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

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

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. $\beta_1 - \beta_2 + \beta_3 = 4$

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

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
```

b.

```
                 2.5 %    97.5 %
(Intercept) 17.5694018 24.170871
depart       0.2989851  0.437265
reds         1.1574748  1.886411
trains       1.7748867  4.272505
```

Trains 係數範圍較大，不精準。

c.

The t-statistic for testing H0: $\beta 3 \geq 2$ is calculated as $t = \frac{\hat{\beta}_3 - 2}{\mathrm{SE}(\hat{\beta}_3)}$

H1: $\beta 3<2$，The degrees of freedom 249-3-1=245.

```
> # Print the t-statistic and p-value
> cat("t-statistic:", t_stat, "\n")
t-statistic: -2.583562
> cat("p-value (one-sided):", p_value, "\n")
p-value (one-sided): 0.005179509
>
> # Decision at 5% significance level
> alpha <- 0.05
> if (p_value < alpha) {
+   cat("Reject the null hypothesis at the 5% significance level.\n")
+ } else {
+   cat("Fail to reject the null hypothesis at the 5% significance level.\n")
+ }
Reject the null hypothesis at the 5% significance level.
```

拒絕 H0，該係數小於 2

d.

The t-statistic for testing H0: $\beta 4 = 3$

H1: $\beta 4 \neq 3$，The degrees of freedom 249-3-1=245.

```
> # Print the t-statistic and p-value
> cat("t-statistic:", t_stat, "\n")
t-statistic: 0.03737444
> cat("p-value (one-sided):", two_tail_p_value, "\n")
p-value (one-sided): 0.9702169
>
> # Decision at 10% significance level
> alpha <- 0.10
> if (two_tail_p_value < alpha) {
+   cat("Reject the null hypothesis at the 10% significance level.\n")
+ } else {
+   cat("Fail to reject the null hypothesis at the 10% significance level.\n")
+ }
Fail to reject the null hypothesis at the 10% significance level.
```

無法拒絕 H0，$\beta 4 = 3$ 顯著

e.

H0: $\beta 2 \cdot 30 \geq 10$

H1: $\beta 2 \cdot < 1/3$

```
> # Print the t-statistic and p-value
> cat("t-statistic:", t_stat, "\n")
t-statistic: 0.9911646
> cat("p-value (one-sided):", p_value, "\n")
p-value (one-sided): 0.8387085
>
> # Decision at 5% significance level
> alpha <- 0.05
> if (p_value < alpha) {
+   cat("Reject the null hypothesis at the 5% significance level.\n")
+ } else {
+   cat("Fail to reject the null hypothesis at the 5% significance level.\n")
+ }
Fail to reject the null hypothesis at the 5% significance level.
```

無法拒絕 H0，$\beta 2 \cdot > 1/3$ 顯著

f.

H0: $\beta 4-3\beta 3\geq 0$

H1: $\beta 4-3\beta 3<0$

```
> cat("Linear combination (beta_4 - 3 * beta_3):", linear_comb, "\n")
Linear combination (beta_4 - 3 * beta_3): -1.542133
> cat("t-statistic:", t_stat, "\n")
t-statistic: -1.825027
> cat("p-value (left-tailed):", p_value, "\n")
p-value (left-tailed): 0.03460731
>
> # Decision at 5% significance level
> alpha <- 0.05
> if (p_value < alpha) {
+   cat("Reject the null hypothesis at the 5% significance level.\n")
+ } else {
+   cat("Fail to reject the null hypothesis at the 5% significance level.\n")
+ }
Reject the null hypothesis at the 5% significance level.
```

拒絕 H0，$\beta 4-3\beta 3<0$。

g.

$$H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \leq 45$$

$$H_1: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 > 45$$

```
> # 輸出結果
> cat("預期通勤時間 E(TIME|X):", expected_time, "分鐘\n")
預期通勤時間 E(TIME|X): 44.06924 分鐘
> cat("t 統計量:", t_stat, "\n")
t 統計量: -1.725964
> cat("p 值 (右尾):", p_value, "\n")
p 值 (右尾): 0.9571926
>
> # 在 5% 顯著性水平下做出決策
> alpha <- 0.05
> if (p_value < alpha) {
+   cat("在 5% 顯著性水平下拒絕原假設。\n")
+ } else {
+   cat("在 5% 顯著性水平下無法拒絕原假設。\n")
+ }
在 5% 顯著性水平下無法拒絕原假設。
```

無法拒絕 H0，$\beta 1+30\beta 2+6\beta 3+\beta 4\leq 45$

h.

虛無假設是正確的，因其無法拒絕 h0，結果為<45m

若假設為相反

H0: $\beta 1+30\beta 2+6\beta 3+\beta 4>45$

H1: $\beta 1+30\beta 2+6\beta 3+\beta 4<=45$

則拒絕 H0，H1 顯著，結果相同。

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 ** 
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 ***
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
```

在 1%. 5%. 10% 皆顯著異於 0 ： b1. b2. b4. b5. b6

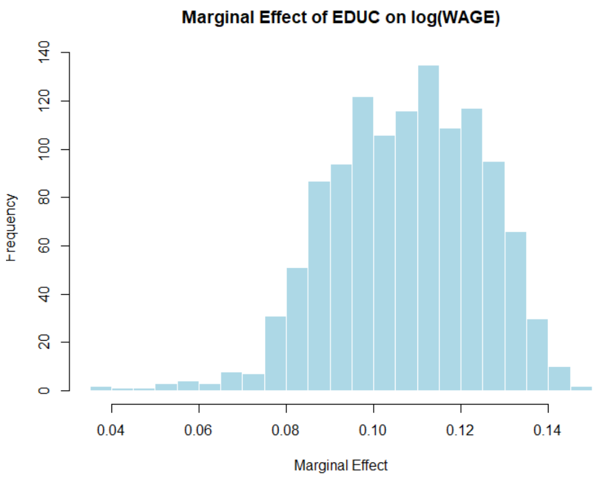
不顯著 ： b3

b.

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

b3 為正，表示隨 educ 增加，邊際效應增加，b6 為負，表示隨 exper 增加，邊際效應減少。

c.



數據為左偏，educ 對於 ln(wage)的邊際效應大部分集中在 0.10-0.12 之間

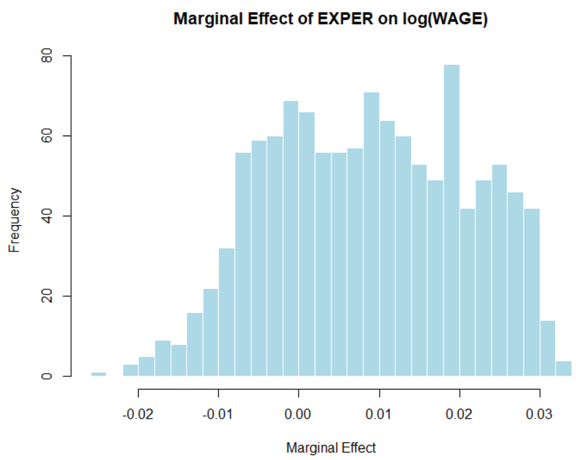
```
        5%        50%        95%
0.08008187 0.10843125 0.13361880
```

d.

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

b5 為負，表示隨 exper 增加，邊際效應減少，b6 為負，表示隨 educ 增加，邊際效應減少。

e.



數據為左偏，exper 對於 ln(wage)的邊際效應大部分集中在 0-0.03 之間

```
          5%          50%         95%
-0.010376212  0.008418878  0.027931151
```

f.

```
> cat("標準誤:", se_delta, "\n")
標準誤: 0.104555
> cat("t 統計量:", t_stat, "\n")
t 統計量: 0.3432123
> cat("p 值:", p_value, "\n")
p 值: 0.3657496
>
> # 檢驗結論
> if (p_value < 0.05) {
+   cat("拒絕虛無假設，David 的預期對數工資大於 Svetlana。\n")
+ } else {
+   cat("不能拒絕虛無假設，沒有足夠證據表明 David 的預期對數工資大於 Svetlana。\n")
+ }
不能拒絕虛無假設，沒有足夠證據表明 David 的預期對數工資大於 Svetlana。
```

g.

```
> cat("t 統計量:", t_stat, "\n")
t 統計量: -0.2957024
> cat("p 值:", p_value, "\n")
p 值: 0.6162456
>
> # 檢驗結論
> if (p_value < 0.05) {
+   cat("拒絕虛無假設，David 的預期對數工資大於 Svetlana。\n")
+ } else {
+   cat("不能拒絕虛無假設，沒有足夠證據表明 David 的預期對數工資大於 Svetlana。\n")
+ }
不能拒絕虛無假設，沒有足夠證據表明 David 的預期對數工資大於 Svetlana。
```

h.

虛無假設 $H_0 : 12\beta_5 - 4\beta_6 = 0$

對立假設 $H_1 : 12\beta_5 - 4\beta_6 \neq 0$

```
> cat("t 統計量:", t_stat, "\n")
t 統計量: -1.027246
> cat("p 值:", p_value, "\n")
p 值: 0.3045129
>
> # 檢驗結論
> if (p_value < 0.05) {
+   cat("拒絕虛無假設，Wendy 和 Jill 的額外經驗邊際效應不相等。\n")
+ } else {
+   cat("不能拒絕虛無假設，沒有足夠證據表明 Wendy 和 Jill 的額外經驗邊際效應不相等。\n")
+ }
不能拒絕虛無假設，沒有足夠證據表明 Wendy 和 Jill 的額外經驗邊際效應不相等。
```

無法拒絕 H0，Wendy 和 Jill 的額外經驗邊際效應相等。

i.

```
> cat("邊際效應為 0 時的 EXPER:", exper_zero, "年\n")
邊際效應為 0 時的 EXPER: 30.67706 年
> cat("EXPER 的標準誤:", se_exper, "\n")
EXPER 的標準誤: 6.329026
> cat("95% 置信區間: [", ci_lower, ", ", ci_upper, "]\n")
95% 置信區間: [ 18.25981 ,  43.09431 ]
```