HW0331

Yung-Jung Cheng

2025-04-06

Q06

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

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

$$\widehat{\text{Cov}}(\mathbf{b}) = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

Using a 5% significance level, and an alternative hypothesis that the equality does not hold, test each of the following null hypotheses:

Q06(a)

$H_0 : \beta_2 = 0$

Ans

Let $c = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $d = 0$.

The test statistic is:

$$t = \frac{c^\top \mathbf{b} - d}{\sqrt{c^\top \widehat{\text{Cov}}(\mathbf{b})c}} = \frac{3 - 0}{\sqrt{4}} = \frac{3}{2} = 1.5$$

Since $|t| = 1.5 < t_{0.025,60} \approx 2.000$, we **fail to reject** H_0.

Conclusion: At the 5% significance level, there is insufficient evidence to conclude that $\beta_2 \neq 0$.

Q06(b)

$H_0 : \beta_1 + 2\beta_2 = 5$

Ans

We define the linear combination:

- $c = [1,\ 2,\ 0]^T$, so that $c^\top \beta = \beta_1 + 2\beta_2$
- Estimate: $c^\top \mathbf{b} = 1 \cdot 2 + 2 \cdot 3 = 8$
- Null value: $c_0 = 5$

Compute the variance of the linear combination:

Using the covariance matrix:

$$\widehat{\mathrm{Cov}}(b) = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

We compute:

$$\mathrm{Var}(c^\top b) = 1^2 \cdot 3 + 2^2 \cdot 4 + 2 \cdot 1 \cdot 2 \cdot (-2)$$
$$= 3 + 16 - 8 = 11$$

Test Statistic:

$$t = \frac{8-5}{\sqrt{11}} \approx \frac{3}{3.3166} \approx 0.9045$$

Critical Value:

From the t-distribution with 60 degrees of freedom:

$$t_{0.975,60} \approx 2.0003$$

Conclusion:

Since $|t| = 0.9045 < 2.0003$, we **fail to reject** the null hypothesis at the 5% significance level. There is not enough evidence to conclude that $\beta_1 + 2\beta_2$ is different from 5.

Q06(c)

$H_0 : \beta_1 - \beta_2 + \beta_3 = 4$

Ans

We define the linear combination:

- $c = [1,\ -1,\ 1]^T$, so that $c^\top \beta = \beta_1 - \beta_2 + \beta_3$
- Estimate: $c^\top \mathbf{b} = 1 \cdot 2 - 1 \cdot 3 + 1 \cdot (-1) = -2$
- Null value: $c_0 = 4$

Compute the variance of the linear combination:

Given the covariance matrix:

$$\widehat{\mathrm{Cov}}(b) = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

We compute:

$$\mathrm{Var}(c^\top b) = 1^2 \cdot 3 + (-1)^2 \cdot 4 + 1^2 \cdot 3$$
$$+ 2 \cdot (1)(-1) \cdot (-2) + 2 \cdot (1)(1) \cdot 1 + 2 \cdot (-1)(1) \cdot 0$$
$$= 3 + 4 + 3 + 4 + 2 + 0 = 16$$

Test Statistic:

$$t = \frac{-2-4}{\sqrt{16}} = \frac{-6}{4} = -1.5$$

Critical Value:

From the t-distribution with 60 degrees of freedom:

$$t_{0.975,60} \approx 2.0003$$

Conclusion:

Since $|t| = 1.5 < 2.0003$, we **fail to reject** the null hypothesis at the 5% significance level.
There is insufficient evidence to conclude that $\beta_1 - \beta_2 + \beta_3$ is different from 4.

Q31

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) – measured in minutes after 6:30 AM,
- the number of red lights (REDS) he encounters, and
- the number of trains (TRAINS) that he has to wait for at the Murrumbeena level crossing.

Observations on these variables for the 249 working days in 2015 appear in the file commute5 . TIME is measured in minutes.

Q31(a)

Estimate the equation:

$$\text{TIME} = \beta_1 + \beta_2 \cdot \text{DEPART} + \beta_3 \cdot \text{REDS} + \beta_4 \cdot \text{TRAINS} + \varepsilon$$

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

Ans

```
# Estimate regression model for commute time
lm_commute_time <- lm(time ~ depart + reds + trains, data = commute5)
summary(lm_commute_time)
```

```
##
## 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
```

The estimated equation is:

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

Interpretation of coefficients:

- **Intercept (20.8701):** Estimated commute time (in minutes) when DEPART, REDS, and TRAINS are all zero — that is, Bill leaves at 6:30 AM and encounters no red lights or trains.

- **DEPART (0.3681):** Each additional minute after 6:30 AM that Bill departs increases expected commute time by 0.3681 minutes, holding other factors constant.

- **REDS (1.5219):** Each additional red light adds 1.5219 minutes to the commute, holding departure time and number of trains constant.

- **TRAINS (3.0237):** Each additional train delay increases commute time by 3.0237 minutes, all else equal.

Q31(b)

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

Ans

```
# Compute 95% confidence intervals for all coefficients
confint(lm_commute_time, 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
```

The 95% confidence intervals for the estimated coefficients are:

- **Intercept**: [17.5694, 24.1709]
- **DEPART**: [0.2990, 0.4373]
- **REDS**: [1.1575, 1.8864]
- **TRAINS**: [1.7749, 4.2725]

Precision of estimates:

- The interval for `DEPART` is narrow, suggesting that its effect is **precisely estimated**.
- The interval for `REDS` is also reasonably tight, indicating **good precision**.
- The interval for `TRAINS` is relatively **wider**, suggesting that the estimate is **less precise**, likely due to more variability or fewer observations involving trains.

In summary, the coefficients for `DEPART` and `REDS` are estimated with higher precision than the coefficient for `TRAINS`.

Q31(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.

Ans

```
# Extract estimate and standard error for reds
b_reds <- coef(lm_commute_time)["reds"]
se_reds <- sqrt(vcov(lm_commute_time)["reds", "reds"])
df <- lm_commute_time$df.residual

# Compute t-statistic
t_stat_reds <- (b_reds - 2) / se_reds

# Compute one-tailed p-value (lower tail)
p_value_reds <- pt(t_stat_reds, df)

# Output result
c(t_value = t_stat_reds, p_value = p_value_reds)
```

```
## t_value.reds p_value.reds
## -2.583561972  0.005179509
```

We test the null hypothesis:

- $H_0: \beta_{\text{REDS}} \geq 2$
- $H_1: \beta_{\text{REDS}} < 2$

Using the estimate $\hat{\beta}_{\text{REDS}} = 1.5219$ and standard error 0.1850, the test statistic is:

$$t = -2.5836, \quad p\text{-value} = 0.0052$$

Conclusion: Since the p-value is less than 0.05, we reject the null hypothesis.
There is significant evidence that the expected delay from each red light is **less than 2 minutes**.

Q31(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.

Ans

```
# Extract estimate and standard error for trains
b_trains <- coef(lm_commute_time)["trains"]
se_trains <- sqrt(vcov(lm_commute_time)["trains", "trains"])
df <- lm_commute_time$df.residual

# Compute t-statistic for H0: beta_trains = 3
t_stat_trains <- (b_trains - 3) / se_trains

# Two-tailed p-value
p_value_trains <- 2 * (1 - pt(abs(t_stat_trains), df))

# Output results
c(t_value = t_stat_trains, p_value = p_value_trains)
```

```
## t_value.trains p_value.trains
##     0.03737444     0.97021689
```

We test the null hypothesis:

- $H_0: \hat{\beta}_{\text{TRAINS}} = 3$
- $H_1: \beta_{\text{TRAINS}} \neq 3$

With $\hat{\beta}_{\text{TRAINS}} = 3.0237$, standard error = 0.6340, and 245 degrees of freedom:

$$t = \frac{3.0237 - 3}{0.6340} = 0.0374$$

The corresponding two-sided p-value is 0.9702.

Conclusion: Since the p-value is much greater than 0.10, we **fail to reject** the null hypothesis.
There is no significant evidence that the expected delay from each train differs from 3 minutes.

Q31(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.)

Ans

```
# Estimate: 30 * beta_depart
b_depart <- coef(lm_commute_time)["depart"]
se_depart <- sqrt(vcov(lm_commute_time)["depart", "depart"])
df <- lm_commute_time$df.residual

# Compute estimated difference in expected time
L_diff <- 30 * b_depart

# Compute standard error of the linear combination
se_L_diff <- sqrt((30^2) * se_depart^2)

# Compute test statistic
t_stat_diff <- (L_diff - 10) / se_L_diff

# Compute one-sided p-value
p_value_diff <- pt(t_stat_diff, df)

# Output result
c(t_value = t_stat_diff, p_value = p_value_diff)
```

```
## t_value.depart p_value.depart
##      0.9911646      0.8387085
```

We test the null hypothesis:

- $H_0: \lambda \geq 10$, where $\lambda = E[\text{TIME}|\text{DEPART} = 60] - E[\text{TIME}|\text{DEPART} = 30]$
- $H_1: \lambda < 10$

Since $\lambda = 30 \cdot \beta_{\text{DEPART}}$, the estimate is:

$$\hat{\lambda} = 30 \cdot 0.3681 = 11.043$$

Standard error:

$$\text{SE}(\hat{\lambda}) = 30 \cdot 0.0351 = 1.053$$

t-statistic:

$$t = \frac{11.043 - 10}{1.053} = 0.9912$$

With a one-sided p-value of 0.8387, we **fail to reject** the null hypothesis at the 5% level.

Conclusion: There is no evidence that the expected commute time increase is less than 10 minutes.
In fact, the data is consistent with the increase being at least 10 minutes or more.

Q31(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.

Ans

```
# Extract estimates and standard errors
b_reds <- coef(lm_commute_time)["reds"]
b_trains <- coef(lm_commute_time)["trains"]

var_reds <- vcov(lm_commute_time)["reds", "reds"]
var_trains <- vcov(lm_commute_time)["trains", "trains"]
cov_reds_trains <- vcov(lm_commute_time)["reds", "trains"]
df <- lm_commute_time$df.residual

# Define linear combination: lambda = beta_trains - 3 * beta_reds
lambda_hat <- b_trains - 3 * b_reds

# Variance of linear combination
var_lambda <- var_trains + 9 * var_reds - 6 * cov_reds_trains
se_lambda <- sqrt(var_lambda)

# t statistic and p-value (left-tail)
t_stat_f <- (lambda_hat - 0) / se_lambda
p_value_f <- pt(t_stat_f, df)

# Output results
c(t_value = t_stat_f, p_value = p_value_f)
```

```
## t_value.trains p_value.trains
##    -1.82502667      0.03460731
```

We define the linear combination:

$$\lambda = \beta_{\text{TRAINS}} - 3 \cdot \beta_{\text{REDS}}$$

We test:

- $H_0: \lambda \geq 0$
- $H_1: \lambda < 0$

From the regression:

- $\hat{\beta}_{\text{TRAINS}} = 3.0237$
- $\hat{\beta}_{\text{REDS}} = 1.5219$
- So: $\hat{\lambda} = 3.0237 - 3 \cdot 1.5219 = -1.5419$

The standard error of $\hat{\lambda}$ is approximately 0.8447, resulting in:

$$t = \frac{-1.5419}{0.8447} = -1.8250$$

With 245 degrees of freedom, the p-value is 0.0346.

Conclusion: Since the p-value is less than 0.05, we **reject the null hypothesis**.
There is significant evidence that the expected delay from a train is **less than three times** the delay from a red light.

Q31(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.) ### Ans

```
# Define x0 = (Intercept, depart, reds, trains)
x0 <- c(1, 30, 6, 1)
b_hat <- coef(lm_commute_time)
df <- lm_commute_time$df.residual

# Predicted value
pred_time <- sum(x0 * b_hat)

# Standard error of the predicted mean
vcov_mat <- vcov(lm_commute_time)
var_pred <- t(x0) %*% vcov_mat %*% x0
se_pred <- sqrt(var_pred)

# t-statistic and one-sided p-value (right-tail)
t_stat_g <- (pred_time - 45) / se_pred
p_value_g <- 1 - pt(t_stat_g, df)

# Output
c(predicted_time = pred_time, t_value = t_stat_g, p_value = p_value_g)
```

```
## predicted_time        t_value        p_value
##      44.0692419     -1.7259635      0.9571926
```

We test whether Bill's expected commute time exceeds 45 minutes when:

- DEPART = 30 (7:00 AM)
- REDS = 6
- TRAINS = 1

We test:

- H_0: $E(\text{TIME}|\mathbf{X}) \leq 45$
- H_1: $E(\text{TIME}|\mathbf{X}) > 45$

The predicted commute time is:

$$E(\widehat{\text{TIME}|\mathbf{X}}) = 44.0692 \text{ minutes}$$

The standard error of the prediction is approximately 0.539, giving:

$$t = \frac{44.0692 - 45}{SE} = -1.726, \quad p\text{-value} = 0.9572$$

Conclusion: Since the p-value is much greater than 0.05, we **fail to reject** the null hypothesis.
There is no evidence that Bill will arrive **after 7:45 AM**. The data supports that he can reasonably expect to arrive **on or before 7:45 AM**.

Q31(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?

Ans

Yes, the hypotheses in part (g) have been set up correctly **from a conservative (risk-averse) perspective**:

- H_0: Bill arrives **on or before** 7:45 AM
- H_1: Bill arrives **after** 7:45 AM

This setup puts the **burden of proof** on the evidence that he will be late.
In this context, **failing to reject** H_0 means we don't have strong evidence he will be late — so we tentatively trust that he will be on time.

If the hypotheses were reversed:

- H_0: Bill arrives **after** 7:45 AM
- H_1: Bill arrives **on or before** 7:45 AM

Then **failing to reject H_0** means we don't have evidence he'll be on time, which might lead us to assume he's at risk of being late — **a more cautious stance**.

Implication:
The choice of hypotheses reflects what kind of **error** we want to avoid.
In situations where **being late is unacceptable**, some may prefer setting:

- H_0: Bill is **late**, and only reject it with strong evidence he'll be on time.

This approach **prioritizes safety** (i.e., avoiding a Type II error where we mistakenly assume he'll be on time).

Thus, both setups can be valid, depending on the risk tolerance and context. The choice should reflect the **real-world consequences** of making each type of error.

===

Q33

This exercise investigates a nonlinear relationship between log-wages and education/experience. We use the following model:

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

The data set `cps5_small` contains a random sample of 1000 observations from the full `cps5` dataset used in the text.

Q33(a)

Which of the estimated coefficients in the regression are significantly different from 0?

Ans

```
# 建立 EDUC^2、EXPER^2、EDUC × EXPER 的交乘項
cps5_small <- cps5_small %>%
  mutate(
    educ_sq = educ^2,
    exper_sq = exper^2,
    educ_exper = educ * exper,
    log_wage = log(wage)
  )

# 配適模型
model_logwage <- lm(log_wage ~ educ + educ_sq + exper + exper_sq + educ_exper, data = cps5_small)

# 顯示結果摘要
summary(model_logwage)
```

```
##
## Call:
## lm(formula = log_wage ~ educ + educ_sq + exper + exper_sq + 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 **
## educ_sq      1.458e-03  9.242e-04   1.578 0.114855
## exper        4.488e-02  7.297e-03   6.150 1.06e-09 ***
## exper_sq    -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
```

We estimate the following nonlinear regression model:

$$\ln(\text{WAGE}) = \beta_1 + \beta_2 \cdot \text{EDUC} + \beta_3 \cdot \text{EDUC}^2 + \beta_4 \cdot \text{EXPER} + \beta_5 \cdot \text{EXPER}^2 + \beta_6 \cdot (\text{EDUC} \cdot \text{EXPER}) + e$$

Based on the regression output:

- The **intercept** is statistically significant at the 1% level (p = 0.0002).
- The **EDUC** coefficient is significant at the 1% level (p = 0.0040).
- The **EDUC²** coefficient is **not** statistically significant (p = 0.1149).
- The **EXPER** coefficient is highly significant (p < 0.0001).
- The **EXPER²** coefficient is highly significant (p < 0.0001).
- The **EDUC × EXPER** interaction term is significant at the 1% level (p = 0.0078).

Conclusion:

All coefficients except for `EDUC²` are significantly different from zero at the 5% significance level. This suggests that education, experience, and their interaction all have statistically significant nonlinear effects on log-wages, though the quadratic effect of education alone is not strongly supported by the data.

Q33(b)

Let $g(\text{EDUC}, \text{EXPER}) = E[\ln(\text{WAGE}) \mid \text{EDUC}, \text{EXPER}]$.
Find the partial derivative of g with respect to EDUC. Give a brief economic interpretation of this partial derivative.

Ans

Given the model:

$$g(\text{EDUC}, \text{EXPER}) = \beta_1 + \beta_2 \cdot \text{EDUC} + \beta_3 \cdot \text{EDUC}^2 + \beta_4 \cdot \text{EXPER} + \beta_5 \cdot \text{EXPER}^2 + \beta_6 \cdot (\text{EDUC} \cdot \text{EXPER})$$

We take the partial derivative with respect to EDUC:

$$\frac{\partial g}{\partial \text{EDUC}} = \beta_2 + 2\beta_3 \cdot \text{EDUC} + \beta_6 \cdot \text{EXPER}$$

Economic interpretation:

This expression represents the **marginal effect of an additional year of education** on the expected log-wage, **holding experience constant**.

- β_2 is the **direct linear effect** of EDUC.
- $2\beta_3 \cdot \text{EDUC}$ captures the **nonlinear (quadratic) curvature** in the return to education.
- $\beta_6 \cdot \text{EXPER}$ reflects that **returns to education may vary depending on experience** — for instance, more experienced individuals may benefit more (or less) from an additional year of schooling.

This shows that the marginal return to education is **not constant**; it depends on both the level of education and experience.

Q33(c)

Using the data, compute the marginal effect of EDUC on $\ln(\text{WAGE})$ for each individual in the sample.
Construct a histogram of the marginal effects.
What is the range of values for the marginal effect of EDUC?

Ans

```r
# 提取估計係數
b <- coef(model_logwage)
beta2 <- b["educ"]
beta3 <- b["educ_sq"]
beta6 <- b["educ_exper"]

# 計算每位樣本的邊際效果
cps5_small <- cps5_small %>%
  mutate(marginal_educ = beta2 + 2 * beta3 * educ + beta6 * exper)

# 畫出直方圖
hist(cps5_small$marginal_educ,
     main = "Histogram of Marginal Effects of EDUC on log(WAGE)",
     xlab = "Marginal Effect",
     col = "lightblue",
     border = "gray")
```


Histogram of Marginal Effects of EDUC on log(WAGE)
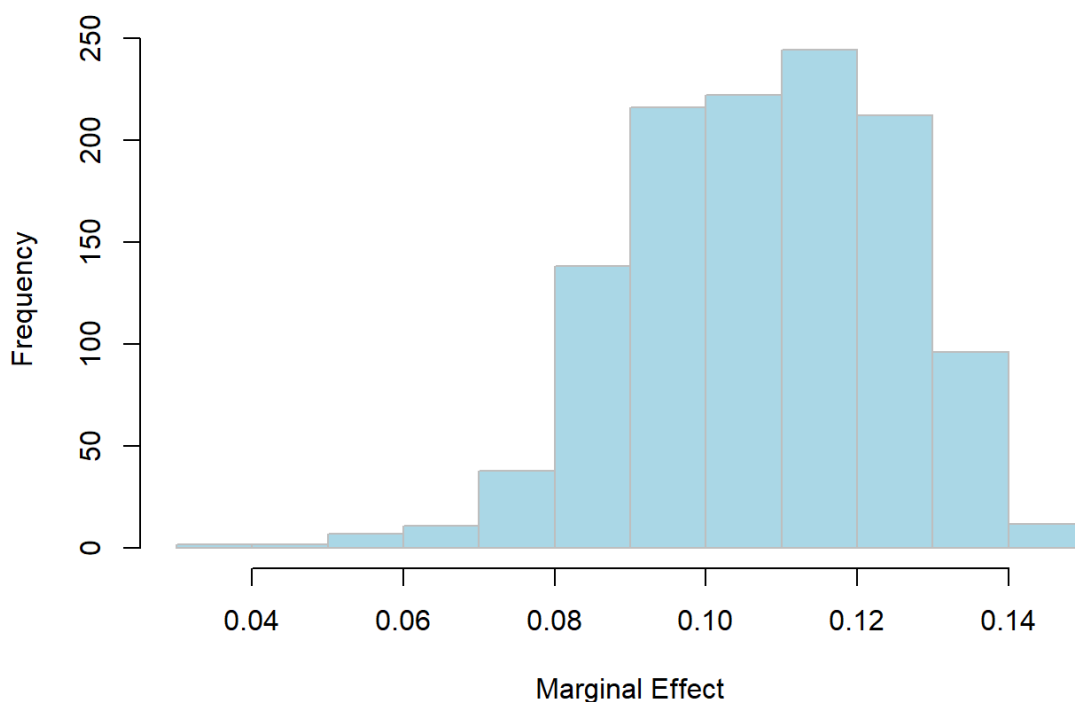
```r
# 顯示範圍
range(cps5_small$marginal_educ)
```

```
## [1] 0.03565419 0.14787053
```

Using the partial derivative from (b), we compute the marginal effect of EDUC on $\ln(\text{WAGE})$ for each individual:

$$\frac{\partial \ln(\text{WAGE})}{\partial \text{EDUC}} = \hat{\beta}_2 + 2\hat{\beta}_3 \cdot \text{EDUC} + \hat{\beta}_6 \cdot \text{EXPER}$$

A histogram of these marginal effects shows that the values are concentrated around 0.10, with a slightly right-skewed shape:

Histogram

The **range of marginal effects** is:

- **Minimum**: 0.0357
- **Maximum**: 0.1479

Interpretation:

The marginal effect of an additional year of education on log-wage varies across individuals, depending on their level of education and experience.
In this dataset, an extra year of schooling is associated with a **3.6% to 14.8% increase** in expected log-wages, assuming other variables are held constant.

Q33(d)

Now find the partial derivative of $g(\text{EDUC}, \text{EXPER})$ with respect to EXPER.
Give a brief interpretation of this quantity.

Ans

Recall the model:

$$g(\text{EDUC}, \text{EXPER}) = \beta_1 + \beta_2 \cdot \text{EDUC} + \beta_3 \cdot \text{EDUC}^2 + \beta_4 \cdot \text{EXPER} + \beta_5 \cdot \text{EXPER}^2 + \beta_6 \cdot (\text{EDUC} \cdot \text{EXPER})$$

Taking the partial derivative with respect to EXPER:

$$\frac{\partial g}{\partial \text{EXPER}} = \beta_4 + 2\beta_5 \cdot \text{EXPER} + \beta_6 \cdot \text{EDUC}$$

Economic interpretation:

This partial derivative measures the **marginal effect of an additional year of labor market experience** on the expected log of wages, **holding education constant**.

- β_4 is the direct linear effect of experience.
- $2\beta_5 \cdot \text{EXPER}$ represents diminishing or increasing marginal returns to experience.
- $\beta_6 \cdot \text{EDUC}$ captures how the return to experience depends on education — for example, workers with more education may receive higher returns for each additional year of experience.

In short, the return to experience is **nonlinear** and **depends on education level**.

Q33(e)

Using the data, compute the marginal effect of EXPER on $\ln(\text{WAGE})$ for each individual in the sample.
Construct a histogram of these effects.
What is the range of values?

Ans

```
# 提取對應係數
beta4 <- coef(model_logwage)["exper"]
beta5 <- coef(model_logwage)["exper_sq"]
beta6 <- coef(model_logwage)["educ_exper"]

# 計算每位樣本對 EXPER 的邊際效果
cps5_small <- cps5_small %>%
  mutate(marginal_exper = beta4 + 2 * beta5 * exper + beta6 * educ)

# 畫出直方圖
hist(cps5_small$marginal_exper,
     main = "Histogram of Marginal Effects of EXPER on log(WAGE)",
     xlab = "Marginal Effect",
     col = "lightgreen",
     border = "gray")
```


Histogram of Marginal Effects of EXPER on log(WAGE)
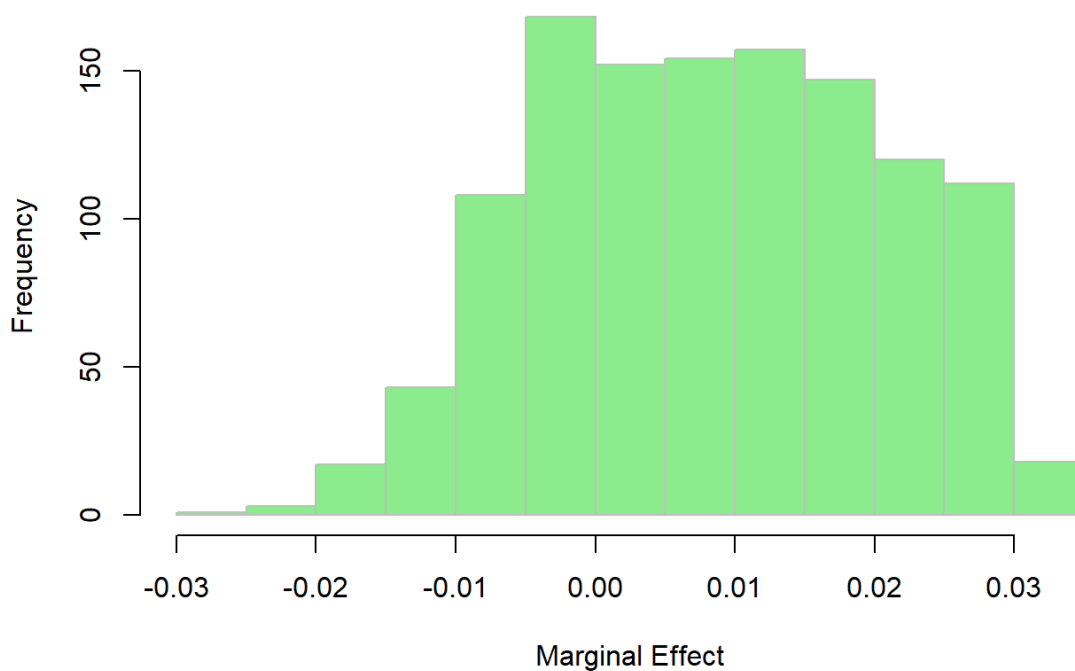
```
# 計算邊際效果範圍
range(cps5_small$marginal_exper)
```

```
## [1] -0.02527874  0.03398886
```

Using the formula derived in (d), we compute the marginal effect of labor market experience:

$$\frac{\partial \ln(\text{WAGE})}{\partial \text{EXPER}} = \hat{\beta}_4 + 2\hat{\beta}_5 \cdot \text{EXPER} + \hat{\beta}_6 \cdot \text{EDUC}$$

A histogram of these marginal effects shows that most values cluster around 0.01 to 0.02, with some negative values as well:

Histogram
Histogram
The **range of marginal effects** is:

- **Minimum**: −0.0253
- **Maximum**: 0.0340

Interpretation:

The return to an additional year of experience is **not constant** — it depends on both the individual's years of experience and their education level.
In this dataset, the marginal effect of experience ranges from **−2.5% to +3.4%** in log-wage terms, suggesting that for some individuals (likely those with very high experience and low education), additional experience may even slightly reduce log-wages, while for others it increases them.

Q33(f)

Let David have 15 years of education and 10 years of experience.
Let Svetlana have 17 years of education and 8 years of experience.
Compare their predicted $\ln(\mathrm{WAGE})$. Who is predicted to have the higher log-wage?

Ans

```r
# 取出係數
b <- coef(model_logwage)

# 建立變數：educ、educ_sq、exper、exper_sq、educ_exper
predict_logwage <- function(educ, exper) {
  educ_sq <- educ^2
  exper_sq <- exper^2
  educ_exper <- educ * exper

  # 套用估計模型
  logwage <- b["(Intercept)"] +
             b["educ"] * educ +
             b["educ_sq"] * educ_sq +
             b["exper"] * exper +
             b["exper_sq"] * exper_sq +
             b["educ_exper"] * educ_exper

  return(logwage)
}

# 分別計算 David 與 Svetlana 的預測 log-wage
logwage_david <- predict_logwage(educ = 15, exper = 10)
logwage_svet <- predict_logwage(educ = 17, exper = 8)

# 輸出結果
c(David = logwage_david, Svetlana = logwage_svet)
```

```
##    David.(Intercept) Svetlana.(Intercept)
##             2.959585             3.173228
```

We use the estimated model to compute predicted log-wages for:

- **David:** 15 years of education, 10 years of experience
- **Svetlana:** 17 years of education, 8 years of experience

Using the nonlinear model:

$$\ln(\mathrm{WAGE}) = \hat{\beta}_1 + \hat{\beta}_2 \cdot \mathrm{EDUC} + \hat{\beta}_3 \cdot \mathrm{EDUC}^2 + \hat{\beta}_4 \cdot \mathrm{EXPER} + \hat{\beta}_5 \cdot \mathrm{EXPER}^2 + \hat{\beta}_6 \cdot (\mathrm{EDUC} \cdot \mathrm{EXPER})$$

we obtain:

- $\ln(\widehat{\mathrm{WAGE}})_{\mathrm{David}} = 2.9596$
- $\ln(\widehat{\mathrm{WAGE}})_{\mathrm{Svetlana}} = 3.1732$

Conclusion:

Svetlana is predicted to have the higher log-wage. This is likely due to her having two additional years of education, which compensates for her slightly lower experience.

Q33(g)

Now suppose that both David and Svetlana gain 8 years of experience.
Compare their predicted log-wages again. Is the ranking the same as in part (f)?

Ans

```
# 更新後的經驗
logwage_david_18 <- predict_logwage(educ = 15, exper = 18)
logwage_svet_16 <- predict_logwage(educ = 17, exper = 16)

# 輸出比較結果
c(David = logwage_david_18, Svetlana = logwage_svet_16)
```

```
##    David.(Intercept) Svetlana.(Intercept)
##             3.092550             3.305006
```

We now update the inputs:

- **David:** EDUC = 15, EXPER = 18
- **Svetlana:** EDUC = 17, EXPER = 16

Using the same estimated model, we obtain the new predictions:

- $\ln(\widehat{\text{WAGE}})_{\text{David}} = 3.0926$
- $\ln(\widehat{\text{WAGE}})_{\text{Svetlana}} = 3.3050$

Conclusion:

Even after both individuals gain 8 more years of experience, **Svetlana** continues to have a higher predicted log-wage than David. This suggests that her **higher level of education** still outweighs David's relative advantage in experience.

Q33(h)

Let Wendy have 13 years of education and 15 years of experience.
Let Jill have 18 years of education and 15 years of experience.
Formulate and test a hypothesis that the marginal effect of experience on log-wages is the same for both individuals.

Ans

```
# 顯示 educ_exper 係數與 p-value
summary(model_logwage)$coefficients["educ_exper", ]
```

```
##      Estimate    Std. Error       t value      Pr(>|t|)
## -0.0010102359  0.0003790763 -2.6649938155  0.0078027810
```

Let:

- Wendy: EDUC = 13, EXPER = 15
- Jill: EDUC = 18, EXPER = 15

The marginal effect of experience is:

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

So the difference in marginal effect between Jill and Wendy is:

$$\lambda_{\text{Jill}} - \lambda_{\text{Wendy}} = \beta_6 \cdot (18 - 13) = 5 \cdot \beta_6$$

We test:

- H_0: $\beta_6 = 0$ (no difference)
- H_1: $\beta_6 \neq 0$ (difference exists)

From the regression output:

- $\hat{\beta}_6 = -0.00101$,
- p-value = 0.0078

Conclusion: Since the p-value is less than 0.05, we **reject the null hypothesis**.
There is significant evidence that Wendy and Jill have **different marginal returns to experience** — the return to experience **decreases more** for individuals with higher education (like Jill).

Q33(i)

For Jill, at what level of experience does the marginal effect of experience on $\ln(\text{WAGE})$ become negative?
Construct a 95% confidence interval for this value.

Ans

```
# 取出估計值
b <- coef(model_logwage)
b4 <- b["exper"]
b5 <- b["exper_sq"]
b6 <- b["educ_exper"]
educ_jill <- 18

# 點估計的 EXPER* 使得邊際效果 = 0
ex_star <- -(b4 + b6 * educ_jill) / (2 * b5)

# 建立 gradient 向量 u = [∂/∂β4, ∂/∂β5, ∂/∂β6]
# ∂/∂β4 = -1 / (2β5)
# ∂/∂β5 = (β4 + 18β6) / (2β5²)
# ∂/∂β6 = -18 / (2β5)
u <- numeric(length(b))
names(u) <- names(b)
u["exper"] <- -1 / (2 * b5)
u["exper_sq"] <- (b4 + b6 * educ_jill) / (2 * b5^2)
u["educ_exper"] <- -educ_jill / (2 * b5)

# 計算變異數與標準差
vcov_mat <- vcov(model_logwage)
var_ex_star <- t(u) %*% vcov_mat %*% u
se_ex_star <- sqrt(var_ex_star)

# 95% 信賴區間
lower <- ex_star - 1.96 * se_ex_star
upper <- ex_star + 1.96 * se_ex_star

# 輸出結果
c(Estimate = ex_star, Std_Dev = se_ex_star, Lower_95 = lower, Upper_95 = upper)
```

```
## Estimate.exper      Std_Dev     Lower_95     Upper_95
##      28.518536     2.261773    24.085461    32.951611
```

We consider Jill, who has 18 years of education. Her marginal effect of experience is given by:

$$\frac{\partial \ln(\text{WAGE})}{\partial \text{EXPER}} = \beta_4 + 2\beta_5 \cdot \text{EXPER} + \beta_6 \cdot 18$$

Setting the marginal effect equal to zero, we solve for the critical value of EXPER:

$$\text{EXPER}^* = \frac{-(\beta_4 + 18 \cdot \beta_6)}{2 \cdot \beta_5} = 28.52$$

Thus, Jill's return to experience becomes negative after approximately **28.52 years** in the labor market.

Using the delta method, we construct a **95% confidence interval** for this value:

$$[24.09, \ 32.95]$$

Conclusion: We are 95% confident that Jill's marginal return to experience becomes negative between **24.1 and 33.0 years** of experience.