functions

06 February, 2023

Write the R code to answer the following questions. You have until the beginning of next class to answer all of the questions below and commit to GitHub, both the .Rmd file and the .pdf.

Question 1

Use the following code block to generate three explanatory variables (`x1`, `x2`, `x3`) and one outcome (`y`).

```
set.seed(12345)
n <- 20 # 20 observations in the sample
x1 <- runif(n, min = -1, max = 1) # Sample uniformly from [-1,1]
x2 <- runif(n, min = -1, max = 1) # Sample uniformly from [-1,1]
x3 <- runif(n, min = -1, max = 1) # Sample uniformly from [-1,1]
noise <- rnorm(n, mean = 0, sd = 0.3) # We have a noisy sample
y <- 1 + x1 + 2*x2 + 3*x3 + noise # generate the response variable
df <- data.frame(x1 = x1,
                 x2 = x2,
                 x3 = x3,
                 y = y)
```

1a

Create a function called `rmse` that reports the root mean squared error (RMSE) of model predictions. Your function must take as inputs a vector `y` and an `lm` output.

For N observations indexed by i, the equation for RMSE is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}}$$

```
#code here
```

1b

Run a bivariate regression with `x1` as the predictor and `y` as the outcome, and then run your `rmse` function using the fitted model.

```
#code here
```

1c

Write a for loop that loops over `x1`, `x2`, and `x3`, calculating a bivariate regression in each iteration, and then calling the `rmse` function you wrote in 1a. Have your loop print the `rmse` output.

```
#code here
```

1d

Turn your for loop into a function such that for **any** x and y, the function runs a bivariate regression, it runs your `rmse` function, and it returns the single RMSE output.

```
#code here
```