Math 300 Lesson 13 Notes

Simple Linear Regression - Related Topics

YOUR NAME HERE

June, 2022

Contents

Objectives

1. Explain and give an example of a confounding variable.

2. In linear regression, explain what "best fit" means and calculate the sum of squared errors.

Reading

Chapter 5.3 - 5.4

Lesson

Work through the learning check LC5.8. Complete code as necessary.

- That correlation does not imply causation is an important idea, but decision makers and humans in general want to know causation. There are other courses that help, DOE and econometrics. We will not explore

- If there is time, play the correlation game.

Setup

```
library(tidyverse)
library(moderndive)
```

```
## Warning: package 'moderndive' was built under R version 4.1.3
```

- Spurious correlations (Objective 1)

Spend some time making sure you understand confounding variables. The spurious correlations website may give you some ideas.

LC 5.8 (Objective 2)

(LC5.8) Note in the following plot there are 3 points marked with dots along with:

- The "best" fitting solid regression line
- An arbitrarily chosen dotted line
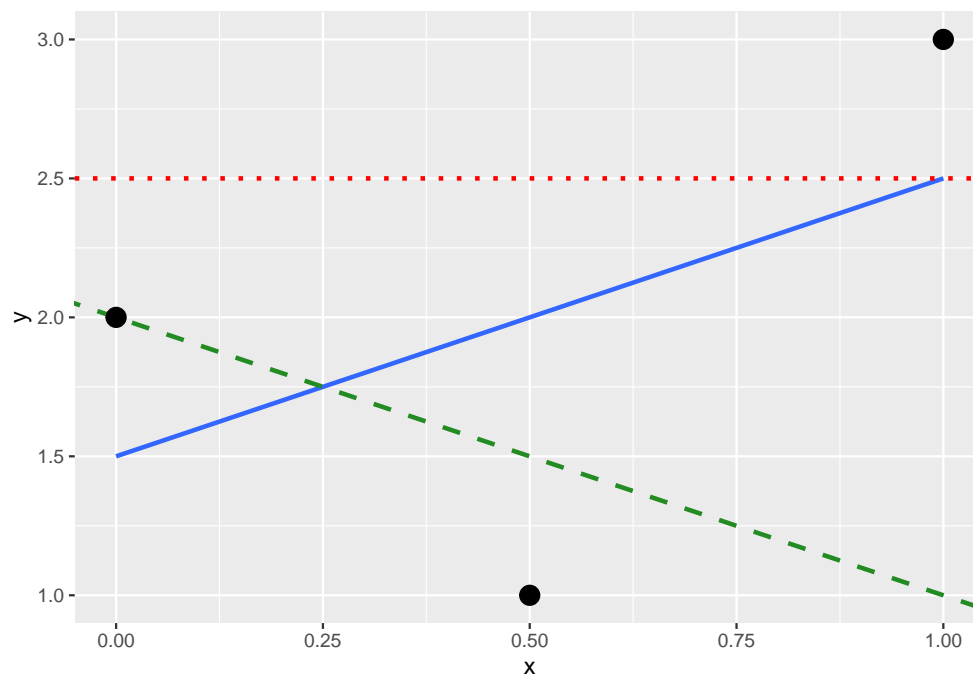- Another arbitrarily chosen dashed line

Figure 1: Regression line and two others.

Compute the sum of squared residuals for each line and show that of these three lines, the regression line has the smallest value.

Solution:

- The "best" fitting solid regression line:

$$\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 = (2.0 - 1.5)^2 + (1.0 - 2.0)^2 + (3.0 - 2.5)^2 = 1.5$$

```
# Complete the code and remove the comment
#sum((c(2, 1, 3)-c(_____,_____,_____))^2)
```

- An arbitrarily chosen dotted line:

$$\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 = (2.0 - 2.5)^2 + (1.00 - 2.5)^2 + (3.0 - 2.5)^2 = 2.75$$

```
# Complete the code and remove the comment
#sum((c(2, 1, 3)-c(_____,_____,_____))^2)
```

- Another arbitrarily chosen dashed line:

$$\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 = (2.0 - 2.0)^2 + (1.0 - 1.5)^2 + (3.0 - 1.0)^2 = 4.25$$

```
# Complete the code and remove the comment
#sum((c(2, 1, 3)-c(_____,_____,_____))^2)
```

Documenting software

- File creation date: 2022-06-23
- R version 4.1.1 (2021-08-10)
- `tidyverse` package version: 1.3.1
- `moderndive` package version: 0.5.4