HOMEWORK 3

Tanyaton Panyadilok
6104641367

CHAPTER 2 : 1,4

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

# of children born

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

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) other factors like income (correlate), preference (not correlate), age (correlate)

ii) No because there are unobserved error which correlated with the level of education. Therefore, β_1 did not measure the 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.

$cigs = 0 \rightarrow \widehat{bwght} = 119.77 - 0.514(0) = 119.77$

$cigs = 20 \rightarrow \widehat{bwght} = 119.77 - 0.514(20) = 109.49$

- The mothers who don't smoke tend to have heavier babies than mothers who smoke

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

Yes, because the more a mother smoke, the lighter weight of the baby.
This refers to a negative relationship between smoking habit and the infant weight.

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

$\widehat{bwght} = 119.77 - 0.514\,cigs = 125$

$cigs = -10.175$

The number of cigarettes can not be a negative number

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)?

It doesn't help because it biased. It violates assumption SLR3.

Chapter 3 : 1, 2

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?

because the smaller hsperc, the higher rank you will get.

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

$$\widehat{colgpa} = 1.392 - 0.0135(20) + 0.00148(1050)$$
$$= 2.676$$

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?

student A : $\widehat{colgpa} = 1.392 - 0.0135(20) + 0.00148(1050)$
$= 2.676$

student B : $\widehat{colgpa} = 1.392 - 0.0135(20) + 0.00148(1190)$
$= 1.392 - 0.27 + 1.7612$
$= 2.8832$

The difference between student A and B ≈ 0.2072

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.

$$\widehat{colgpa} = 1.392 - 0.0135\underbrace{(hsperc)}_{\text{fixed}} + 0.00148(sat)$$

$0.5 = 0.00148 \cdot \Delta SAT$

The effect is too large $\rightarrow \Delta SAT = 337.84$

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.)

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

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

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

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?

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

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

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