HW Q1 / Q2 :

* 試證當 $k=2$ 時，b 和 CH2 提到關於 b_1, b_2 的公式及抽樣分配結果相同

$$Y = X\beta + e$$

$\because k=2$

其中 Y 為 $N\times 1$ 之 vector，且 $X = \begin{pmatrix} 1 & X_{12} \\ 1 & X_{22} \\ \vdots & \vdots \\ 1 & X_{N2} \end{pmatrix}$

X $N\times K$

β $K\times 1$

e $N\times 1$

$$X'X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{12} & X_{22} & \cdots & X_{N2} \end{pmatrix} \begin{pmatrix} 1 & X_{12} \\ 1 & X_{22} \\ \vdots & \vdots \\ 1 & X_{N2} \end{pmatrix}$$

$$= \begin{pmatrix} N & \sum_{i=1}^{N} X_{i2} \\ \sum_{i=1}^{N} X_{i2} & \sum_{i=1}^{N} X_{i2}^2 \end{pmatrix}$$

$$(X'X)^{-1} = \frac{1}{N\sum_{i=1}^{N} X_{i2}^2 - \left(\sum_{i=1}^{N} X_{i2}\right)^2} \begin{pmatrix} \sum_{i=1}^{N} X_{i2}^2 & -\sum_{i=1}^{N} X_{i2} \\ -\sum_{i=1}^{N} X_{i2} & N \end{pmatrix}$$

↓

即行列式 $=\Delta$

$$X'Y = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{12} & x_{22} & \cdots & x_{N2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{i2} y_i \end{pmatrix}$$

$$\tilde{b} = (X'X)^{-1}(X'Y) = \frac{1}{\Delta} \begin{pmatrix} \sum_{i=1}^{N} x_{i2}^2 & -\sum_{i=1}^{N} x_{i2} \\ -\sum_{i=1}^{N} x_{i2} & N \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{i2} y_i \end{pmatrix}$$

$$= \frac{1}{\Delta} \begin{pmatrix} \sum_{i=1}^{N} x_{i2}^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_{i2} \sum_{i=1}^{N} x_{i2} y_i \\ -\sum_{i=1}^{N} x_{i2} \sum_{i=1}^{N} y_i + N \sum_{i=1}^{N} x_i y_i \end{pmatrix}$$

且 $\operatorname{Cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \dfrac{\sigma^2}{\Delta} \begin{pmatrix} \sum_{i=1}^{N} x_{i2}^2 & -\sum_{i=1}^{N} x_{i2} \\ -\sum_{i=1}^{N} x_{i2} & N \end{pmatrix}$

因此 $\begin{cases} b_1 = \dfrac{1}{\Delta} \left(\sum_{i=1}^{N} x_{i2}^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_{i2} \sum_{i=1}^{N} x_{i2} y_i \right) \\ b_2 = \dfrac{1}{\Delta} \left(-\sum_{i=1}^{N} x_{i2} \sum_{i=1}^{N} y_i + N \sum_{i=1}^{N} x_i y_i \right) \end{cases}$

$$
\begin{cases}
\operatorname{Var}(b_1) = \dfrac{\sigma^2}{\Delta} \displaystyle\sum_{i=1}^{N} x_i^2 \\[2ex]
\operatorname{Var}(b_2) = \dfrac{\sigma^2}{\Delta} N
\end{cases}
$$

$$
\tilde{b} \sim N\left(\hat{\beta}, \ \sigma^2 \begin{pmatrix} \dfrac{\sum x_i^2}{\Delta} \\[2ex] \dfrac{N}{\Delta} \end{pmatrix} \right)
$$

5.3 Consider the following model that relates the percentage of a household's budget spent on alcohol *WALC* to total expenditure *TOTEXP*, age of the household head *AGE*, and the number of children in the household *NK*.

$$WALC = \beta_1 + \beta_2 \ln(TOTEXP) + \beta_3 NK + \beta_4 AGE + e$$

This model was estimated using 1200 observations from London. An incomplete version of this output is provided in Table 5.6.

$= N$

TABLE 5.6 **Output for Exercise 5.3**

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

a. Fill in the following blank spaces that appear in this table.
  i. The *t*-statistic for b_1.
  ii. The standard error for b_2.
  iii. The estimate b_3.
  iv. R^2.
  v. $\hat{\sigma}$.

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

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

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

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.

a.

(i) $t_1 = \dfrac{1.4515}{2.2019} = 0.6592$

(ii) $SE(b_2) = \dfrac{2.7648}{5.7103} = 0.48442$

(iii) $\hat{b_3} = -3.9376 \times 0.3695 = -1.4558$

(iv) 先求 $TSS = (6.39547)^2 \times (1{,}200 - 1) = 49{,}024.9$

且已知 $SSE = 46{,}221.62$

$\Rightarrow R^2 = 1 - \dfrac{46{,}221.62}{49{,}024.9} = 0.0573$

(a) $\hat{\sigma} = \sqrt{MSE} = \sqrt{\frac{SSE}{n-K}} = \sqrt{\frac{46,221.62}{1,200-4}} = 6.216$

b.

$b_2 = 2.7648$：表示家庭總支出每增加1%，酒精支出占比會增加 2.7648 百分點

$b_3 = -1.4558$：表示每多一個小孩，家庭酒精支出占比少 1.46 百分點

$b_4 = -0.1503$：表示家庭負責人每年年齡增加，酒精支出占比減少 0.1503 百分點

c.

$$\begin{cases} \hat{\beta}_4 = -0.1503 \\ SE(\hat{\beta}_4) = 0.0235 \end{cases}$$

在 $N=1,200$ 下，適用大樣本，採用Z分配

$$C.I. = \hat{\beta}_4 \pm z_{0.025} \times SE(\hat{\beta}_4)$$

$$= -0.1503 \pm 1.96 \times 0.0235$$

$$= (-0.1963, -0.1043)$$

<u>解釋</u>：我們有95%的信心認為，家庭負責人年齡

對酒精支出占比的影響係數落在
-0.1963 到 -0.1043 之間

d、

在 $\alpha = 0.05$ 下，

$\hat{\beta}_1$ 之 p-value $= 0.5099 > 0.05 \Rightarrow$ 不顯著
$\hat{\beta}_2$ 之 p-value $= 0.0000 < 0.05 \Rightarrow$ 顯著
$\hat{\beta}_3$ 之 p-value $= 0.0001 < 0.05 \Rightarrow$ 顯著
$\hat{\beta}_4$ 之 p-value $= 0.0000 < 0.05 \Rightarrow$ 顯著

$\Rightarrow$ 除常數項以外，皆顯著

e、

即 ① $\begin{cases} H_0: \beta_3 = -2 \\ H_a: \beta_3 \neq -2 \end{cases}$

② $\alpha = 0.05$

③ $z = \dfrac{\hat{\beta}_3 - (-2)}{SE(\hat{\beta}_3)}$ （∵大樣本下）

④ $z_1 = \dfrac{-1.4558 - (-2)}{0.3695} = 1.4725$

⑤ $z_1 = 1.4725 < z_{0.025} = 1.96$

$\therefore$ do not reject H_0

$\Rightarrow$ 即在 $\alpha = 0.05$ 下，無法否定 NK係數 $= 2$，即

無法否定每多一位孩子會使支出減少2個百分點