

> **summary**(lm(SALARY ~ PUB, data = lab4))

Regression table:

Call: lm(formula = SALARY ~ PUB, data = lab4)

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|---------|--------|--------|---------|
| -21638.8 | -8327.3 | 697.2 | 7456.6 | 20322.9 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|--------------|
| (Intercept) | 47940.4 | 2511.7 | 19.09 | < 2e-16 *** |
| PUB | 1148.2 | 146.1 | 7.86 | 2.58e-11 *** |
| --- | | | | |

\hat{b}_0 S_{b0} \hat{b}_1 S_{b1}

Significantly differ from zero

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9707 on 73 degrees of freedom

Multiple R-squared: 0.4584, Adjusted R-squared: 0.4509

F-statistic: 61.78 on 1 and 73 DF, p-value: 2.583e-11

> **anova**(lm(SALARY ~ PUB, data = lab4))

Analysis of Variance Table

$$R^2 = \text{RSS}/\text{TSS} = \text{RSS}/(\text{RSS}+\text{ESS})$$

Response: SALARY

Sum of Squares Regression : RSS

$$\text{df1} = p = 1$$

Mean Square Regression: $\text{RSS}/\text{df1}$

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|-----------|----|------------|------------|---------|---------------|
| PUB | 1 | 5821421788 | 5821421788 | 61.775 | 2.583e-11 *** |
| Residuals | 73 | 6879172447 | 94235239 | | |

Sum of Squares Error : ESS

Mean Square Error: $\text{ESS}/\text{df2}$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$$\text{df2} = n - 1 - p = 75 - 1 - 1 = 73$$

$$F = (\text{RSS}/\text{df1})/(\text{ESS}/\text{df2})$$

A significant F-value means that X can account for a substantial proportion of variation in Y statistically. Statistically!

```
> vcov( lm(SALARY ~ PUB, data= lab4) )
```

Variance/Covariance matrix:

| | (Intercept) | PUB |
|-------------|-------------|------------|
| (Intercept) | 6308830.5 | -328359.66 |
| PUB | -328359.7 | 21340.53 |

The square root of the diagonals is same as “standard error” in the regression table.

```
> vcov.lab4<- vcov( lm(SALARY ~ PUB, data= lab4) )
```

```
> sqrt(diag( vcov.lab4 ) )
```

| | (Intercept) | PUB |
|----------|-------------|---------|
| S_{b0} | 2511.739 | 146.084 |
| S_{b1} | | |

> **confint** (lm(SALARY ~ PUB, data= lab4))

Confidence interval (CI) :

| | 2.5 % | 97.5 % | |
|-------------|------------|-----------|--------------|
| (Intercept) | 42934.4632 | 52946.239 | CI for b_0 |
| PUB | 857.0364 | 1439.326 | CI for b_1 |

We believe that the probability is 0.95 that:

(1)The “true value of b_0 ” would be between 42934.46 and 52946.24.

(2)The “true value of b_1 ” would be between 857.04 and 1439.33.

CI is another version of hypothesis testing. If the null hypothesized value does not fall in the interval, we would reject the null hypothesis.