DSA 8020 R Lab 1: Simple Linear Regression

Whitney

January 08, 2024

Contents

Leaning Tower of Pisa

The dataset `PisaTower.csv` provides annual measurements of the lean (the difference between where a point on the tower would be if the tower were straight and where it actually is) from 1975 to 1987. We would like to characterize lean over time by fitting a simple linear regression.

Load the dataset

Code:

```
PisaTower <- read.csv("PisaTower.csv")
head(PisaTower)
```

```
##     lean year
## 1 2.9642 1975
## 2 2.9644 1976
## 3 2.9656 1977
## 4 2.9667 1978
## 5 2.9673 1979
## 6 2.9688 1980
```

Descriptive analysis

Numerical summary

Provide some numerical summaries to describe the response and the predictor variables, respectively, as well as their relationship.

Code:

Graphical summary

Provide graphical summaries through plots to describe the response and predictor variables, respectively, as well as their relationship.

Code:

Question: Describe the direction, strength, and the form of the relationship.

Answer:

Simple linear regression

1. Identify the response variable, the predictor variable, and the sample size.

Answer:

2. Fit a simple linear regression.

Code:

3. Write down the fitted linear regression model.

Answer:

4. What is $\hat{\sigma}$, the estimate of σ?

Answer:

5. Find a 95% confidence interval for β_1.

Code:

6. Test the following hypothesis: $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$ with $\alpha = 0.05$

Answer:

7. Construct a 90% confidence interval for E[`lean`] in year 1984

Code:

8. Use residuals to check model assumptions.

Code:
Answer:

9. Would it be a good idea to use the fitted linear regression equation to predict `lean` in year 2010? Explain your answer.

Answer: