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. $\begin{cases} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{cases}$

$t = \dfrac{b_2 - \beta_2}{\sqrt{se(b_2)}} = \dfrac{3-0}{\sqrt{4}} = \dfrac{3}{2}$

$t_{0.975,60} = 2 \Rightarrow RR: \{t: t > 2 \text{ or } t < -2\}$

$\Rightarrow t = 1.5 < t_{0.975,60} \Rightarrow$ fails to reject H_0

b. $\begin{cases} H_0: \beta_1 + 2\beta_2 = 5 \\ H_1: \beta_1 + 2\beta_2 \neq 5 \end{cases}$

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

$Var(b_1 + 2b_2) = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

$= \begin{bmatrix} (3-4+0) & (-2+8+0) & (1+0+0) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

$= \begin{bmatrix} -1 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = 11$

$t = \dfrac{8-5}{\sqrt{11}} = 0.9045 \qquad t_{0.975,60} = 2$

$RR: \{t: t > 2 \text{ or } t < -2\} \Rightarrow$ fails to reject H_0

c. $\begin{cases} H_0: \beta_1 - \beta_2 + \beta_3 = 4 \\ H_1: \beta_1 - \beta_2 + \beta_3 \neq 4 \end{cases}$

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

$Var(b_1 - b_2 + b_3) = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} (3+2+1) & (-2-4+0) & (1-0+3) \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} 6 & -6 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 16$

$t = \dfrac{-2-4}{\sqrt{16}} = -1.5 \qquad t_{0.975,60} = 2$

$RR: \{t: t > 2 \text{ or } t < -2\} \Rightarrow$ fails to reject H_0.

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.

```
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
```

$\beta_1 = 20.8701$ ：當 depart, reds 和 trains 皆等於零時，通勤時間為 20.8701 分

$\beta_2 = 0.3681$ ：當 reds 和 trains 維持不變下，depart 每增加一分鐘，通勤時間增加 0.3681 分

$\beta_3 = 1.5291$ ：當 depart 和 trains 維持不變下，reds 每增加一次，通勤時間增加 1.5291 分

$\beta_4 = 3.0237$ ：當 depart 和 reds 維持不變下，trains 每增加一次，通勤時間增加 3.0237 分

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
```

這些區間都相對窄，因此這些係數的估計相對準確。

c.
$$\begin{cases} H_0: \beta_3 \geq 2 \\ H_1: \beta_3 < 2 \end{cases}$$

```
> qt(0.05, 245)
[1] -1.651097
```
$\Rightarrow RR: \{t: t < -1.6511\}$

```
> cat(t_value_reds)
-2.583562
```
< -1.6511

$\Rightarrow$ rejects H_0

$\Rightarrow$ The expected delay from reds is less than 2 minutes.

d.
$$\begin{cases} H_0: \beta_4 = 3 \\ H_1: \beta_4 \neq 3 \end{cases}$$

```
> qt(0.05, 245)
[1] -1.651097
```
$\Rightarrow RR: \{t: t > 1.6511 \text{ or } t < -1.6511\}$

```
> cat(t_value_trains)
0.03737444
```
< 1.6511

$\Rightarrow$ fails to reject H_0

$\Rightarrow$ The expect delay from each trains is 3 minutes.

e.
$$\begin{cases} H_0: \beta_2 \times 30 \geq 10 \\ H_1: \beta_2 \times 30 < 10 \end{cases} \Rightarrow \begin{cases} H_0: \beta_2 \geq 0.3333 \\ H_1: \beta_2 < 0.3333 \end{cases}$$

```
> cat(t_value_depart)
0.9921142> qt(0.05, 245)
[1] -1.651097
```
$\Rightarrow RR: \{t: t < -1.6511\}$

$\Rightarrow$ fails to reject H_0

$\Rightarrow$ Delaying departure time by 30 minutes increase expected travel time at lease by 10 minutes.

f.
$$\begin{cases} H_0: \beta_4 \geq 3\beta_3 \\ H_1: \beta_4 < 3\beta_3 \end{cases} \Rightarrow \begin{cases} H_0: \beta_4 - 3\beta_3 \geq 0 \\ H_1: \beta_4 - 3\beta_3 < 0 \end{cases}$$

```
> cat(t_value_f)
-1.825027> qt(0.05, 245)
[1] -1.651097
```
$\Rightarrow RR: \{t: t < -1.6511\}$

$\Rightarrow$ Reject H_0

$\Rightarrow$ The expected delay from a train is less than three times the delay from a red light.

g.
$$\begin{cases} H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \leq 45 \\ H_1: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 > 45 \end{cases}$$

```
> cat(t_value_g)
-1.725964> qt(0.95, 245)
[1] 1.651097
```
$\Rightarrow RR = \{t: t > 1.651097\}$

$\Rightarrow$ reject H_0

$\Rightarrow$ Bill will get to University after 7:45 A.M.

h.
$$\begin{cases} H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \geq 45 \\ H_1: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 < 45 \end{cases}$$

$RR: \{t: t < -1.651097\} \Rightarrow$ reject H_0.

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.

```
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
```

所有變數下 5% 下顯著異於 0.
而 EDU² 在 12% 下顯著異於 0.

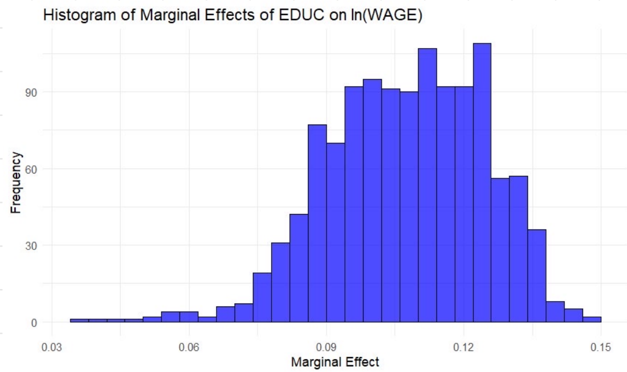
b. $\dfrac{\partial E[\ln(wage)|EDUC, EXPER]}{\partial EDUC}$

$= \beta_2 + 2\beta_3 EDUC + \beta_6 EXPER$

$= 0.08954 + 2 \times 0.002916\, EDUC - 0.001010\, EXPER$

The marginal effect of education increases as the level of EDUC ↑, but EXPER ↓.

c.




```
> quantile(data_df$me_educ, probs = c(0.05, 0.5, 0.95))
        5%        50%        95%
0.08008187 0.10843125 0.13361880
> min(data_df$me_educ)
[1] 0.03565419
> max(data_df$me_educ)
[1] 0.1478705
```

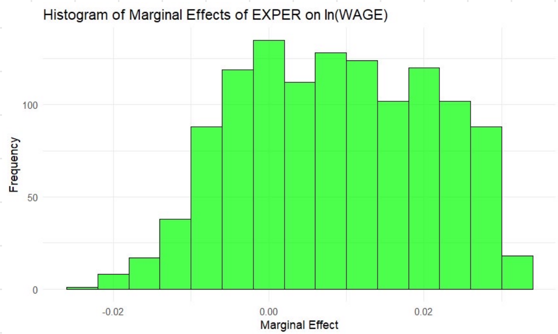
d. $\dfrac{\partial E[\ln(WAGE \mid EDUC, EXPER)]}{\partial EXPER}$

$= \beta_4 + 2\beta_5 + \beta_6 EDUC$

$= 0.044879 + 0.000936\,EXPER + 0.001010\,EDUC$

The marginal effect of EXPER ↑ as EDUC ↑ and as EXPER ↑

e.




```
> quantile(data_df$me_exper, probs = c(0.05, 0.5, 0.95))
          5%          50%          95%
-0.010376212  0.008418878  0.027931151
> min(data_df$me_exper)
[1] -0.02527874
> max(data_df$me_exper)
[1] 0.03398886
```

f. $\begin{cases} H_0: \beta_1 + \beta_2 \times 16 + \beta_3 \times 16^2 + \beta_4 \times 18 + \beta_5 \times 18^2 + \beta_6 \times (16 \times 18) \\ \quad \geq \beta_1 + \beta_2 \times 17 + \beta_3 \times 17^2 + \beta_4 \times 8 + \beta_5 \times 8^2 + \beta_6 \times (17 \times 8) \\ H_1: \beta_1 + \beta_2 \times 16 + \beta_3 \times 16^2 + \beta_4 \times 18 + \beta_5 \times 18^2 + \beta_6 \times (16 \times 18) \\ \quad < \beta_1 + \beta_2 \times 17 + \beta_3 \times 17^2 + \beta_4 \times 8 + \beta_5 \times 8^2 + \beta_6 \times (17 \times 8) \end{cases}$

$\Rightarrow \begin{cases} H_0: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 \geq 0 \\ H_1: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 < 0 \end{cases}$

```
> cat(t_value)
1.669902> qt(0.05,1194)
[1] -1.646131
```

$\Rightarrow RR: \{t : t < -1.6461\}$

⇒ fails to reject H_0

⇒ There is insufficient evidence to conclude that David's log-wage is greater.

g. $\begin{cases} H_0: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 \geq 0 \\ H_1: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 < 0 \end{cases}$

```
> cat(t_value_new)
-2.062365
```

$RR: \{t : t < -1.6461\}$

⇒ reject H_0

⇒ David's log-wage is greater.

∵ EXPER's marginal effect is decreasing.

h. $\begin{cases} H_0: 12\beta_5 - 4\beta_6 = 0 \\ H_1: 12\beta_5 - 4\beta_6 \neq 0 \end{cases}$

```
> cat(t_value_me)
-1.027304
> qt(0.975,1194)
[1] 1.961953
```

$\Rightarrow RR: \{t: |t| > 1.96120\}$

$\Rightarrow$ fails to reject H_0

$\Rightarrow$ There is no evidence to suggest the marginal effect from extra experience are different for Jill and Wendy.

i. $$\frac{\partial E[\ln(Wage) \mid EDUC, EXPER]}{\partial EXPER}$$

$= \beta_4 + 2\beta_5 \times EXPER + \beta_6 \times 16 = 0$

$\Rightarrow 0.04488 + 2 \times -0.000468 \cdot EXPER + -0.001010 \times 16 = 0$

$\Rightarrow EXPER = 19.6771$

```
> cat(lowbg,upbg)
15.95776 23.39636
```