HW0324_313707002_羅芷羚
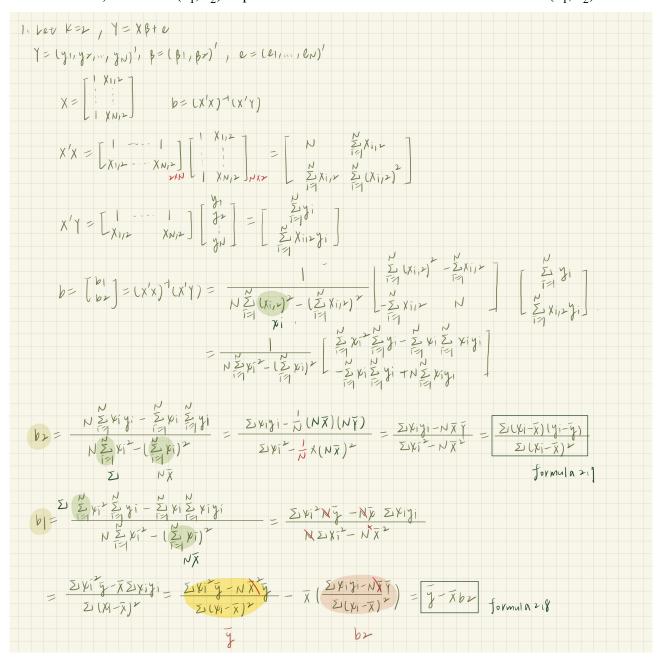
1. Let $K = 2$, show that (b_1, b_2) in p. 29 of slides in Ch5 reduces to the formula of (b_1, b_2)

1. Let $K=2$, $Y = X\beta + e$

$Y = (y_1, y_2, \cdots, y_N)'$, $\beta = (\beta_1, \beta_2)'$, $e = (e_1, \cdots, e_N)'$

$$X = \begin{bmatrix} 1 & X_{1,2} \\ \vdots & \vdots \\ 1 & X_{N,2} \end{bmatrix} \qquad b = (X'X)^{-1}(X'Y)$$

$$X'X = \begin{bmatrix} 1 & \cdots & 1 \\ X_{1,2} & \cdots & X_{N,2} \end{bmatrix}_{2\times N} \begin{bmatrix} 1 & X_{1,2} \\ \vdots & \vdots \\ 1 & X_{N,2} \end{bmatrix}_{N\times 2} = \begin{bmatrix} N & \sum_{i=1}^{N} X_{i,2} \\ \sum_{i=1}^{N} X_{i,2} & \sum_{i=1}^{N} (X_{i,2})^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & \cdots & 1 \\ X_{1,2} & & X_{N,2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} X_{i,2} y_i \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = (X'X)^{-1}(X'Y) = \frac{1}{N\sum_{i=1}^{N}(X_{i,2})^2 - (\sum_{i=1}^{N} X_{i,2})^2} \begin{bmatrix} \sum_{i=1}^{N}(X_{i,2})^2 & -\sum_{i=1}^{N}X_{i,2} \\ -\sum_{i=1}^{N}X_{i,2} & N \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} X_{i,2} y_i \end{bmatrix}$$

$$= \frac{1}{N\sum_{i=1}^{N} X_i^2 - (\sum_{i=1}^{N} X_i)^2} \begin{bmatrix} \sum_{i=1}^{N} X_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} X_i y_i \\ -\sum_{i=1}^{N} X_i \sum_{i=1}^{N} y_i + N\sum_{i=1}^{N} X_i y_i \end{bmatrix}$$

$$b_2 = \frac{N\sum_{i=1}^{N} X_i y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} y_i}{N\sum_{i=1}^{N} X_i^2 - (\sum_{i=1}^{N} X_i)^2} = \frac{\sum X_i y_i - \frac{1}{N}(N\bar{X})(N\bar{Y})}{\sum X_i^2 - \frac{1}{N}\times(N\bar{X})^2} = \frac{\sum X_i y_i - N\bar{X}\bar{Y}}{\sum X_i^2 - N\bar{X}^2} = \boxed{\frac{\sum(X_i-\bar{X})(y_i-\bar{y})}{\sum(X_i-\bar{X})^2}}$$

formula 2.7

$$b_1 = \frac{\sum_{i=1}^{N} X_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} X_i \sum_{i=1}^{N} X_i y_i}{N\sum_{i=1}^{N} X_i^2 - (\sum_{i=1}^{N} X_i)^2} = \frac{\sum X_i^2 N\bar{y} - N\bar{X}\sum X_i y_i}{N\sum X_i^2 - N^2\bar{X}^2}$$

$$= \frac{\sum X_i^2 \bar{y} - \bar{X}\sum X_i y_i}{\sum(X_i-\bar{X})^2} = \frac{\sum X_i^2 \bar{y} - N\bar{X}^2\bar{y}}{\sum(X_i-\bar{X})^2} - \bar{X}\left(\frac{\sum X_i y_i - N\bar{X}\bar{Y}}{\sum(X_i-\bar{X})^2}\right) = \boxed{\bar{y} - \bar{X}b_2}$$

formula 2.8

$\bar{y}$ $\qquad$ b_2

2. Let $K = 2$, show that $cov(b_1, b_2)$ in p. 30 of slides in Ch5 reduces to the formula

2. $Var(b) = \sigma^2(X'X)^{-1}$

由上一題的結果 可得 $\sigma^2(X'X)^{-1} = \frac{\sigma^2}{N\sum X_i^2 - (\sum X_i)^2}\begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & N \end{bmatrix} = \begin{bmatrix} \dfrac{\sigma^2 \sum X_i^2}{N\sum X_i^2 - (\sum X_i)^2} & \dfrac{-\sigma^2 \sum X_i}{N\sum X_i^2 - (\sum X_i)^2} \\ \dfrac{-\sigma^2 \sum X_i}{N\sum X_i^2 - (\sum X_i)^2} & \dfrac{\sigma^2 N}{N\sum X_i^2 - (\sum X_i)^2} \end{bmatrix}$

$Var(b_1|x)$ $\qquad$ $Cov(b_1,b_2|x)$

$Cov(b_2,b_1|x)$ $\qquad$ $Var(b_2|x)$

$$Var(b_1|x) = \frac{\sigma^2 \sum X_i^2}{N\sum X_i^2 - (\sum X_i)^2} = \frac{\sigma^2 \sum X_i^2}{N\sum X_i^2 - N^2\bar{X}^2} = \frac{\sigma^2 \sum X_i^2}{N(\sum X_i^2 - N\bar{X}^2)}$$

$$= \sigma^2\left[\frac{\sum X_i^2}{N(\sum(X_i-\bar{X})^2)}\right]$$

$$Var(b_2|x) = \frac{\sigma^2 N}{N\sum X_i^2 - (\sum X_i)^2} = \frac{\sigma^2 N}{N\sum X_i^2 - N^2\bar{X}^2} = \sigma^2 \frac{1}{\sum X_i^2 - N\bar{X}^2} = \frac{\sigma^2}{\sum(X_i-\bar{X})^2}$$

$$Cov(b_1,b_2|x) = \frac{-\sigma^2 \sum X_i}{N\sum X_i^2 - (\sum X_i)^2} = -\sigma^2\frac{N\bar{X}}{N\sum X_i^2 - N^2\bar{X}^2} = \sigma^2\left[\frac{-\bar{X}}{\sum(X_i-\bar{X})^2}\right]$$

5.3 a. Fill in the following blank spaces that appear in this table.

Dependent Variable:WALC				
Included observations:1200				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.4515	2.2019	**0.6594**	0.5099
ln(TOTEXP)	2.7648	**0.4840**	5.7103	0.0000
NK	**-1.4549**	0.3695	-3.9376	0.0001
AGE	-0.1503	0.0235	-6.4019	0.0000
R-squared	**0.0575**	Mean dependent var	6.19434	
S.E. of regression	**6.2167**	S.D. dependent var	6.39547	
Sum squared resid	46221.62			

Ans.

i. The t-statistic for b_1=1.4515/2.2019=0.6594
ii. The standard error for b_2=2.7648/5.7103=0.4840
iii. The estimate b_3=0.3695×(-3.9376)= -1.4549
iv. $R^2 = 1 - 46221.62/(6.39547^2 \times 1199) = 0.0575$
v. $\hat{\sigma} = [\dfrac{46221.62}{(1200-4)}]^{\frac{1}{2}} = 6.2167$

b. Interpret each of the estimates b_2, b_3, and b_4

Ans.

$\beta_2 = 2.7648$：For every 1% increase in total household expenditure, the budget share spent on alcohol increases by approximately 0.027648 percentage points .

$\beta_3 = -1.4549$：Each additional child in the household is associated with a decrease of approximately 1.4549 percentage points in the alcohol budget share.

$\beta_4 = -0.1503$：For each additional year of the household head's age, the alcohol budget share decreases by approximately 0.1503 percentage points.

c. Compute a 95% interval estimate for β_4.What does this interval tell you?

Ans.

-0.1503-1.96×0.0235=-0.1964 ,-0.1503+1.96×0.0235=-0.1042
The 95% interval estimate for β_4 is [-0.1964,-0.1042]
For each additional year of the household head's age, the alcohol budget share decreases by an amount between 0.1042 and 0.1964 percentage points.

d. Are each of the coefficient estimates significant at a 5% level? Why?

Ans.

Except for intercept, all coefficient estimates are significant at a 5% level since their p-values are all less than 0.05.

e.Test the hypothesis that the addition of an extra child decreases the mean budget share of alcohol by 2 percentage points against the alternative that the decrease is not equal to 2 percentage points. Use a 5% significance level.

Ans.

null hypothesis : $H_0 = -2$

Alternative hypothesis : $H_1 \neq -2$

test statistic $t = \dfrac{-1.4549 - (-2)}{0.3695} = 1.4752$

We fail to reject H_0 since $|1.4752| < 1.96$

Each additional child in the household is leads to a decline in the alcohol budget share that is different from two percentage points.

5.23 a. What signs would you expect on the cofficients β_2, β_3, and β_4?

Ans.

β_2(QUANT):The expected sign of β_2 is **negative**, because larger quantities in a single sale are typically associated with lower prices per gram.

β_3 (QUAL):The expected sign of β_3 is **positive**, since higher purity (quality) cocaine should command a higher price.

β_4 (TREND):The expected sign of β_4 is **uncertain.** If supply increases over time or law enforcement becomes more effective, the market price might decline. If demand increases over time, the market price might rise.

b. Use your computer software to estimate the equation. Report the results and interpret the coefficient estimates. Have the signs turned out as you expected?

Ans.

$\widehat{PRICE} = 90.8467 - 0.06QUANT + 0.1162QUAL - 2.3546TREND$

The estimated values for β_2 = -0.0600 Every additional gram sold in a transaction ,the price per gram decreases by about $0.06. This matches our expectation of a negative sign.

The estimated values for β_3 = 0.01162 Every additional number of grams of cocaine in a given sale, the price per gram increases by about $0.1162. This matches our expectation of a positive sign.

The estimated values for β_4 = - 2.3546 Every additional year, the price per gram decreases by about $2.3546.This imply supply increases more effective then demand.

```
Call:
lm(formula = price ~ quant + qual + trend, data = cocaine)

Residuals:
    Min      1Q  Median      3Q     Max
-43.479 -12.014  -3.743  13.969  43.753

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 90.84669    8.58025  10.588 1.39e-14 ***
quant       -0.05997    0.01018  -5.892 2.85e-07 ***
qual         0.11621    0.20326   0.572   0.5700
trend       -2.35458    1.38612  -1.699   0.0954 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.06 on 52 degrees of freedom
Multiple R-squared:  0.5097,    Adjusted R-squared:  0.4814
```

c.What proportion of variation in cocaine price is explained jointly by variation in quantity, quality, and time?
Ans.
$R^2 = 0.5097$

d.It is claimed that the greater the number of sales, the higher the risk of getting caught. Thus, sellers are willing to accept a lower price if they can make sales in larger quantities. Set up H_0 and H_1 that would be appropriate to test this hypothesis. Carry out the hypothesis test.
Ans.
null hypothesis: $H_0 : \beta_2 \geq 0$
Alternative hypothesis: $H_0 : \beta_2 < 0$
test statistic $= -5.892$，$|-5.892| > 1.675 = t_{(0.95,52)}$
Reject null hypothesis, means that an increase in the quantity of cocaine sold in a given transaction is associated with a decrease in the price per gram.

e. Test the hypothesis that the quality of cocaine has no influence on expected price against the alternative that a premium is paid for better-quality cocaine.
Ans.
null hypothesis: $H_0 : \beta_3 = 0$
Alternative hypothesis: $H_0 : \beta_3 > 0$
test statistic $= 0.572$，$|-5.892| < 1.675 = t_{(0.95,52)}$
We can't reject null hypothesis, means that a premium isn't paid for better-quality cocaine.

f. What is the average annual change in the cocaine price? Can you suggest why price might be changing in this direction
Ans.
The average annual change in the price of cocaine is a decrease of approximately $2.35 per gram per year over the period from 1984 to 1991. In other words, each year the expected price of cocaine falls by about $2.35 on average.
I think he possible reason are increased supply or technological improvements.
If more cocaine entered the market during this period, possibly due to expanded production or trafficking, prices would naturally fall due to higher availability and if technological improvements cocaine became easier or cheaper to produce and transport, producers may have been willing to accept lower prices while still maintaining profitability.