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

a. $H_0: \beta_2 = 0$ $\quad t = \frac{3}{\sqrt{4}} = 1.5 < t_c = 2.000298$

$H_1: \beta_2 \neq 0$ $\quad \therefore$ We failed to reject H_0, means $\beta_2 = 0$

b. $H_0: \beta_1 + 2\beta_2 = 5$ $\quad t = \frac{b_1 + 2b_2 - 5}{se(b_1 + 2b_2)} \sim t_{60} = 0.904534 < 2.000298$

$H_1: \beta_1 + 2\beta_2 \neq 5$ $\quad \therefore$ We failed to reject H_0, means $\beta_1 + 2\beta_2 = 5$

c. $H_0: \beta_1 - \beta_2 + \beta_3 = 4$ $\quad t = \frac{b_1 - b_2 + b_3 - 4}{se(b_1 - b_2 + b_3)} \sim t_{60} = -1.5 > -2.000298$

$H_1: \beta_1 - \beta_2 + \beta_3 \neq 4$ $\quad \therefore$ We failed to reject H_0, means $\beta_1 - \beta_2 + \beta_3 = 4$

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|)
β1 (Intercept)  20.8701     1.6758  12.454  < 2e-16 ***
β2 depart        0.3681     0.0351  10.487  < 2e-16 ***
β3 reds          1.5219     0.1850   8.225 1.15e-14 ***
β4 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
```

$\widehat{TIME} = 20.8701 + 0.3681\,DEPART + 1.5219\,REDS + 3.0237\,TRAINS$

β_1 Bill's expected commute time when he leaves Carnegie at 6:30AM and encounters no red lights and no trains is estimated to be 20.87 minutes.

β_2 If Bill leaves later than 6:30AM, the increase in his expected traveling time is estimated to be 3.7 minutes for every 10 minutes that his departure time is later than 6:30AM (assuming the number of red lights and trains are constant).

β_3 The expected increase in traveling time from each red light, with departure time and number of trains held constant, is estimated to be 1.52 minutes.

β_4 The expected increase in traveling time from each train, with departure time and number of red lights held constant, is estimated to be 3.02 minutes.

b.

```
                   2.5 %    97.5 %
β1 (Intercept) 17.5694018 24.170871
β2 depart       0.2989851  0.437265
β3 reds         1.1574748  1.886411
β4 trains       1.7748867  4.272505
```

c. $H_0: \beta_3 \geq 2$ $\quad t = \frac{1.5219 - 2}{0.185} = -2.5843 < -1.6511 = t_c$

$H_1: \beta_3 < 2$ $\quad \therefore$ We reject H_0, means that the expected delay from each red light is less than 2 mins.

d. $H_0: \beta_4 = 3$ $\quad t = \frac{3.0237 - 3}{0.634} = 0.0374$ $\quad t_c = \pm 1.6511$

$H_1: \beta_4 \neq 3$ $\quad -t_c < t < t_c$ $\quad \therefore$ We failed to reject H_0, means that the expected delay from each train is 3 mins.

e. $H_0: \beta_2 \geq \frac{1}{3}$ $t = \frac{0.3681 - \frac{1}{3}}{0.0351} = 0.9905 > -1.6511 = t_c$

$H_1: \beta_2 < \frac{1}{3}$ ∴ We failed to reject H_0, means that delaying departure time by 30 mins. increases expected travel time by at least 10 mins.

f. $H_0: \beta_4 \geq 3\beta_3 \Rightarrow \beta_4 - 3\beta_3 \geq 0$ $t = -1.825 < -1.6511$

$H_1: \beta_4 < 3\beta_3 \Rightarrow \beta_4 - 3\beta_3 < 0$ ∴ We reject H_0, means that the expected delay from a train is less than 3 times the delay from a red light.

g. $H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \leq 45 \Rightarrow \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 - 45 \leq 0$ $t = -1.726 < 1.6511 = t_c$

$H_1: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 > 45 \Rightarrow \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 - 45 > 0$ ∴ We failed to reject H_0, means that we can't conclude that Bill we get to the University after 7:45 AM.

h. $H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 \geq 45$ $t = -1.726 < -1.6511$

$H_0: \beta_1 + 30\beta_2 + 6\beta_3 + \beta_4 < 45$ ∴ We reject H_0, means Bill's expected commute time which can make him be on time for the meeting.

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.

```
              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
```

All coefficient estimates are significantly different from zero at 1 % level.

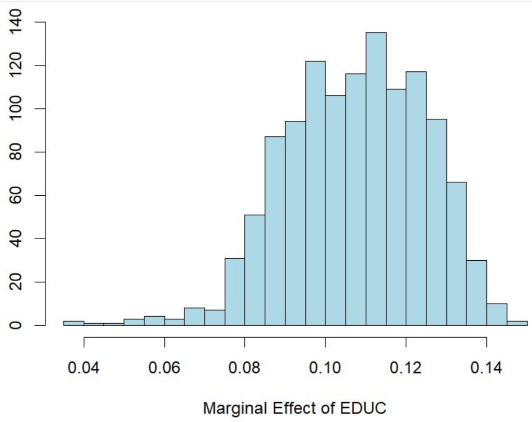
Except $EDUC^2$ is significant at 12 % level.

b.

$ME_{EDUC} = \beta_2 + 2\beta_3 EDUC + \beta_6 EXPER$ $\widehat{ME}_{EDUC} = 0.08954 + 0.002916\,EDUC - 0.00101\,EXPER$

The marginal effect of education increases as level of education increases, but decreases with the level of experience.

c.

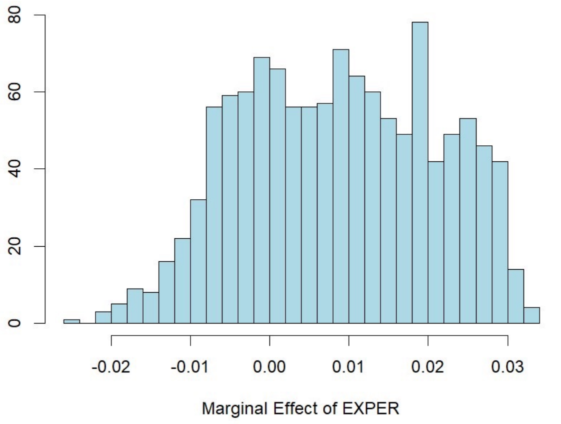


5%	50%	95%
0.08008187	0.10843125	0.13361880

d. $ME_{EXPER} = \beta_4 + 2\beta_5 EXPER + \beta_6 EDUC$ $\widehat{ME}_{EXPER} = 0.04488 - 0.000936 EXPER - 0.00101 EDUC$

The marginal effect of experience decreases as level of education increases, and as the years of experience increase

e.



5%	50%	95%
0.08008187	0.10843125	0.13361880

f. $H_0: -\beta_2 - 33\beta_3 + 10\beta_4 + 260\beta_5 + 152\beta_6 \ge 0$ $t = -1.6699 > t_c = -1.6463$

$H_1: -\beta_2 - 33\beta_3 + 10\beta_4 + 260\beta_5 + 152\beta_6 < 0$ ∴ We fail to reject H_0, we can't conclude that David's log-wage is greater.

g. $H_0: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 \ge 0$ $t = -2.0624 < t_c = -1.6463$

$H_1: -\beta_2 - 33\beta_3 + 10\beta_4 + 420\beta_5 + 144\beta_6 < 0$ ∴ We reject H_0, we can conclude that David's log-wage is greater.

h. $H_0: 12\beta_5 - 4\beta_6 = 0$ $t = -1.027$ $t_c = \pm 1.962$

$H_1: 12\beta_5 - 4\beta_6 \neq 0$ $\because -1.962 < 1.027 < 1.962$ ∴ We fail to reject H_0, there's no evidence to say that the marginal effects from extra experience are different for Jill & Wendy.

i. The estimate would be 19.677 & 95% interval = [15.9578 , 23.3964]