# Statistical Inference - Project - Simulation Exercise

Author: Wilson Mok

## Overview

In the simulation below, I will investigate the exponential distribution in R and compare it the Central Limit Theorem. We will compare the simulated results with the theorietical results for the following 3 questions: 1. Show the sample mean and compare it to the theoretical mean of the distribution. 2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. 3. Show that the distribution is approximately normal.

## Simulation

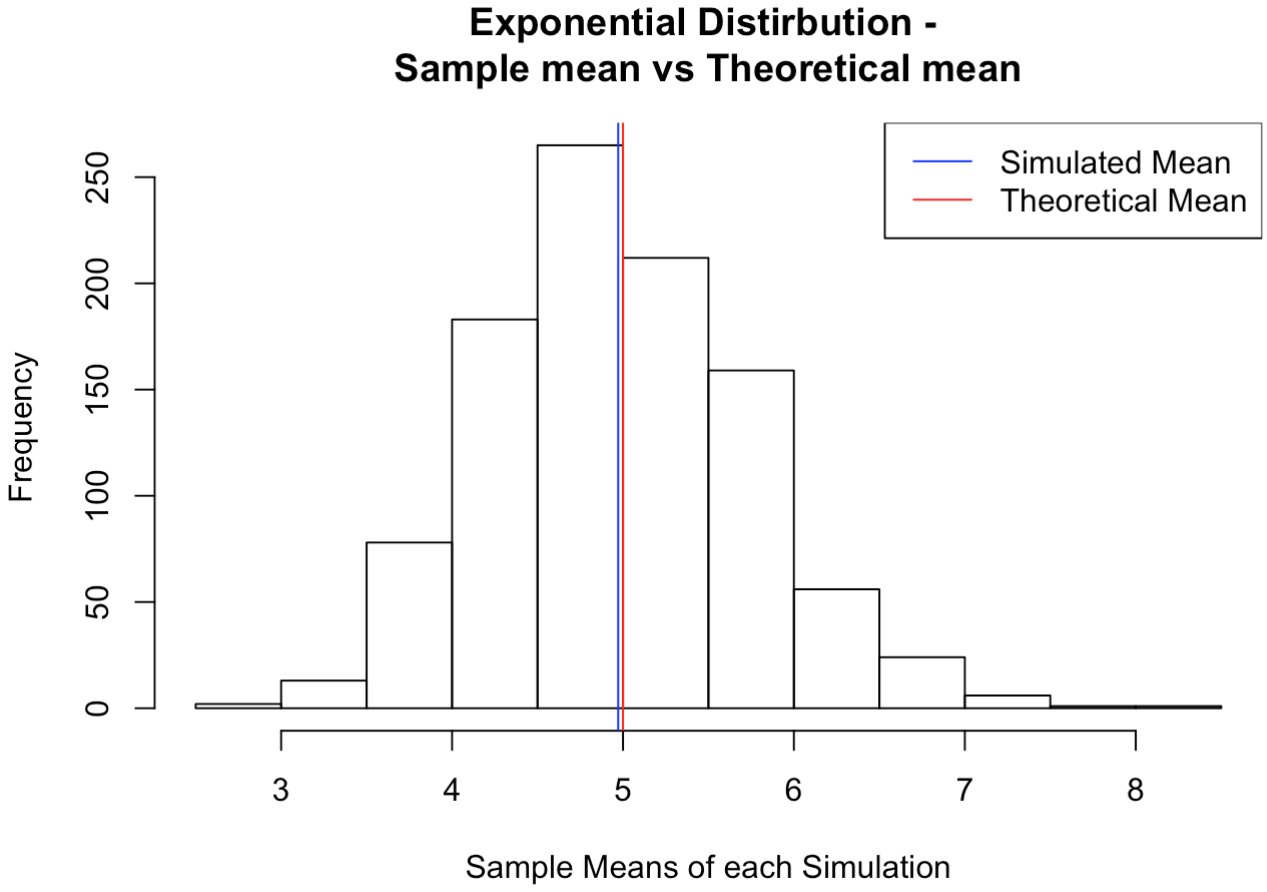
The exponential distribution can be simulated in R using rexp(n,lambda) where lambda is the rate parameter. In this experiement we will be set lambda = 0.2 for all the simulations.

```
set.seed(12345)

numOfSim <- 1000
lambda <- 0.2
n <- 40

mns = NULL
for (i in 1:numOfSim)
    mns = c(mns,mean(rexp(n,lambda)))
```

## Sample Mean versus Theoretical Mean

The theoretical mean is 1/lambda. We will calcuate and comapre the theoretical and simulated mean.

```
theoMean <- 1/lambda
simMean <- mean(mns)

hist(mns,
     main="Exponential Distirbution - \nSample mean vs Theoretical mean",
     xlab="Sample Means of each Simulation",
     ylab="Frequency"
)
abline(v=simMean, col='blue', lwd=1)
abline(v=theoMean, col='red', lwd=1)
legend('topright', c("Simulated Mean", "Theoretical Mean"), col=c("blue","red"),
       lty=c(1,1))
```




```
theoMean
```

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

```
simMean
```

```
## [1] 4.971972
```

Conclusion: The simulated mean is very similar to the theoetical mean.

# Sample Variance versus Theoretical Variance

The theoretical variance is (Std)^2/n. The theoretical standard deviation is 1/lambda. We will calcuate and comapre the theoretical and simulated variances.

```
theoVar <- (1/lambda)^2/n
simVar <- var(mns)

theoVar
```

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

```
simVar
```

```
## [1] 0.5954369
```

Conclusion: The simulated variance is very similar to the theoetical variance.

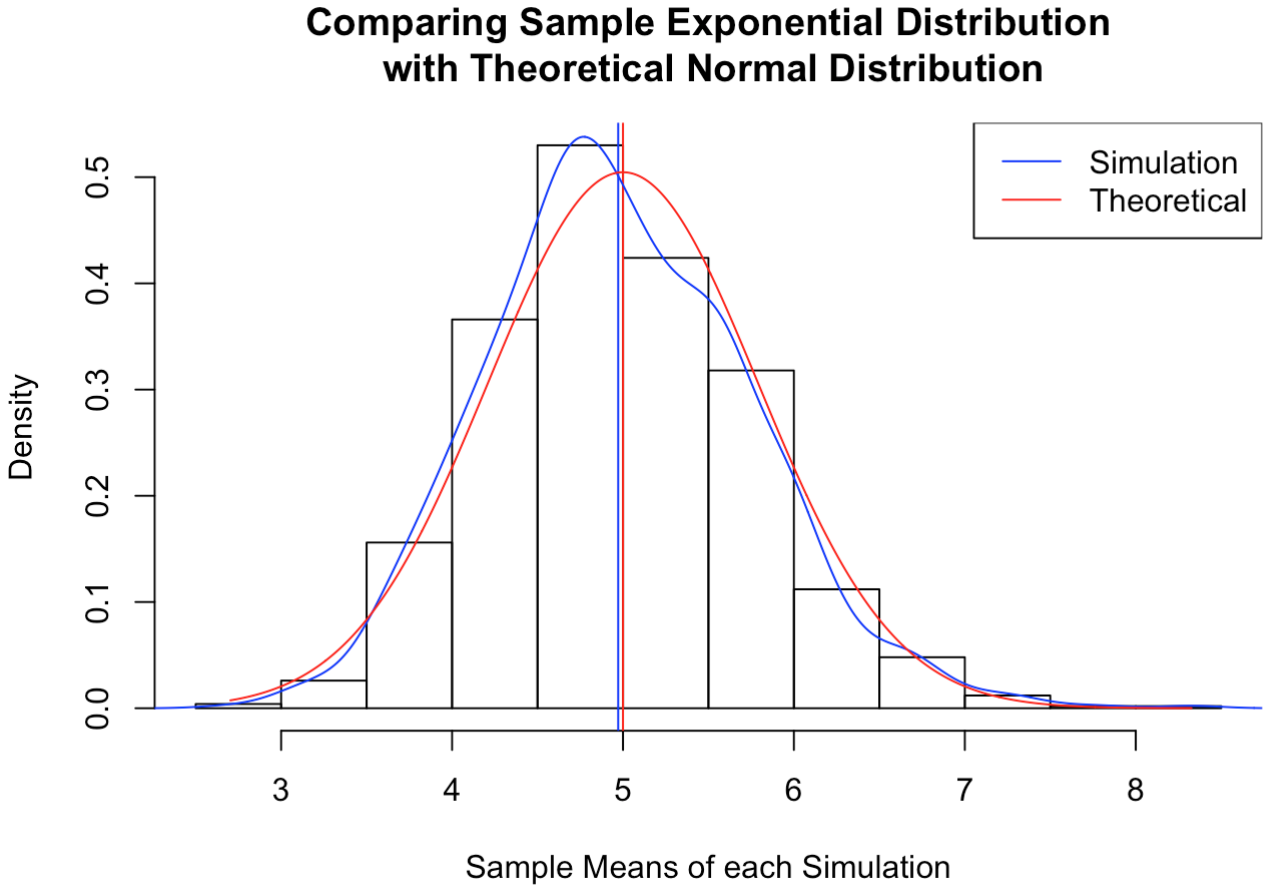
# Show simulated data is approximately normal

In order investage if the simulated data is approximately normal, we have to generate sample normal data with the same mean and standard distribution as the simulated exponential data. We can do this by using r function dnorm.

```
theoSd <- 1/lambda/sqrt(n)
theoSd
```

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

```
normPoints <- seq(min(mns), max(mns), length=numOfSim)
normData <- dnorm(normPoints, mean=theoMean, sd=theoSd)

hist(mns, prob=T,
     main="Comparing Sample Exponential Distribution\n with Theoretical Normal Distribut
ion",
     xlab="Sample Means of each Simulation",
     ylab="Density")
abline(v=simMean, col='blue', lwd=1)
abline(v=theoMean, col='red', lwd=1)

lines(density(mns), col="blue", lwd=1)
lines(normPoints, normData, pch=22, col="red", lwd=1)
legend('topright', c("Simulation", "Theoretical"), col=c("blue","red"),
       lty=c(1,1))
```




Conclusion: The simulated curve is very similar to the theoretical standard normal curve. Given this information, I believe the simulated data is approximately normal.