

1. Problem

Using the data provided in `regression.csv` estimate a linear regression of y on x_1 and x_2 . Answer the following questions.

- Proportion of variance explained (in percent):
- F-statistic:
- Characterize in your own words how the response y depends on the regressors x_1 and x_2 .
- Upload the R script you used to analyze the data.

Solution

The presented results describe a semi-logarithmic regression.

Call:

```
lm(formula = log(y) ~ x1 + x2, data = d)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -2.68802 | -0.67816 | -0.01803 | 0.68866 | 2.35064 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|--------------|
| (Intercept) | -0.06802 | 0.13491 | -0.504 | 0.616 |
| x1 | 1.37863 | 0.13351 | 10.326 | 9.34e-15 *** |
| x2 | -0.21449 | 0.13995 | -1.533 | 0.131 |

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

Residual standard error: 1.052 on 58 degrees of freedom

Multiple R-squared: 0.6511, Adjusted R-squared: 0.6391

F-statistic: 54.12 on 2 and 58 DF, p-value: 5.472e-14

The mean of the response y increases with increasing x_1 . If x_1 increases by 1 unit then a change of y by about 296.94 percent can be expected. Also, the effect of x_1 is significant at the 5 percent level.

Variable x_2 has no significant influence on the response at 5 percent level.

The R-squared is 0.6511 and thus 65.11 percent of the variance of the response is explained by the regression.

The F-statistic is 54.12.

- Proportion of variance explained: 65.11 percent.
- F-statistic: 54.12.
- Characterization: semi-logarithmic.
- R code.