YaRrr! The Pirate's Guide to R

Nathaniel D. Phillips

2017-03-17

Contents

Chapter 1

Preface

Chapter 2

Getting Started

Chapter 3

Jump In!

Chapter 4

The Basics

Chapter 5

Scalers and vectors

Chapter 6

Vector functions

Chapter 7

Indexing Vectors with [ ]

Chapter 8

Matrices and Dataframes

Chapter 9

Importing, saving and managing data

Chapter 10

Advanced dataframe manipulation

Chapter 11

Plotting (I)

Chapter 12

Plotting (II)

Chapter 13

Hypothesis Tests

Chapter 14

ANOVA

Chapter 15

Regression

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
## Loading required package: jpeg
## Loading required package: BayesFactor
## Loading required package: coda
## Loading required package: Matrix
## ************
## Welcome to BayesFactor 0.9.12-2. If you have questions, please contact Richard Morey (richarddmorey@gmail
##
## Type BFManual() to open the manual.
## ************
## yarrr v0.1.5. Citation info at citation('yarrr'). Package guide at yarrr.guide()
## Email me at Nathaniel.D.Phillips.is@gmail.com
```

Figure 15.1: Insert funny caption here.

Pirates like diamonds. Who doesn't?! But as much as pirates love diamonds, they hate getting ripped off. For this reason, a pirate needs to know how to accurately assess the value of a diamond. For example, how much should a pirate pay for a diamond with a weight of 2.0 grams, a clarity value of 1.0, and a color gradient of 4 out of 10? To answer this, we'd like to know how the attributes of diamonds (e.g.; weight, clarity, color) relate to its value. We can get these values using linear regression.

15.1 The Linear Model

The linear model is easily the most famous and widely used model in all of statistics. Why? Because it can apply to so many interesting research questions where you are trying to predict a continuous variable of interest (the *response* or *dependent variable*) on the basis of one or more other variables (the *predictor* or *independent variables*).

The linear model takes the following form, where the x values represent the predictors, while the beta values represent weights.

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... \beta_n x_n$

For example, we could use a regression model to understand how the value of a diamond relates to two independent variables: its weight and clarity. In the model, we could define the value of a diamond as $\beta_{weight} \times weight + \beta clarity \times clarity$. Where β_{weight} indicates how much a diamond's value changes as a function of its weight, and $\beta_{clarity}$ defines how much a diamond's value change as a function of its clarity.

15.2 Linear regression with `lm()`

To estimate the beta weights of a linear model in R, we use the `lm()` function. The function has three key arguments: `formula`, and `data`

Chapter 16

Solutions

Chapter 17

Placeholder

Bibliography