Title: Exploration of the Exponential Distribution by Allen Seol 7/21/15
(Coursera Statistical Inference Class for the Data Science Specialization)

Overview:
For this project, we will investigate the exponential distribution in R and compare it with the Central Limit Theorem (CLT).

Simulations:
We will use rexp(n,lambda) to simulate 1000 simulations with lambda set at 0.2 for all simulations. We will investigate the averages of 40 exponentials.

```
#Parameters and setting seed
set.seed(1337)
lambda = .2
n = 40
sims = NULL ##Matrix with simulations
avg = NULL ##vector with Averages

#1000 Simluations of 40 simulations of exponential distribution
for(i in 1:1000) {
  sims <- rbind(sims, rexp(n,.2))
}
```

Sample Mean Versus Theoretical Mean:
1. Show the sample mean and compare it to the theoretical mean of the distribution.

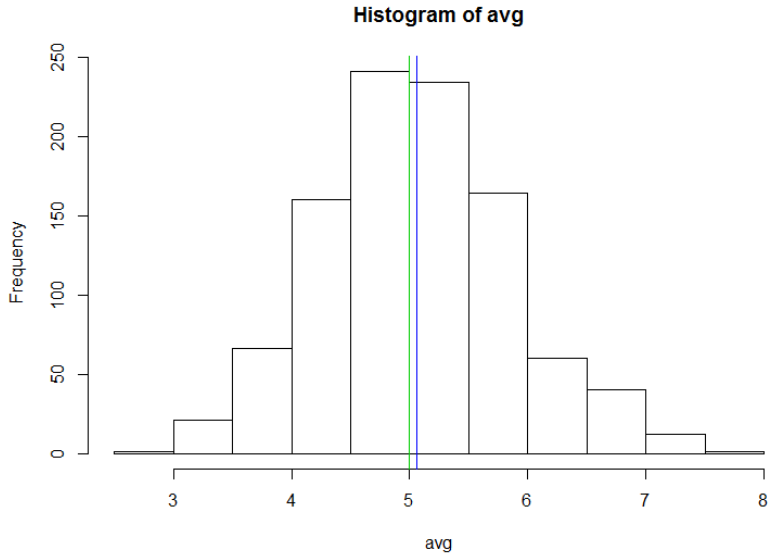
The theoretical mean of an expotential distribution: Mu = 1 /lambda ; Mu = 1 / 0.2 = 5

```
#Run Apply Function to get mean of each simulations of 40 samples
avg = apply(sims,1,mean)

#mean of the 1000 simulations
mean_avg <- mean(avg)
print("Mean of 1000 simulations:")
print(mean_avg)
print("Theoretical mean is 5")
```

Output:
```
[1] "Mean of 1000 simluations:"
[1] 5.055995
[1] "Theoretical mean is 5"
```

When we review the histogram of the means, we can see that the mean of 40 samples from our 1000 simulations closely approximates 5, which is the theoretical mean. The blue line is the simulation mean and the green line is the theoretical mean.



Sample Variance Versus Theoretical Variance:
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

The theoretical variances of an exponential distribution is expected to be 1/lambda^2. Hence with a lambda of 0.2, we should expect 1 / 0.2^2 or 25 for variance. We can see below that the variance is close to 25 in our simulation.

```
#variance of the 1000 simulations
var_avg <- mean(apply(sims,1,var))
print("Variance of 1000 simulations:")
print(var_avg)
print("Theoretical Variance is 25")
```

Output:

```
[1] "Variance of 1000 simulations:"
[1] 25.84083
[1] "Theoretical Variance is 25"
```

Distribution:
3. Show that the distribution is approximately normal. We will use a quantile-quantile plot to see if the distribution of means of 40 samples of an exponential distribution has a normal distribution. As we can see below, it is indeed likely normal, given the linear regression of the qqplot. A line going through the 1st and 3rd quantile is provided in red.

```
#Histogram of means of 40 samples x 1000 times
hist(avg)
abline(v=5,col=3) ##Green line Shows Theoretical Mean of 5
abline(v=mean_avg,col=4) ##Blue Line shows Actual Mean Calculated

#qqplot of values
qqnorm(avg)
qqline(avg,col=2)
```