

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.

- What kinds of factors are contained in *u*? Are these likely to be correlated with level of education?
- Will a simple regression analysis uncover the *ceteris paribus* effect of education on fertility? Explain.

→ keeping other factor constant

i) ปัจจัยที่ measure ได้ : income , age , SAT score , น้ำหนัก , number of sibling

income (positive) , age (positive) , SAT score (positive) , น้ำหนัก (no correlation) , number of sibling (negative)

ii) เวลา education เปลี่ยนตัวแปรอื่นเปลี่ยน

= other factor change as well which effect kids

ปกติ $E(u) = 0$

2. In the simple linear regression model $y = \beta_0 + \beta_1 x + u$, suppose that $E(u) \neq 0$

. Letting $\alpha_0 = E(u)$, show that the model can always be rewritten with the same slope, but a new intercept and error, where the new error has a zero expected value.

PROVE

$$y = \beta_0 + \beta_1 x + u$$

take summation
แล้วนำตัว n

$$\sum c = nc$$

$$\frac{\sum c}{n} = \frac{nc}{n}$$

$$\frac{\sum y}{n} = \frac{\sum \beta_0}{n} + \frac{\sum \beta_1 x}{n} + \frac{\sum u}{n}$$

$$\bar{y} = \beta_0$$

นี่จ้ะ

(iii)

iv. How much of the variation in GPA for these eight students is explained by ACT? Explain.

หา R^2

ดูว่า $x(AC)$ explain $y(GPA)$
ได้กี่เปอร์เซ็นต์?

$$R^2 = 1 - \frac{SSB}{SST} = 1 - \frac{\sum \hat{u}^2}{\sum (y_i - \bar{y})^2}$$

$$= 1 - \frac{0.434725}{1.02875} = 0.5774 \%$$

อธิบายได้ 57.74 %

GPA is explained by ACT about 57.74 %

แปลว่าส่วนที่เหลือไม่ได้มาจาก ACT มาจากอย่างอื่น

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:

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

- What is the predicted birth weight when *cigs* = 0? What about when *cigs* = 20 (one pack per day)? Comment on the difference.
- Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.
- To predict a birth weight of 125 ounces, what would *cigs* have to be? Comment.
- 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)?

ตัว hat = ตัวที่ถูก explain

i) Put 0 and 20 to the function

$$\begin{aligned}\widehat{bwght} &= 119.77 - 0.514(0) \\ &= 119.77\end{aligned}$$

$$\begin{aligned}\widehat{bwght} &= 119.77 - 0.514(20) \\ &= 109.49\end{aligned}$$

∴ เด็กจะหนักน้อยกว่า ถ้าแม่สูบบุหรี่

ii)

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

Does not capture everything because there are some other factors
ex) health, mother's age

iii)

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

$$125 = 119.77 - 0.514 \text{ cigs}$$
$$\text{cigs} = -10.1751$$

ตัดสิน = มี factor อื่น, regression explain ทุกกรณีไม่ได้.
estimator = biased (ความสัมพันธ์กับ *cigs*)

iv)

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)?

จาก 1388
1180 คน ไม่สูบบุ.

$X = 0$ ตั้งฟังก์ชัน
 $\therefore$ sample คำนวณค่า
mean of *cigs* จะต่ำใกล้เคียง 0

1 ใน assumption ที่นำมา sample ชื่อ "SLR 3"

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

$$\text{college GPA} = 1.392 - .0135 \text{ hspc} + .00148 \text{ sat}$$

$n = 4,137, R^2 = .273,$

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?

1) The higher percentile, the lower college GPA

The lower percentile, the higher rank your college GPA

$$\begin{aligned} \text{ii)} \quad \hat{\text{colgpa}} &= 1.392 - 0.0135(20) + 0.00148(1050) \\ &= 2.676 \end{aligned}$$

- 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?

→ หา partial ก่อน
แล้วทำเป็น change *

If A and B have same percentile from HS, keep *hsperc* constant

$$\frac{\partial \text{colgpa}}{\partial \text{SAT}} = 0.00148$$

Ceteris paribus,

$$\begin{aligned} \therefore \Delta \text{ in college gpa} &= 0.00148 \Delta \text{SAT} \\ &= 0.00148(140) \end{aligned}$$

$$= 0.2072 \quad \rightarrow \text{เกรดขึ้น 0.2072.}$$

The predicted different in college GPA for these two student is 0.2072. It is only a small increase

iv. Holding *hspcr* 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.

$$0.5 = 0.00148 * \Delta SAT$$

$$337.84 = \Delta SAT$$

$\therefore$ It is too exaggerate to increase SAT score for 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.)

$$1) \frac{\Delta edu}{\Delta sib} = -0.094$$

Keeping other factor constant, the increase in *sibs* in 1 reduce year of edu decrease for 0.094 ปี.

$$\Delta edu = -0.094 \Delta sibling$$

$$1 = -0.094 \Delta sib$$

$$\Delta sib = -10.64 \# \quad \therefore \text{ต้องลดพี่น้องไป 10.64 คน}$$

เพื่อเพิ่ม education 1 ปี
ceteris paribus

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?

$$2) \frac{\Delta edu}{\Delta meduc} = 0.131$$

/ ค่าของ father & mother

$$3) 0.131 (4) + 0.210 (4) = 1.364$$

OR

$$10.36 - 0.094 sib + 0.131 meduc + 0.210 feduc$$

ใส่ค่าของ father and mother.