Statistical Inference Project 2 - Part 1

Daniel

1/19/2020

Loading and preprocessing the data

Set Global Options

```
knitr::opts_chunk$set(warnings=FALSE, message=FALSE, echo = TRUE)
```

Overview

The goal of this part of the project is to compare the exponential distribution wiht the Central Limit Theorem (CLM). This will be done using the R function rexp(n, lambda), with the distribution of 40 averages and 1000 simulations

Question 1

Show the sample mean and compare it to the thoretical mean of the distribution

Simulation

```
set.seed(100)
lambda <- 0.2
num_sim <- 1000
n <- 40

# Perform simulation
simulation <- matrix(rexp(num_sim *n, lambda), num_sim, n)
```

```
# Find the means of the samples
row_means <- rowMeans(simulation)
# Find the mean of the means
sample_mean <- mean(row_means)
# Find the Theoretical mean
mean_theo <- 1/lambda
sample_mean
```

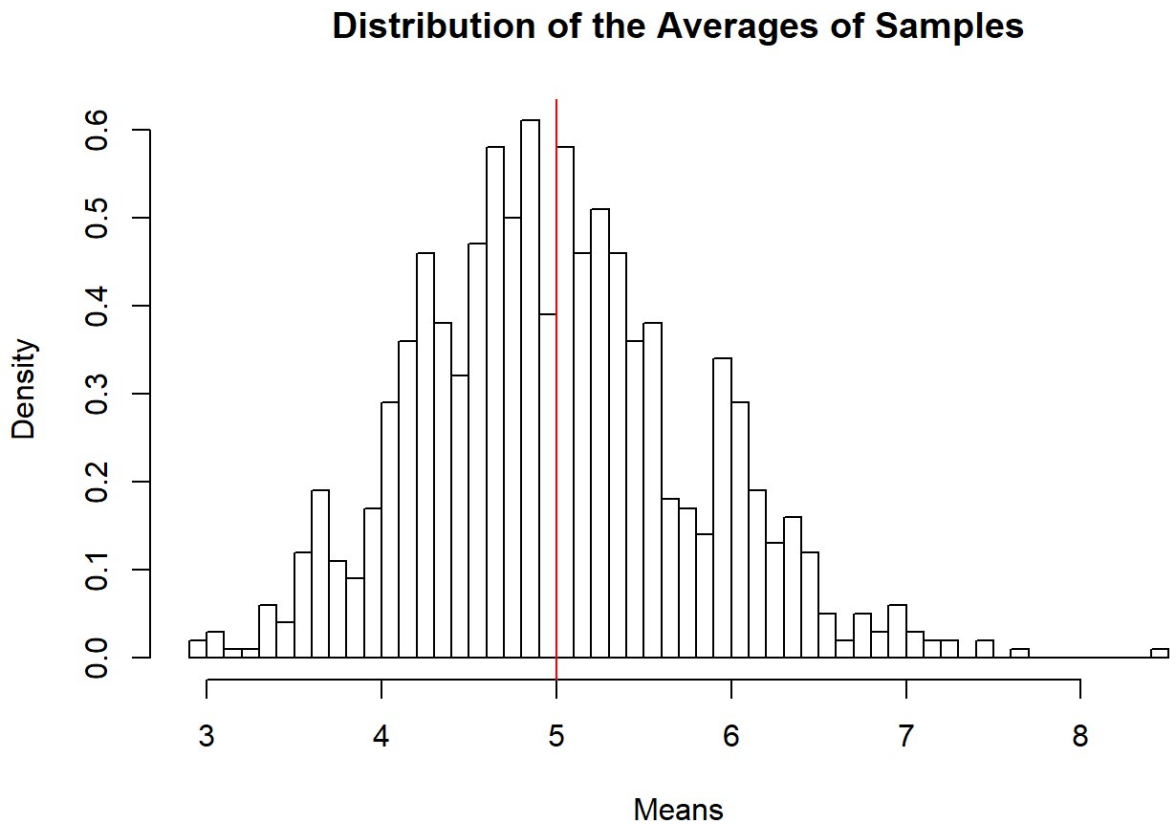
```
## [1] 4.999702
```

```
mean_theo
```

```
## [1] 5
```

The sample mean and theoretical mean are very similar. The sample mean is given by 4.99 and the theoretical mean is 5.

```
hist(row_means, breaks = 40, prob = TRUE, main = "Distribution of the Averages of Samp
les",
     xlab = "Means")
# Sample mean
abline(v = sample_mean, col = "blue")
# Theoretical mean
abline(v = 1/lambda, col = "red")
```



Question 2

Show how variable the sample is (via varience) and compare it to the theoretical variance of the distribution

```
sample_variance <- sd(row_means)^2
theoretical_variance <- ((1/lambda)/sqrt(n))^2
# Call both values
sample_variance
```

```
## [1] 0.6335302
```

```
theoretical_variance
```

```
## [1] 0.625
```

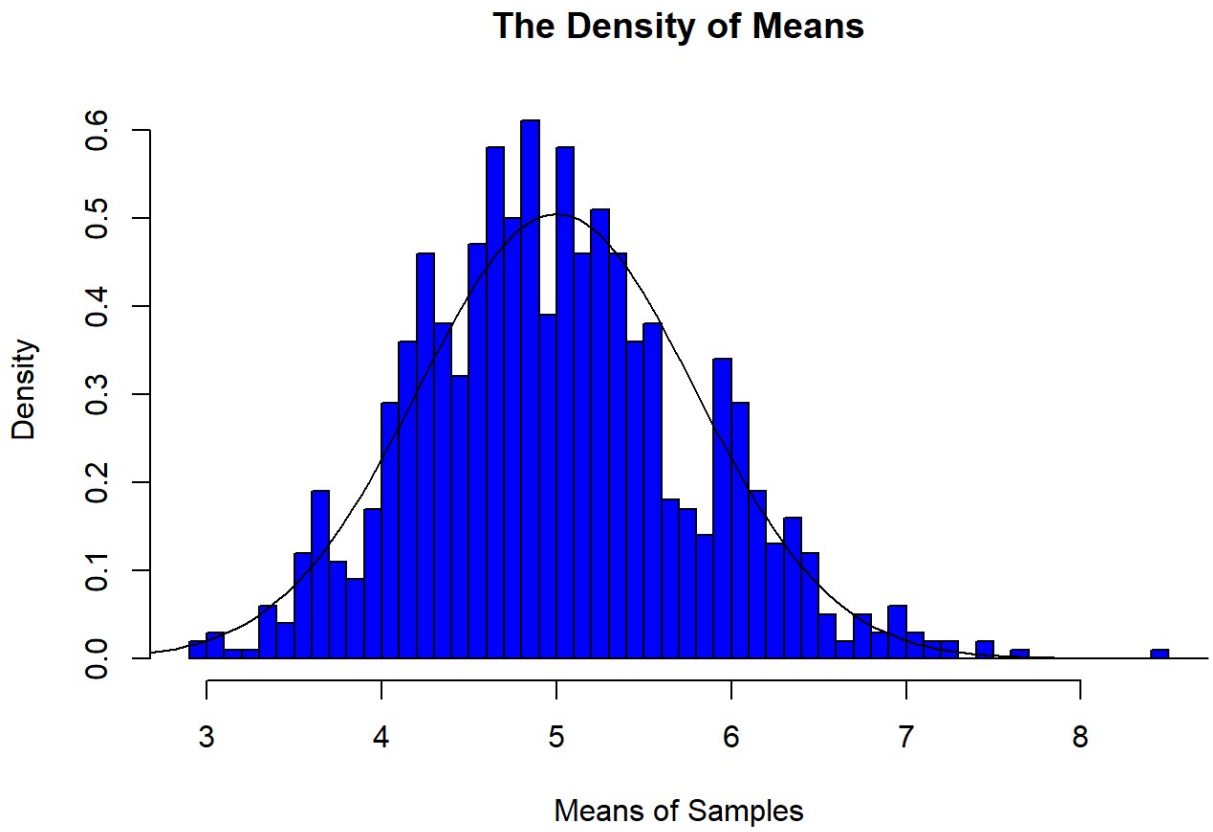
The variance for the sample 0.63353 is and the variance for the theoretical is 0.625

Question 3

Show that the distribution is approximately normal

We will plot a histogram using the sample means and use a line plot to show that the shape of the graph is normal

```
x <- seq(0,10,length = 256)
y <- dnorm(x, mean = 1/lambda, sd = 1/lambda/sqrt(n))
hist(row_means, breaks = 40, prob = T, col = "blue", main = "The Density of Means", xl
ab = "Means of Samples", ylab = "Density")
lines(x, y, pch = 22, col = "black")
```



We can see here that the sample means follow a theoretical normal distribution fairly well.