

Applied Statistical Methods - Solution 3

AUTHOR
Peter von Rohr

PUBLISHED
March 11, 2024

WEBR STATUS
● Ready!

Problem 1: Simple Linear Regression

Use the data on `Width` and `Height` that we collected during the class to fit a linear regression model to the data. Use the formula derived in class to compute estimates \hat{b}_0 and \hat{b}_1 for the intercept (b_0) and the slope (b_1) that define the linear regression line. Verify the obtained results using the function `lm()` in R.

Tasks

- Read data from https://charlotte-ngs.github.io/asmasss2024/data/asm_width_height.csv using the function `read.table()`

▶ Run Code



```
1 # read data to data.frame
2 s_data_url_p1 <- "https://charlotte-ngs.github.io/asmasss2024/data/a
3 df_width_height <- read.table(s_data_url_p1, header = T, sep = ",")
4 df_width_height
```

	Width	Height
1	82	185
2	65	168
3	76	168
4	80	193
5	78	180
6	70	181
7	72	182
8	70	169
9	65	165
10	73	170

- Compute estimate \hat{b}_1 for the slope b_1 using the formula from class given as

$$\hat{b}_1 = \frac{(xy)_{\cdot} - N\bar{x} \cdot \bar{y}}{(x^2)_{\cdot} - N\bar{x}^2}$$

▶ Run Code



```
1 # estimate of slope
2 N <- nrow(df_width_height)
3 x_y_dot <- sum(df_width_height$Width * df_width_height$Height)
4 x_dot_bar <- mean(df_width_height$Width)
5 y_dot_bar <- mean(df_width_height$Height)
6 x2_dot <- sum(df_width_height$Width^2)
7 b_1_hat <- (x_y_dot - N * x_dot_bar * y_dot_bar) / (x2_dot - N * x_dot_bar^2)
8 b_1_hat
```

[1] 1.119009

- Compute estimate \hat{b}_0 for the slope b_0 using the formula from class given as

$$\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}.$$

▶ Run Code



```
1 # estimate intercept
2 b_0_hat <- y_dot_bar - b_1_hat * x_dot_bar
3 b_0_hat
```

[1] 94.30042

- Validation with R

▶ Run Code



```
1 # validation in R using lm()
2 lm_wei_wid <- lm(Height ~ Width, data = df_width_height)
3 summary(lm_wei_wid)
```

Call:

```
lm(formula = Height ~ Width, data = df_width_height)
```

Residuals:

Min	1Q	Median	3Q	Max
-11.345	-3.232	-1.321	5.589	9.179

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	94.3004	29.0898	3.242	0.0118 *
Width	1.1190	0.3968	2.820	0.0225 *

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

Residual standard error: 6.996 on 8 degrees of freedom

Multiple R-squared: 0.4985, Adjusted R-squared: 0.4358

F-statistic: 7.953 on 1 and 8 DF, p-value: 0.02249