Coursera Data Science Project: Statistical Inference

Rob Rolleston

August 23, 2015

Introduction

This is the project for the statistical inference class. In it, I will use simulation to explore inference and do some simple inferential data analysis. The project consists of two parts:

1. A simulation exercise.
2. Basic inferential data analysis.

Simulation Excercise

tbd

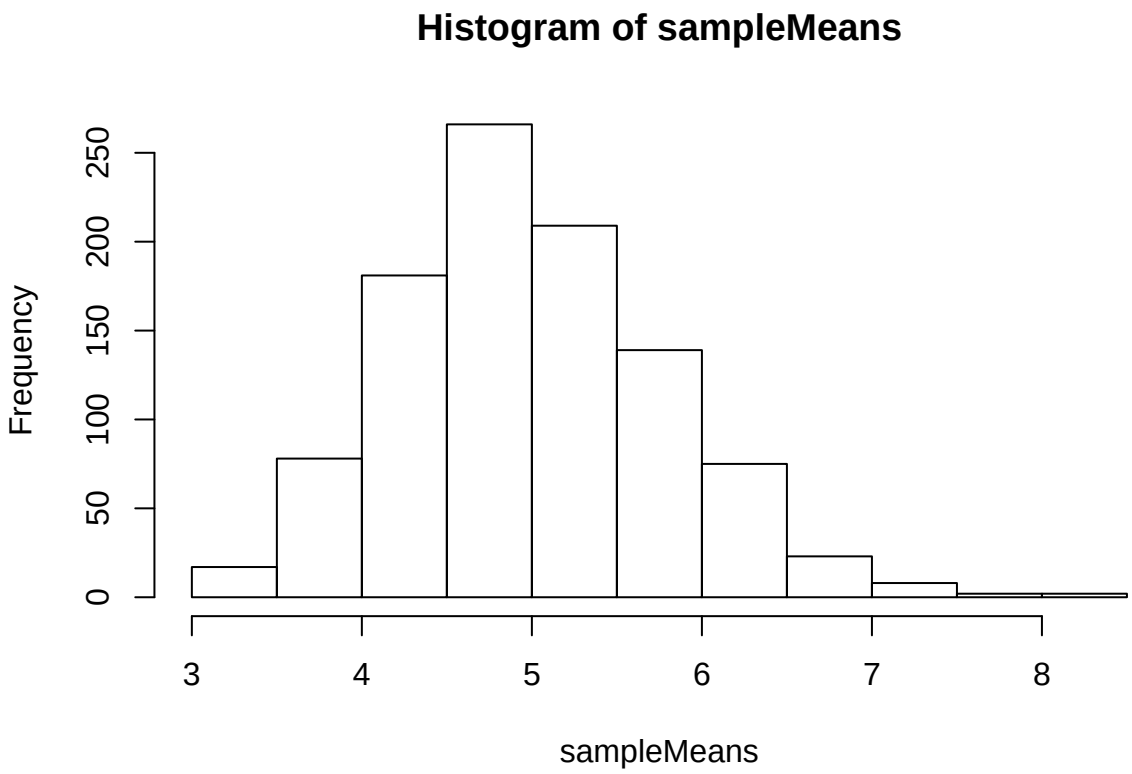
Simulations

```
nosim <- 1000
n <- 40
lambda <- 0.2
populationMean <- 1./lambda
populationSD <- 1./lambda
sampleData <- matrix(rexp(nosim * n, rate=lambda), nosim)
```

Sample Mean versus Theoretical Mean

```
sampleMeans <- apply(sampleData, 1, mean)
sampleMean <- mean(sampleMeans)
sampleSD <- sd(sampleMeans)

inferredSD <- populationSD/sqrt(n)
```



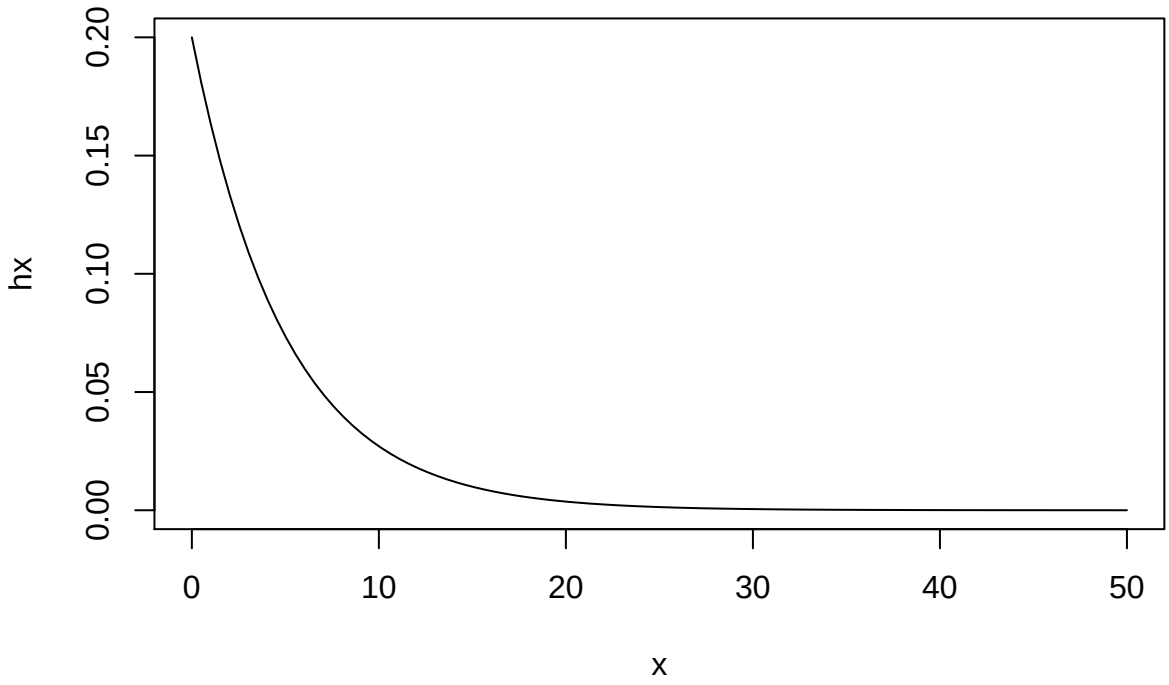
Sample Variance versus Theoretical Variance

tbd

Distribution

tbd

```
x <- seq(0,50, length=100)
hx <- dexp(x, rate=lambda)
```



Basic Inferential Data Analysis

tbd

Load and explore data

tbd

```
library(datasets)
toothGrowth <- ToothGrowth
```

Data Summary

tbd

Compare tooth growth by supp and dose

tbd

Conclusions

tbd