Q5.6

5.6 Suppose that, from a sample of 63 observations, the least squares estimates and the corresponding estimated covariance matrix are given by

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} \qquad \widehat{\operatorname{cov}}(b_1, b_2, b_3) = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

Using a 5% significance level, and an alternative hypothesis that the equality does not hold, test each of the following null hypotheses:

a. $\beta_2 = 0$
b. $\beta_1 + 2\beta_2 = 5$
c. $\beta_1 - \beta_2 + \beta_3 = 4$

(a)

H0: b2 = 0, H1: b2 != 0

SE(b2) = sqrt(Var(b2)) = 2

t = (3-0)/2 = 1.5 < $t_{0.025, 60}$ = 2.00 (查表)

t 不在拒絕域內，所以無法拒絕 H0

(b)

H0: b1+2b2 = 5, H1: b1+2b2 != 5

t = (L^Tb - c) / sqrt(L^T cov_matrix L), L = [1; 2; 0], c = 5

L^Tb = 1(2)+2(3)+0(−1)=2+6+0=8

L^T cov_matrix =(1×3+2×(−2)+0×1,1×(−2)+2×4+0×0,1×1+2×0+0×3)

=(3−4+0,−2+8+0,1+0+0) =(−1,6,1)= (3 - 4 + 0, -2 + 8 + 0, 1 + 0 + 0)

= (-1, 6, 1)=(3−4+0,−2+8+0,1+0+0) = (−1,6,1)

(−1,6,1) * L = (−1×1)+(6×2)+(1×0)=−1+12+0 = 11

t = (8-5) / sqrt(11) = 0.905 < $t_{0.025, 60}$ = 2.00 (查表)

t 不在拒絕域內，所以無法拒絕 H0，表示 $b_1 + 2b_2$可能等於 5

(c)

H0: b1 − b2 + b3 = 4, H1: b1 − b2 + b3 != 4

$t = (L^T b - c) / \text{sqrt}(L^T \text{cov\_matrix } L)$, $L = [1; -1; 1]$, $c = 4$

$L^T b = 1(2)+(-1)(3)+1(-1)=2-3-1 = -2$

L^T cov_matrix = (1×3+(−1)×(−2)+1×1,1×(−2)+(−1)×4+1×0,1×1+(−1)×0+1×3) =(3+2+1,−2−4+0,1+0+3)=(6,−6,4)= (3 + 2 + 1, -2 - 4 + 0, 1 + 0 + 3) = (6, -6, 4)=(3+2+1,−2−4+0,1+0+3) = (6,−6,4)

(6,−6,4) * L = (6×1)+(−6×−1)+(4×1)=6+6+4 = 16

$t = (-2-4) / \text{sqrt}(16) = -1.5 \quad < \quad t_{0.975, 60} = -2.00$ (查表)

t 不在拒絕域內，所以無法拒絕 H0，表示 $b_1 - b_2 + b_3$可能等於 4

Q5.31

(a)

```
Call:
lm(formula = time ~ depart + reds + trains, data = commute5)

Residuals:
     Min       1Q   Median       3Q      Max 
-18.4389  -3.6774  -0.1188   4.5863  16.4986 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  20.8701     1.6758  12.454  < 2e-16 ***
depart        0.3681     0.0351  10.487  < 2e-16 ***
reds          1.5219     0.1850   8.225 1.15e-14 ***
trains        3.0237     0.6340   4.769 3.18e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.299 on 245 degrees of freedom
Multiple R-squared:  0.5346,	Adjusted R-squared:  0.5289 
F-statistic: 93.79 on 3 and 245 DF,  p-value: < 2.2e-16
```

截距 b1：當 DEPART, REDS, TRAINS 都是 0 時（6:30 準時出門、無紅燈、無火車），預測所需時間。

b2：出門每晚 1 分鐘，預期通勤時間會增加（或減少）幾分鐘。

b3：每個紅燈預計增加多少分鐘。

b4：每列火車預計增加多少分鐘。

(b)

```
> confint(model, level = 0.95)
                2.5 %    97.5 %
(Intercept) 17.5694018 24.170871
depart       0.2989851  0.437265
reds         1.1574748  1.886411
trains       1.7748867  4.272505
```

信賴區間由窄到寬為 DEPART, REDS, TRAINS, Intercept，符合他們標準誤的大小順序，其中 DEPART、REDS 的區間相對較小(不到 1)，估計相對較準確。

(c)

H0: b3 >= 2, H1: b3 < 2

```
> if (abs(t_value) > abs(qt(0.05, df = 245))) {
+   print("reject H0")
+ }
[1] "reject H0"
> t_value
     reds
-2.583562
> p_value
       reds
0.005179509
```

t 值落在拒絕域中，所以拒絕 H0，支持每個紅燈的影響小於 2 分鐘

(d)

H0: b4 = 3, H1: b4 != 3

```
> t_value
    trains
0.03737444
> p_value
   trains
0.9702169
```

T 值未落在拒絕域，無法拒絕 H0

(e)

H0: 30*b2 >= 10, H1: 30*b2<10

```
> t_value
   depart
0.9911646
> p_value
   depart
0.8387085
```

沒有落在拒絕域，所以拒絕 H0

(f)

H0: b4 – 3*b3 >= 0, H1: b4 – 3*b3 < 0

```
t =  -1.825027
```

落在拒絕域，拒絕 H0，表示 trains 的影響小於 reds 的三倍影響

(g)

E(TIME | X) = b1 + 30 * depart + 1*trains

H0: E(TIME | X) <= 45, H1: E(TIME | X) > 45

```
> t_value
          [,1]
[1,] -1.725964
> p_value
          [,1]
[1,] 0.9571926
```

無法拒絕 H0，Bill 有機會準時抵達

(h)

在不可遲到的假設下，最不想發生的事是以為會準時，實際遲到（Type I error），所以應該要將準時設為 H1

H0: E(TIME | X) >= 45, H1: E(TIME | X) < 45

反轉後，critical value = -1.6511。我們拒絕 H0，表示我們支持 Bill 能在 7:45 前抵達

Q5.33

(a)

```
Call:
lm(formula = log(wage) ~ educ + I(educ^2) + exper + I(exper^2) 
    I(educ * exper), data = cps5_small)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.6628 -0.3138 -0.0276  0.3140  2.1394 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      1.038e+00  2.757e-01   3.764 0.000175 ***
educ             8.954e-02  3.108e-02   2.881 0.004038 ** 
I(educ^2)        1.458e-03  9.242e-04   1.578 0.114855    
exper            4.488e-02  7.297e-03   6.150 1.06e-09 ***
I(exper^2)      -4.680e-04  7.601e-05  -6.157 1.01e-09 ***
I(educ * exper) -1.010e-03  3.791e-04  -2.665 0.007803 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4638 on 1194 degrees of freedom
Multiple R-squared:  0.3227,	Adjusted R-squared:  0.3198 
F-statistic: 113.8 on 5 and 1194 DF,  p-value: < 2.2e-16
```

educ^2 在 10%、5%、1% 顯著水準下都不顯著

其他變數全部在 1%顯著水準下顯著

(b)

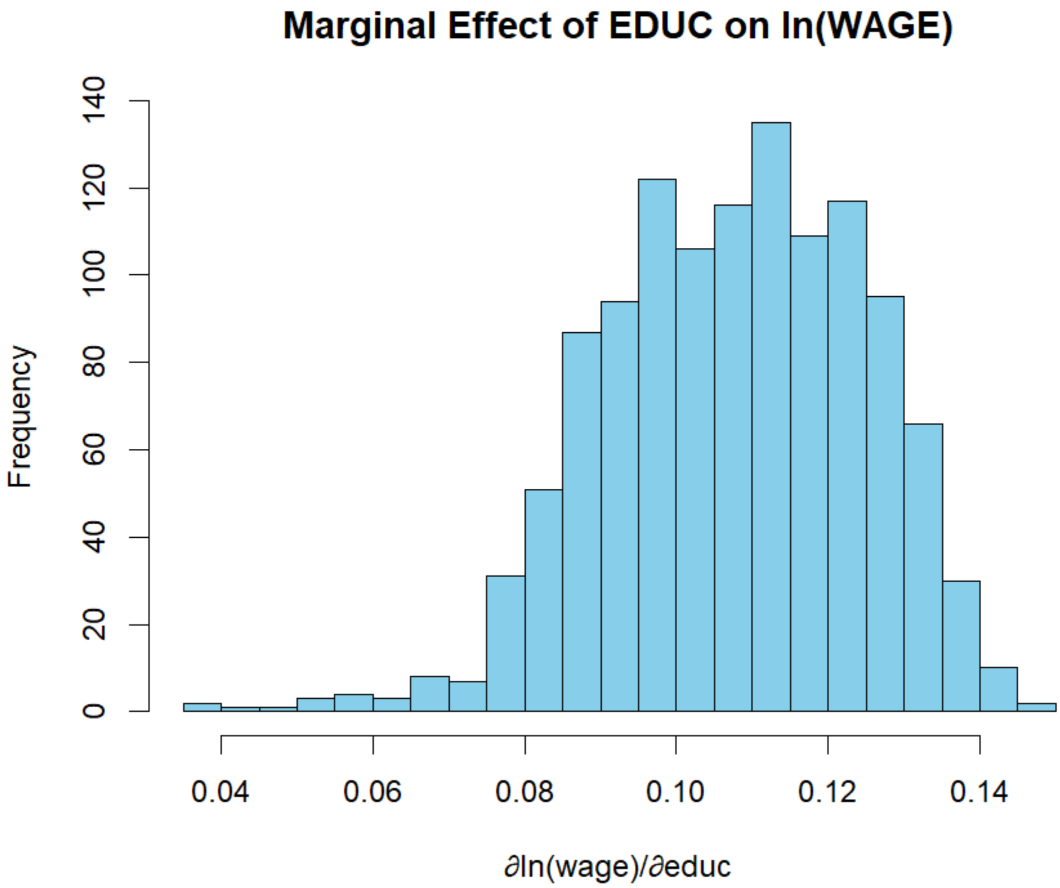
$\partial E[\ln(wage)]/\partial educ = b2 + 2b3 \cdot educ + b6 \cdot exper$

$= 0.0895 + 2(0.00146) \cdot educ - 0.00101 \cdot exper$

當教育年數（educ）增加時，$educ^2$的係數會讓邊際效果變大

與 exper 的交乘項（負）會使教育的效果因經驗較多而遞減

(c)




```
> quantile(marginal_effect_educ, probs = c(0.05, 0.5, 0.95))
        5%        50%        95%
0.08008187 0.10843125 0.13361880
```

集中在 0.9~0.13 之間

(d)

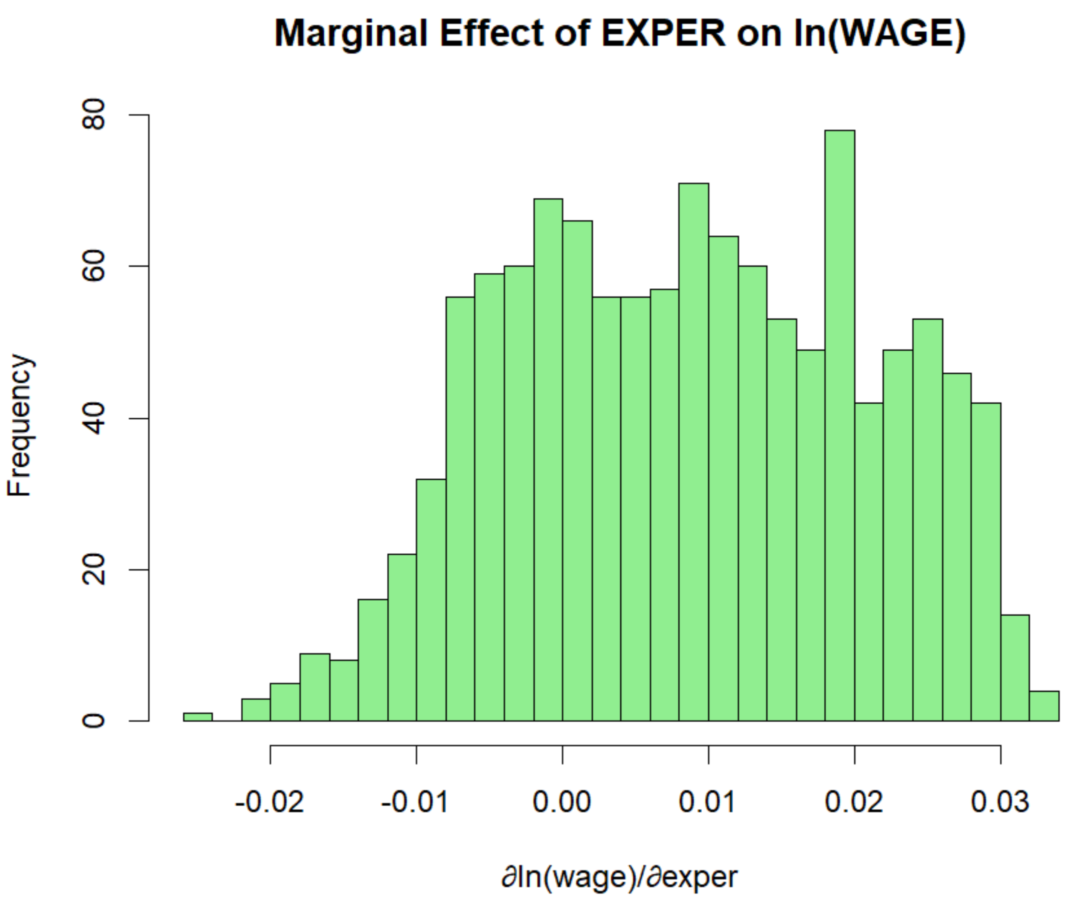
∂E[ln(wage)]/∂exper = b4 + 2b5·exper + b6·educ

= 0.04488 + 2(−0.000468)·exper − 0.00101·educ

隨著 educ 增加，邊際效應減少。

而隨著 exper 增加，邊際效應也是減少。

(e)




```
> quantile(marginal_effect_exper, probs = c(0.05, 0.5, 0.95))
          5%          50%          95% 
-0.010376212  0.008418878  0.027931151 
```

多集中在 0~0.02 之間

(f)

diff=E[ln(WAGE)David]−E[ln(WAGE)Svetlana]

=b2(17−16)+b3(172−162)+b4(8−18)+b5(82−182)+b6(17·8−16·18)= \beta_2(17 - 16) + \beta_3(17^2 - 16^2) + \beta_4(8 - 18) + \beta_5(8^2 - 18^2) + \beta_6(17 \cdot 8 - 16 \cdot 18)=b2(17−16)+b3(172−162)+b4(8−18)+b5(82−182)+b6 (17·8−16·18) =b2+33b3−10b4−260b5−152b6= \beta_2 + 33\beta_3 - 10\beta_4 - 260\beta_5 - 152\beta_6=b2+33b3−10b4−260b5−152b6

H0: Svetlana 的期望 log-wage ≥ David, diff≤0

H1: David 的期望 log-wage > Svetlana, diff>0

```
> t_stat
          [,1]
[1,] -1.669902
> p_value
          [,1]
[1,] 0.9523996
> qt(0.95, df = 1194)
[1] 1.646131
```

無法拒絕 H0

沒有足夠證據說明 David 的預期 log-wage 高於 Svetlana

(g)

New diff = b2+(172−162)b3+(16−26)b4+(162−262)b5+(17·16−16·26)b6

=b2+33b3−10b4−600b5−160b6= \beta_2 + 33\beta_3 - 10\beta_4 - 600\beta_5 - 160\beta_6=b2+33b3−10b4−600b5−160b6

H0: Svetlana 的期望 log-wage ≥ David, diff≤0

H1: David 的期望 log-wage > Svetlana, diff>0

```
> t_stat_new
          [,1]
[1,] 2.062365
> p_value_new
           [,1]
[1,] 0.01969445
> qt(0.05, df = 1194)
[1] -1.646131
```

落在拒絕域，經過 8 年後，David 在經驗上的提升讓他的薪資預期超過 Svetlana，顯著改變原本的情況

(h)

H0:(β4+2β5·17+β6·12) = (β4+2β5·11+β6·16)

H1:(β4+2β5·17+β6·12) != (β4+2β5·11+β6·16)

```
> cat("t = ", tva_L3, "\n")
t =  -1.027304
> cat("critical value = ", tcr_two, "\n")
critical value =  1.961953
```

沒有落在拒絕域，無法拒絕 H0

(i)

β4+2β5·EXPER+β6·16 < 0，此時邊際效應變為負數

```
b4 <- coef(model)["exper"]
b5 <- coef(model)["I(exper^2)"]
b6 <- coef(model)["I(educ * exper)"]

educ_val <- 16
current_exper <- 11
g <- -(b4 + educ_val*b6)/(2*b5) - current_exper
g
```

將對應的值帶入後得到

```
> g <- -(b4 + educ_val*b6)/(2*b5) - current_exper
> g
   exper
19.67706
```

還需 19.67706 年

```
var(g): 3.593728
 se(g): 1.895713
```

```
> qt(0.05, df = 1194)
[1] -1.646131
```

```
Extra years: 19.67706  with 95% interval estimates [ 15.95776 ,  23.39636 ]
```

信賴區間介於 15.958~23.396 之間