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

$\to \sigma_{11} = \sigma_1^2$

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

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) $\begin{cases} H_0: \beta_2 = 0 \\ H_a: \beta_2 \neq 0 \end{cases}$

$$t = \frac{b_2 - \beta_2}{SE(b_2)} = \frac{3-0}{\sqrt{Var(b_2)}} = \frac{3}{\sqrt{4}} = 1.5, \quad t_{(0.025, 63-3)} = 2.0.$$

$\begin{array}{r} 19 \\ -8 \\ \hline 11 \end{array}$

$t = 1.5 < t_{(0.025, 60)} = 2.0 \Rightarrow$ do not reject Ho

$\begin{array}{r} 64 \\ +9 \\ \hline 73 \end{array}$

(b) $\begin{cases} H_0: \beta_1 + 2\beta_2 = 5 \\ H_a: \beta_1 + 2\beta_2 \neq 5 \end{cases}$

$b_1 + 2b_2 = 2+2\times 3 = 8$

$$se(b_1 + 2b_2) = \sqrt{1^2 Var(b_1) + 2^2 Var(b_2) + 2\cdot 1\cdot 2\cdot Cov(b_1, b_2)} = \sqrt{3 + 4\times 4 + 4\times(-2)} = \sqrt{11}$$

$$t = \frac{(b_1 + 2b_2) - (\beta_1 + 2\beta_2)}{se(b_1 + 2b_2)} = \frac{8-5}{\sqrt{11}} = 0.9045$$

$\because t = 0.9045 < 2.003 \Rightarrow$ do not reject Ho

(c) $\begin{cases} H_0: \beta_1 - \beta_2 + \beta_3 = 4 \\ H_a: \beta_1 - \beta_2 + \beta_3 \neq 4. \end{cases}$

$b_1 - b_2 + b_3 = 2 - 3 + (-1) = -2$

$$se(b_1 - b_2 + b_3) = \sqrt{Var(b_1) + Var(b_2) + Var(b_3) - 2\cdot 1\cdot 1\cdot Cov(b_1, b_2) + 2\cdot 1\cdot 1\cdot Cov(b_1, b_3) - 2\cdot 1\cdot 1\cdot Cov(b_2, b_3)}$$

$$= \sqrt{10 - (-4) + 2 - 0} = \sqrt{16} = 4$$

$$t = \frac{(b_1 - b_2 + b_3) - (\beta_1 - \beta_2 + \beta_3)}{se(b_1 - b_2 + b_3)} = \frac{-2-4}{4} = -1.5 < -2.0 \Rightarrow \text{do not reject Ho.}$$

$(|t| < |t_{0/2, df=60}|)$

5.31 Each morning between 6:30 AM and 8:00 AM Bill leaves the Melbourne suburb of Carnegie to drive to work at the University of Melbourne. The time it takes Bill to drive to work (*TIME*), depends on the departure time (*DEPART*), the number of red lights that he encounters (*REDS*), and the number of trains that he has to wait for at the Murrumbeena level crossing (*TRAINS*). Observations on these variables for the 249 working days in 2015 appear in the file *commute5*. *TIME* is measured in minutes. *DEPART* is the number of minutes after 6:30 AM that Bill departs.

a. Estimate the equation

$$TIME = \beta_1 + \beta_2 DEPART + \beta_3 REDS + \beta_4 TRAINS + e$$

Report the results and interpret each of the coefficient estimates, including the intercept β_1.

b. Find 95% interval estimates for each of the coefficients. Have you obtained precise estimates of each of the coefficients?

c. Using a 5% significance level, test the null hypothesis that Bill's expected delay from each red light is 2 minutes or more against the alternative that it is less than 2 minutes.

d. Using a 10% significance level, test the null hypothesis that the expected delay from each train is 3 minutes against the alternative that it is not 3 minutes.

e. Using a 5% significance level, test the null hypothesis that Bill can expect a trip to be at least 10 minutes longer if he leaves at 7:30 AM instead of 7:00 AM, against the alternative that it will not be 10 minutes longer. (Assume other things are equal.)

f. Using a 5% significance level, test the null hypothesis that the expected delay from a train is at least three times greater than the expected delay from a red light against the alternative that it is less than three times greater.

g. Suppose that Bill encounters six red lights and one train. Using a 5% significance level, test the null hypothesis that leaving Carnegie at 7:00 AM is early enough to get him to the university on or before 7:45 AM against the alternative that it is not. [Carry out the test in terms of the expected time $E(TIME|\mathbf{X})$ where $\mathbf{X}$ represents the observations on all explanatory variables.]

h. Suppose that, in part (g), it is imperative that Bill is not late for his 7:45 AM meeting. Have the null and alternative hypotheses been set up correctly? What happens if these hypotheses are reversed?

(a)

```
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
```

β_1: 當Bill depart at 6:30AM, 且reds, trains = 0, 則 commute time = 20.87 minutes

β_2: 每晚1分鐘出發, commute time + 0.368 mins (22s)

β_3: 多1個 red → Time + 1.52 mins

β_4: 每多等一班 train, time + 3.02 mins

(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
```

→ not preise
→ preise
→ preise
→ relative not preise

(c)

$$\begin{cases} H_0: \beta_{reds} \geq 2 \\ H_a: \beta_{reds} < 2 \end{cases}$$

$$t = \frac{b_3 - \beta_3}{se(b_3)} = \frac{1.5219 - 2}{0.1850} = -2.5843 < -1.651 \Rightarrow$$ reject H_0 that 1 red light + 2 mins of times or more.

(d) $\begin{cases} H_0: \beta_4 = 3 \\ H_1: \beta_4 \neq 3 \end{cases}$ $t = \dfrac{3.0237 - 3}{0.6340} = 0.0374 < 1.6511 \Rightarrow$ do not reject Ho that delay time for waiting 1 train is 3 mins.

(e) $\begin{cases} H_0: 30\beta_2 \geq 10 \\ H_a: 30\beta_2 < 10 \end{cases} \Rightarrow \begin{cases} H_0: \beta_2 \geq \frac{1}{3} \\ H_a: \beta_2 < \frac{1}{3} \end{cases} \Rightarrow t = \dfrac{0.3861 - \frac{1}{3}}{0.0351} = 0.9905 > -1.6511$

$\Rightarrow$ do not reject Ho expect a trip be at least 10 mins longer if leaves 30mins earlier.

(f) $\begin{cases} H_0: \beta_4 \geq 3\beta_3 \\ H_a: \beta_4 < 3\beta_3 \end{cases}$ $t = \dfrac{3.0237 - 3 \times 1.5219}{0.842} = -1.8249 < -1.6511 \Rightarrow$ reject Ho a expect time from train delay is greater than 3 times than red lights.

(g) $\begin{cases} H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \leq 45 \\ H_a: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 > 45 \end{cases}$, $t = \dfrac{44.0629 - 45}{0.5392} = -1.736 < 1.6511 \Rightarrow$ reject Ho that 30min depart earlier + 6 reds + 1 train can arrive before 7:45 A.M.

(h) Having a commute time < 45 mins should be Ha since it's more imparative

$\begin{cases} H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \geq 45 \\ H_a: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 < 45 \end{cases}$

5.33 Use the observations in the data file *cps5_small* to estimate the following model:

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EDUC^2 + \beta_4 EXPER + \beta_5 EXPER^2 + \beta_6 (EDUC \times EXPER) + e$$

a. At what levels of significance are each of the coefficient estimates "significantly different from zero"?

b. Obtain an expression for the marginal effect $\partial E[\ln(WAGE)|EDUC, EXPER]/\partial EDUC$. Comment on how the estimate of this marginal effect changes as *EDUC* and *EXPER* increase.

c. Evaluate the marginal effect in part (b) for all observations in the sample and construct a histogram of these effects. What have you discovered? Find the median, 5th percentile, and 95th percentile of the marginal effects.

d. Obtain an expression for the marginal effect $\partial E[\ln(WAGE)|EDUC, EXPER]/\partial EXPER$. Comment on how the estimate of this marginal effect changes as *EDUC* and *EXPER* increase.

e. Evaluate the marginal effect in part (d) for all observations in the sample and construct a histogram of these effects. What have you discovered? Find the median, 5th percentile, and 95th percentile of the marginal effects.

f. David has 17 years of education and 8 years of experience, while Svetlana has 16 years of education and 18 years of experience. Using a 5% significance level, test the null hypothesis that Svetlana's expected log-wage is equal to or greater than David's expected log-wage, against the alternative that David's expected log-wage is greater. State the null and alternative hypotheses in terms of the model parameters.

g. After eight years have passed, when David and Svetlana have had eight more years of experience, but no more education, will the test result in (f) be the same? Explain this outcome?

h. Wendy has 12 years of education and 17 years of experience, while Jill has 16 years of education and 11 years of experience. Using a 5% significance level, test the null hypothesis that their marginal effects of extra experience are equal against the alternative that they are not. State the null and alternative hypotheses in terms of the model parameters.

i. How much longer will it be before the marginal effect of experience for Jill becomes negative? Find a 95% interval estimate for this quantity.

(a)

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	**
educ2	1.458e-03	9.242e-04	1.578	0.114855	
exper	4.488e-02	7.297e-03	6.150	1.06e-09	***
exper2	-4.680e-04	7.601e-05	-6.157	1.01e-09	***
educ_exper	-1.010e-03	3.791e-04	-2.665	0.007803	**

all variables are at 1% significant except for edu^2.

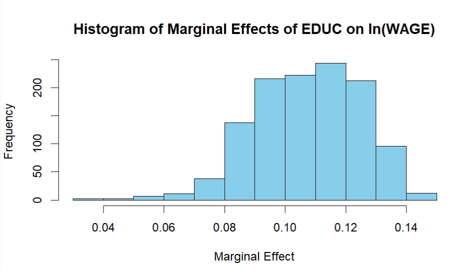
(b) $$\frac{\partial[\ln(WAGE|EDUC,EXPER)]}{\partial EDUC} = \beta_2 + 2\beta_3 \cdot EDUC + \beta_6 EDUC$$

$$= 0.08954 + 2 \times 0.001458\, EDUC + (-0.00101) \times EXPER$$

when: EDUC↑, ME↑

EXPER↑, ME↓

(C)



Median: 0.1.84

5th Percentile: 0.0801

95th Percentile: 0.1336

bell-shaped且左偏，大多数人的ME為8%~14%，

且 almost no negative effect.

(d) $$\frac{\partial E(\ln(WAGE))}{\partial EXPER} = \beta_4 + 2\beta_5 EXPER + \beta_6 EDUC$$

$$= 0.04488 + 2 \times (-0.000468)(EXPER) + (-0.00101) \cdot EDUC$$

when: EXPER↑, ME↓

EDUC↑, ME↓

(f) $\begin{cases} H_0: \beta_1+17\beta_2+17^2\beta_3+8\beta_4+8^2\beta_5+8\times17\beta_6 \le \beta_1+16\beta_2+16^2\beta_3+18\beta_4+18^2\beta_5+18\times16\beta_6 \\ H_a: \beta_1+17\beta_2+17^2\beta_3+8\beta_4+8^2\beta_5+8\times17\beta_6 > \beta_1+16\beta_2+16^2\beta_3+18\beta_4+18^2\beta_5+18\times16\beta_6 \end{cases}$

$t = \dfrac{-0.09588-0}{0.057489} = -1.67 < 1.646$ ⇒ do not reject H_0 that Svetlana's log-wage is = or > than David's.

(g) $\begin{cases} H_0: \beta_1+17\beta_2+17^2\beta_3+16\beta_4+16^2\beta_3+16\times17\beta_6 \le \beta_1+16\beta_2+16^2\beta_3+26\beta_4+26^2\beta_5+16\times26\beta_6 \\ H_a: \beta_1+17\beta_2+17^2\beta_3+16\beta_4+16^2\beta_3+16\times17\beta_6 > \beta_1+16\beta_2+16^2\beta_3+26\beta_4+26^2\beta_5+16\times26\beta_6 \end{cases}$

$t = 2.0627 > 1.646$ ⇒ reject H_0 that Svetlana's log-wage is = or > than David's.

(h) ME of EDUC $= \beta_4 + 2\beta_5 EXPER + \beta_6 EDUC$

$H_0: \beta_4+2\times17\beta_5+12\beta_6 = \beta_4+2\times11\beta_5+16\beta_6$

$H_a: \beta_4+2\times17\beta_5+12\beta_6 \ne \beta_4+2\times11\beta_5+16\beta_6$

$t = -1.02713 > -1.962$ ⇒ do not reject H_0 that their ME are equal.

(i) $\beta_4 + 2\beta_5(11+x) + 16\beta_6 = 0$

$x = 19.677$. → 19.667 more years than his EXPER ME will be negative.

```
> cat("Jill 需要額外增加的經驗估計:", x_hat, "年\n")
Jill 需要額外增加的經驗估計: 19.67706 年
> cat("95% 信賴區間: (", CI_lower, ",", CI_upper, ")\n")
95% 信賴區間: ( 15.96146 , 23.39265 )
```

95% C.I.