

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

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

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

$$\hat{y} = \beta_1 + \beta_2 X_1 + \beta_3 X_2$$

$$a. H_0: \beta_2 = 0 / H_1: \beta_2 \neq 0$$

$$t(63-3, 0.975) = 2.003$$

$$t^* = \frac{3-0}{\sqrt{4}} = 1.5$$

$$1.5 < 2.003$$

Not reject H_0

$$b. H_0: \beta_1 + 2\beta_2 = 5 / H_1: \beta_1 + 2\beta_2 \neq 5$$

$$\text{Var}(b_1 + 2b_2) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 11$$

$$t(63-3, 0.975) = 2.003$$

$$t^* = \frac{2+2(3)-5}{\sqrt{11}} = 0.9043$$

$$0.9043 < 2.003$$

Not reject H_0

$$c. H_0: \beta_1 - \beta_2 + \beta_3 = 4 / H_1: \beta_1 - \beta_2 + \beta_3 \neq 4$$

$$\text{Var}(b_1 - b_2 + b_3)$$

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

$$= 16$$

$$t \sim t(63-3, 0.975) = 2.003$$

$$t^* = \frac{2-3-1-4}{4} = -1.5$$

$$-1.5 < 2.003$$

Not 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 .

- Find 95% interval estimates for each of the coefficients. Have you obtained precise estimates of each of the coefficients?
- 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.
- 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.
- 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.)
- 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.
- 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 $E(TIME|X)$ where X represents the observations on all explanatory variables.]
- 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. $\beta_1 = 20.8701$ 在 6:30 出发, 没红橙黄火车, 折返时间有 20.8701
 $\beta_2 = 0.3681$ 其他不要下, 晚 1 分钟出发, 折返时间增 0.3681 分钟
 $\beta_3 = 1.5219$ 其他不要下, 多 1 个红橙, 折返时间增 1.5219 分钟
 $\beta_4 = 3.0237$ 其他不要下, 多 1 个火车, 折返时间增 3.0237 分钟

```
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

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

t = -2.583562

$$c. H_0: \beta_3 \geq 2 / H_1: \beta_3 < 2$$

$$t = -2.56 < -1.65 \quad \text{reject } H_0$$

t = 0.03737444

$$d. H_0: \beta_4 = 3 / H_1: \beta_4 \neq 3$$

$$t = 0.037 < 1.65 \quad \text{no reject } H_0$$

t = 0.9911646

$$e. H_0: 60\beta_2 - 30\beta_3 \geq 10$$

$$H_1: 60\beta_2 - 30\beta_3 < 10$$

$$t = 0.9911 > -1.66 \quad \text{not reject } H_0$$

t = -1.830017

$$f. H_0: \beta_4 \geq 3\beta_3$$

$$H_1: \beta_4 < 3\beta_3$$

$$t = -1.825 < -1.65 \quad \text{reject } H_0$$

t = -1.725964

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

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

$$t = -1.7259 > 1.651 \quad \text{no reject } H_0$$

t = -1.725964

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

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

$$t = -1.726 < -1.651 \quad \text{reject } H_0$$

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

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

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

a. 除了 $educ^2$, 所有係數都在 $\alpha=1$ 下顯著異於 0

$$b. \frac{\partial E(\ln(WAGE))}{\partial EDUC} = 0.0824 + 0.001458 \cdot EDUC - 0.00101 \cdot EXPER$$

$$d. \frac{\partial E(\ln(WAGE))}{\partial EXPER} = 0.004488 - 0.000468 \cdot EXPER - 0.00101 \cdot EDUC$$

Appendix 5A Derivation of Least Squares Estimators

- 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?
- 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.
- 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), 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

$$H_0: \beta_2 + 33\beta_3 - 10\beta_4 - 260\beta_5 - 152\beta_6 \leq 0$$

$$H_1: \beta_2 + 33\beta_3 - 10\beta_4 - 260\beta_5 - 152\beta_6 > 0$$

```
{差異(David - Svetlana): -0.03588456
> cat("t 值:", round(t_val, 4), "; p
t 值: -1.6699 ; p 值: 0.9524
```

$$H_0: \beta_2 + 33\beta_3 - 10\beta_4 - 420\beta_5 - 144\beta_6 \leq 0$$

$$H_1: \beta_2 + 33\beta_3 - 10\beta_4 - 420\beta_5 - 144\beta_6 > 0$$

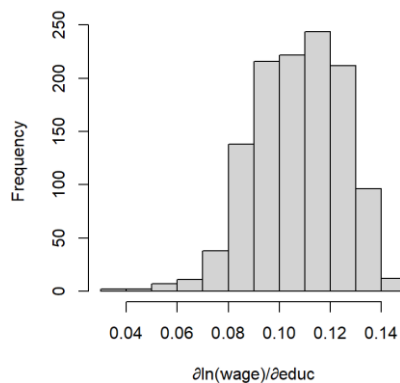
```
差異(David - Svetlana): 0.03091716
> cat("t 值:", round(t_val, 4), "; p
t 值: 2.062365 ; p 值: 0.01958648
```

$$h. H_0: 12\beta_5 - 4\beta_6 = 0$$

$$H_1: 12\beta_5 - 4\beta_6 \neq 0$$

```
差異(David - Svetlana): -0.001575327
> cat("t 值:", round(t_val, 4), "; p
t 值: -1.027304 ; p 值: 0.8478613
```

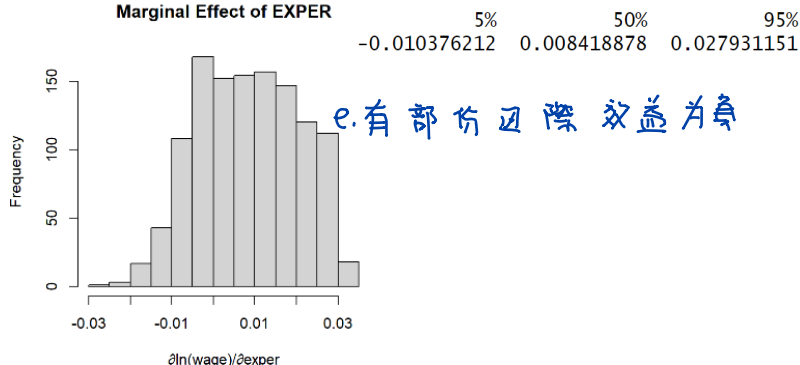
Marginal Effect of EDUC



第 5 百分位數: 0.08008187
> cat("中位數:", quantile
中位數: 0.1084313
> cat("第 95 百分位數:",
第 95 百分位數: 0.1336188

c. EDUC ↑, WAGE ↑
呈鐘型

Marginal Effect of EXPER



e. 有部分邊際效益為負

∴

再過多久經驗邊際效益變為負值: 19.67706
> cat("95% 置信區間: [", ci_lower, ", ", ci_upper, "]\n")
95% 置信區間: [15.96146 , 23.39265]