Statistical Inference Course Project Part 1

Muhammad Imran Saeed

Installation and setup..

```
listOfPackages <- c("ggplot2","plyr","tinytex")
newPackages <- listOfPackages[!(listOfPackages %in% installed.packages()[,"Package"])]

if(length(newPackages)) {
    message(sprintf("Going to install package(s)[ %s ]\n", newPackages))
    install.packages(newPackages)
} else {
    message("All packages are already installed. Will skip install pahse")
}
```

Installing and lodaing all the packages needed for this assignment. Will install only if not installed.

```
## All packages are already installed. Will skip install pahse
```

```
## Load all of the package needed for this project.
for(package in listOfPackages){
  library(package, character.only = TRUE)
}
knitr::opts_chunk$set(echo = TRUE)
```

Comparison of Means Sample Mean vs Theoretical Mean of Distribution

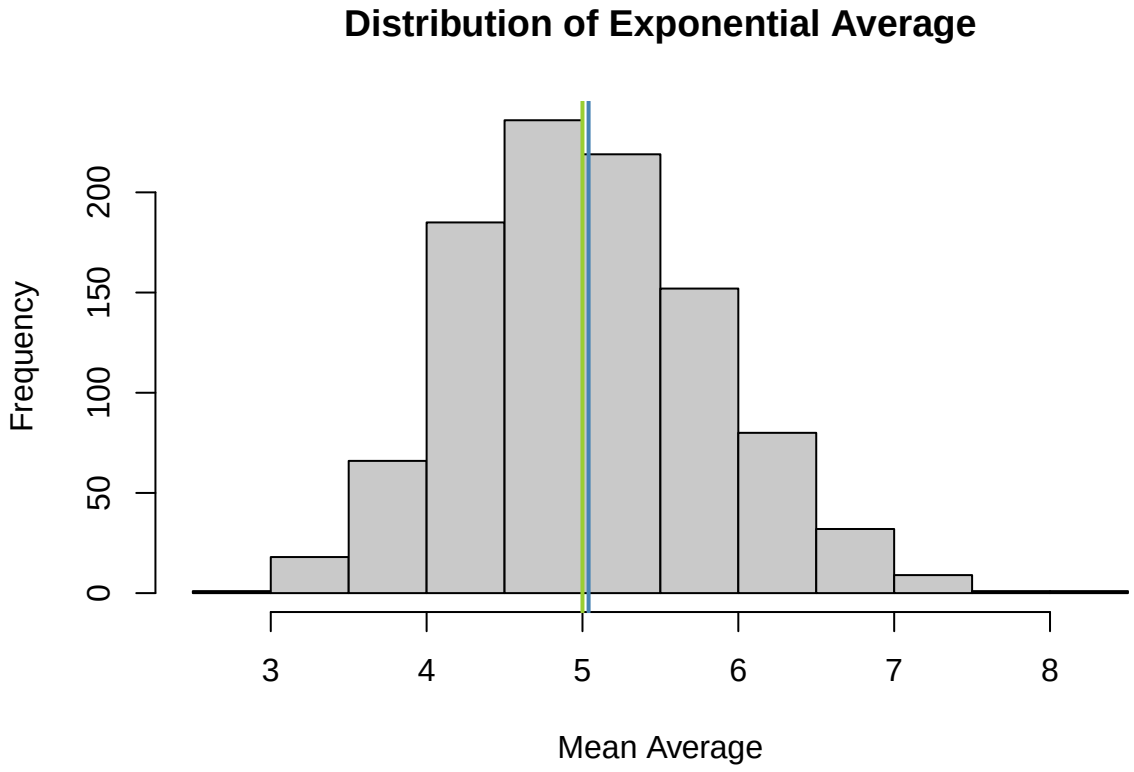
```
# Sample Mean
sampleMean <- mean(meanSimulationData) # Mean of sample means
print (paste("Sample Mean = ", sampleMean))
```

```
## [1] "Sample Mean =  5.03930581733264"
```

```
# Theoretical Mean
# the expected mean of the exponential distribution of rate = 1/lambda
theoreticalMean <- (1/lambda)
print (paste("Theoretical Mean = ", theoreticalMean))
```

```
## [1] "Theoretical Mean =  5"
```

```r
# Histogram shows differences
hist(meanSimulationData, col="grey79", xlab = "Mean Average", main="Distribution of Exponential Average
abline(v = theoreticalMean, col="yellowgreen",lwd=2)
abline(v = sampleMean, col="steelblue",lwd=2)
```



Question 2: Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution

Calculating theoretical and sample variance

```r
# sample deviation & variance
sampleDeviation <- sd(meanSimulationData)
sampleDeviation
```

```
## [1] 0.7953828
```

```r
sampleVariance <- sampleDeviation^2
sampleVariance
```

```
## [1] 0.6326338
```

```r
# theoretical deviation & variance
theoreticalDev <- (1/lambda)/sqrt(n)
theoreticalDev
```

```
## [1] 0.7905694
```

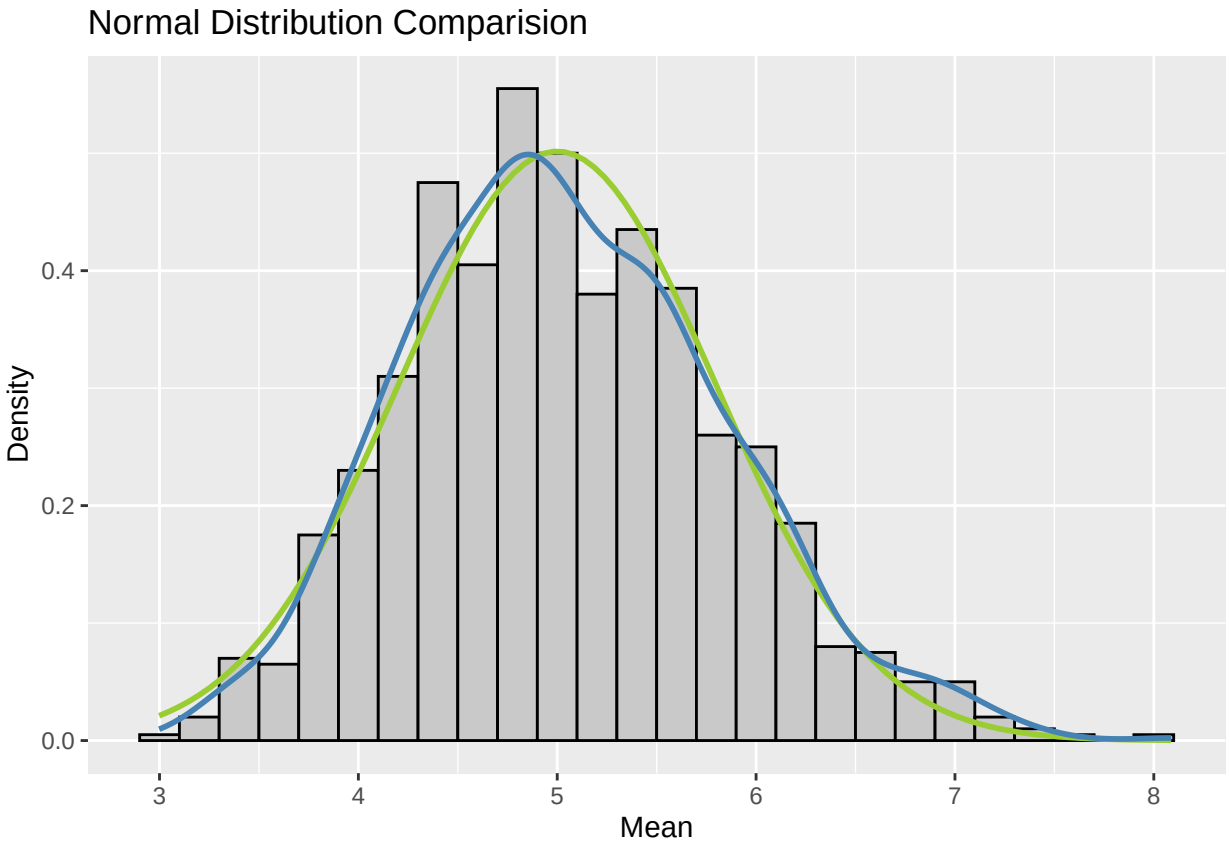
```
theoreticalVariance <- ((1/lambda)*(1/sqrt(n)))^2
theoreticalVariance
```

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

Question 3: Show that the distribution is approximately normal Histogram with Density and sample means:

```
meansDataFrame <- data.frame(meanSimulationData)
tMeansDataFrame <- data.frame(theoreticalMean)
g <- ggplot(meansDataFrame, aes(x = meanSimulationData)) +

 geom_histogram(binwidth = .2, color="black", fill="grey79" , aes(y=..density..))+
            stat_function(fun=dnorm, args=list(mean=theoreticalMean, sd=sd(meanSimulationData)),
                          color="yellowgreen", size =1) +
            stat_density(geom = "line", color = "steelblue", size =1)  +
            labs(x="Mean", y= "Density",
                    title="Normal Distribution Comparision")
g
```



The above plot indicated that density curve is similar to normal distribution curve.

Q-Q Normal Plot also indicates the normal distribution

```r
qqnorm(meanSimulationData)
 qqline(meanSimulationData, col = "steelblue",lwd=2)
```