

# Lab #13 - Resampling Methods

*Econ 224*

*October 23rd, 2018*

## Introduction

In this lab you will work through Section 5.3 of ISL and record your code and results in an RMarkdown document. I have added section headings below to help you organize your results. You do not have to submit this lab, so you don't have to type up a detailed description of what you've done. However, I'd suggest that you write down some notes for your own future reference. These will be helpful on the problem set.

You do not need to follow the code in ISL exactly: feel free to use your preferred coding style. In particular, rather than using the `attach` command suggested by ISL, feel free to instead use `dplyr` or related commands to produce more readable and error-resistant code. If you choose to do this, you will need to select the rows of a tibble *by position*, i.e. by row index. You can do this using the `slice` function in `dplyr`.

You will need the ISLR package for the lab, so please install it if you have not done so already. Note that this lab uses two different datasets: `Auto` and `Portfolio`. Both of these are included with ISLR. Make sure to read the documentation for each dataset in the R help files before proceeding.

## The Validation Set Approach

Work through section 5.3.1 of ISL and add your code and results below.

```
library(ISLR)
library(tidyverse)
#----- create training and test samples
set.seed(1)
train_indices <- sample(nrow(Auto), 196)
train <- Auto %>% slice(train_indices)
test <- Auto %>% slice(-train_indices)

#----- fit regressions using training sample
fit1 <- lm(mpg ~ horsepower, train)
fit2 <- lm(mpg ~ poly(horsepower, 2), train)
fit3 <- lm(mpg ~ poly(horsepower, 3), train)

#----- predict test data
predicted1 <- predict(fit1, test)
predicted2 <- predict(fit2, test)
predicted3 <- predict(fit3, test)

#----- compare MSE of predictions
test %>% summarize(MSE1 = mean((mpg - predicted1)^2),
                  MSE2 = mean((mpg - predicted2)^2),
                  MSE3 = mean((mpg - predicted3)^2))
```

	MSE1	MSE2	MSE3
1	26.14142	19.82259	19.78252

```
#----- try re-running with a different seed
```

## Leave-One-Out Cross-Validation

Work through section 5.3.2 of ISL and add your code and results below.

```
library(boot)
#----- Function to automate LOO-CV for polynomial fits
cv_poly <- function(degree) {
  glm_fit <- glm(mpg ~ poly(horsepower, degree), data = Auto)
  cv_error <- cv.glm(Auto, glm_fit)
  cv_error$delta
}
sapply(1:5, cv_poly)
```

```
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 24.23151 19.24821 19.33498 19.42443 19.03321
[2,] 24.23114 19.24787 19.33448 19.42371 19.03242
```

## k-Fold Cross-Validation

Work through section 5.3.3 of ISL and add your code and results below.

## The Bootstrap

Work through section 5.3.4 of ISL and add your code and results below.