and let *educ* denote years of education for the woman. A simple model relating fertility to years of education is

$$kids = \beta_0 + \beta_1 educ + u,$$

where u is the unobserved error.

i. What kinds of factors are contained in u? Are these likely to be correlated with level of education?

ii. Will a simple regression analysis uncover the ceteris paribus effect of education on fertility? Explain.

(i) income, preference, ~~culture~~, age, height
(+) (not relate) (+) (not relate)

(ii) no, b/c there are others factor in an observed error which correlated w/ level of education.
Thus, β_1 will not reflect only the effect of level of education alone.

---

4. The data set BWGHT contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (*bwght*), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (*cigs*). The following simple regression was estimated using data on $n = 1{,}388$ *births*:

$$\widehat{bwght} = 119.77 - 0.514\, cigs$$

i. What is the predicted birth weight when $cigs = 0$? What about when $cigs = 20$ (one pack per day)? Comment on the difference.

ii. Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.

iii. To predict a birth weight of 125 ounces, what would *cigs* have to be? Comment.

iv. The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?

(i)
- cigs = 0; $\widehat{bwght}$ = 119.77
- cigs = 20; $\widehat{bwght}$ = 109.49
- 10.28 differences, mothers who don't smoke, kids tend to weight heavier.

(ii) Yes, the more a mother smoke during pregnancy the lighter the infant will weight. This implies negative effect on infant's health.

(iii) $125 = 119.77 - 0.514 cigs$
$cigs = -10.175$
∴ it doesn't make sense, the data & regression line doesn't fit in w/ bwght.

(iv) bias, b/c the data collected was the same (85%) on non-smokers.
Violating assumption SLR 3.

3. The following model is a simplified version of the multiple regression model used by Biddle and Hamermesh (1990) to study the tradeoff between time spent sleeping and working and to look at other factors affecting sleep:

$$sleep = \beta_0 + \beta_1 totwrk + \beta_2 educ + \beta_3 age + u,$$

where *sleep* and *totwrk* (total work) are measured in minutes per week and *educ* and *age* are measured in years. (See also Computer Exercise C3 in Chapter 2.)

i. If adults trade off sleep for work, what is the sign of β_1?

ii. What signs do you think β_2 and β_3 will have?

iii. Using the data in SLEEP75, the estimated equation is

$$\widehat{sleep} = 3,638.25 - .148\ totwrk - 11.13\ educ + 2.20\ age$$
$$n = 706,\ R^2 = .113.$$

If someone works five more hours per week, by how many minutes is *sleep* predicted to fall? Is this a large tradeoff?

iv. Discuss the sign and magnitude of the estimated coefficient on *educ*.

v. Would you say *totwrk*, *educ*, and *age* explain much of the variation in *sleep*? What other factors might affect the time spent sleeping? Are these likely to be correlated with *totwrk*?

(i) β_1 would be negative
The more sleep, the less work.

(ii) β_2 & β_3 are negative.
the more educations you thrive the less hrs you sleep. The older you get the more work you do the less you sleep.

(iii) $\Delta sleep = -.148\ \Delta totwrk$

$= -.148(300)$

$\Delta sleep = -44.4$

Yes, b/c 5 additional hours of working = 44.4 minutes of losing sleep.