

# Homework 03.11 作業解答

Use the `lm()` function to perform a simple linear regression on the Boston data set with `medv` as the response and `rm` as the predictor. Use the `summary()` function to print the results. Comment on the output. The simple linear model:

$$(medv) = \beta_0 + \beta_1(rm) + \varepsilon, E(\varepsilon) = 0, Var(\varepsilon) = \sigma^2$$

1. Find out  $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2$ .

```
library(ISLR2)
```

```
## Warning: 套件 'ISLR2' 是用 R 版本 4.3.3 來建造的
```

```
r<-lm(medv~rm,data=Boston)
summary(r)
```

```
##
## Call:
## lm(formula = medv ~ rm, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -23.346  -2.547   0.090   2.986  39.433
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -34.671      2.650  -13.08  <2e-16 ***
## rm              9.102      0.419   21.72  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.616 on 504 degrees of freedom
## Multiple R-squared:  0.4835, Adjusted R-squared:  0.4825
## F-statistic: 471.8 on 1 and 504 DF,  p-value: < 2.2e-16
```

$$\hat{\beta}_0 = -34.671,$$

$$\hat{\beta}_1 = 9.102,$$

$$\hat{\sigma} = 6.616.$$

2. Provide an interpretation of  $\hat{\beta}_1$ .

該區域的平均房間數每增加一間，房價的中位數預期增加 9102 美元 ( $1000 \times \hat{\beta}_1$ )。

3. Based on the results, is there a relationship between the predictor and the response? Use the p-value in the printed result, and write down the null and alternative hypotheses.

虛無假設 ( $H_0$ ) 與對立假設 ( $H_1$ ) 分別為:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

此假設檢定的  $p\text{-value} < 2 \times 10^{-16}$  , 因此  $\beta_1$  顯著不為零 , 此筆資料有足夠的證據顯示 predictor 和 response 相關。

4. Find out the 95% confidence interval of  $\beta_0, \beta_1$ .

```
round(confint(r),digits=3)
```

```
##           2.5 %   97.5 %  
## (Intercept) -39.877 -29.465  
## rm          8.279    9.925
```

$\beta_0$  的 95% 信賴區間為 (-39.877, -29.465) ; ,  $\beta_1$  的 95% 信賴區間為 (8.279, 9.925)。

5. According to the printed results of `summary()`, write down the equation describing the prediction for  $y_i$  based on the value of  $x_i$ .

$$\hat{y}_i = -34.674 + 9.102x_i$$

6. If we have  $rm$  equal to 6, what is the predicted value of  $medv$  based on  $rm = 6$ ?

```
new=data.frame("rm"=6)  
predict(r,new)
```

```
##           1  
## 19.94203
```

```
#-34.674 + 9.102*6
```

$$\hat{y}_i = -34.674 + 9.102 \times 6 = 19.938(\text{or } 19.942)$$

故當  $rm = 6$  ,  $medv$  的預測值為 19.938。