HW 0324 111700027 林育炘

1. Let $k=2$, and $Y_i = b_1 + b_2 X_i + e_i$, $b_2 = \frac{\sum(X_i-\bar{X})(y_i-\bar{y})}{\sum(X_i-\bar{X})^2}$, $b_1 = \bar{y} - b_2\bar{X}$

$$X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} \Rightarrow X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix}$$

$$\text{so } b = (X'X)^{-1}(X'Y) = \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix} \begin{bmatrix} \sum y_i \\ \sum X_i y_i \end{bmatrix} = \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 \sum y_i - \sum X_i \sum X_i y_i \\ -\sum X_i \sum y_i + n\sum X_i y_i \end{bmatrix}$$

$$b_2 = \frac{n\sum X_i y_i - \sum X_i y_i}{n\sum X_i^2 - (\sum X_i)^2} = \frac{\sum X_i y_i - \frac{\sum X_i y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}} = \frac{\sum(X_i-\bar{X})(y_i-\bar{y})}{\sum(X_i-\bar{X})^2}$$

$$b_1 = \frac{\sum X_i^2 y_i - \sum X_i \sum X_i y_i}{n\sum X_i^2 - (\sum X_i)^2} = \frac{\bar{y} n \sum X_i^2 - n\bar{X}\cdot\sum X_i y_i}{n\sum X_i^2 - n\bar{X}^2} = \frac{\sum X_i^2 y_i - \sum X_i \bar{X}^2 \bar{y}}{\sum(X_i-\bar{X})^2} - \bar{X}\frac{\sum X_i y_i - n\bar{X}\bar{y}}{\sum(X_i-\bar{X})^2}$$

$$= \bar{y} - b_2\bar{X}$$

2. $$\text{var}(b) = \sigma^2(X'X)^{-1} = \sigma^2 \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix} = \frac{\sigma^2}{n\sum(X_i-\bar{X})^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\sigma^2 \sum X_i^2}{n\sum(X_i-\bar{X})^2} & \frac{\sigma^2(-\bar{X})}{\sum(X_i-\bar{X})^2} \\ \frac{\sigma^2(-\bar{X})}{\sum(X_i-\bar{X})^2} & \frac{\sigma^2}{\sum(X_i-\bar{X})^2} \end{bmatrix}$$

therefore: $\text{var}(b_1|X) = \sigma_{11}^2 = \frac{\sigma^2 \sum X_i^2}{n\sum(X_i-\bar{X})^2}$

$\text{var}(b_2|X) = \sigma_{22}^2 = \frac{\sigma^2}{\sum(X_i-\bar{X})^2}$

$\text{cov}(b_1, b_2|X) = \sigma_{12}^2 = \frac{\sigma^2(-\bar{X})}{\sum(X_i-\bar{X})^2}$

5.3

a. (i) $t_{stat} = \frac{1.4515}{2.2019} = 0.6592$

(ii) $se(b_2) = \frac{2.7648}{5.7103} = 0.4842$

(iii) $b_3 = 0.3695 \times (-3.9376) = -1.4549$

(iv) $SST = 6.39547(1200-1) = 49041.5418$

$R^2 = 1 - \frac{46221.62}{49041.5418} = 0.0575$

(v) $\hat{\sigma} = \sqrt{\frac{46221.62}{1200-4}} = 6.217$

b. $b_2 = 2.7648$: 1 unit increase in total expenditure will lead to 2.7648 unit increase on share of alcohol expenditure. when other factors are held.

$b_3 = -1.45494$: 1 unit increase in number of children will lead to 1.4549 unit decrease on share of alcohol expenditure. when other factors are held.

$b_4 = -0.1503$: 1 unit increase in age of household will lead to 0.1503 unit decrease on share of alcohol expenditure. when other factors are held.

c. $CI = [-0.1503 - 1.96 \times 0.0235, -0.1503 + 1.96 \cdot 0.0235] = [-0.1964, -0.1042]$

There is a 95% confidence that increase 1 year of age the share of alcohol expenditure is estimated decrease between -0.1964 with -0.1042.

d. Except for the intercept

All coefficient are significantly different from 0 at 5% level.

e. $H_0: \beta_3 = -2$

$H_1: \beta_3 \neq -2$, $t = \frac{-1.4575 + 2}{0.3695} = 1.475 < 1.96$

$\Rightarrow$ do not reject H_0. There is no sufficient evidence to say that the decrease is different from 2%

5.24

a. β_2 is negative, because larger quantity will make price lower

β_3 is positive, because larger quality will make price larger

β_4 is not sure, because increasing supply leads fall in price and increasing demand leads fall in price

b.

```
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 
F-statistic: 18.02 on 3 and 52 DF,  p-value: 3.806e-08
```

quantity increase 1 unit ⇒ price decrease 0.05997 with other factor is held.

quality increase 1 unit ⇒ price increase 0.11621 with other factor is held.

time increase 1 unit ⇒ price decrease 2.35458 with other factor is held.

⇒ All the signs are same as expected.

$\widehat{PRICE} = 90.84669 - 0.05997 QUANT + 0.11621 QUAL - 2.3478 TREND + e$

c. $R^2 = 0.5097$

d. $H_0: \beta_2 \geq 0$

查表 $t_{0.05,54}$

$H_1: \beta_2 < 0$, $t = -5.892 < -1.6736$

⇒ reject H_0, there is sufficient evidence to indicate that if quality gets larger and the price will be lower

```
> qt(0.05, df = 54)
[1] -1.673565
> qt(0.95, df = 54)
[1] 1.673565
```

e. $H_0: \beta_3 \leq 0$

查表 $t_{0.95,54}$

$H_1: \beta_3 > 0$, $t = 0.572 < 1.6736$

⇒ do not reject H_0, there is no sufficient evidence to show that if quality gets larger the price will be larger

f. average $= -2.3548$

it may because the development of technology.