Statistical Inference Course Project Pt.1

Junaid Khan

March 25, 2018

Overview :

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials

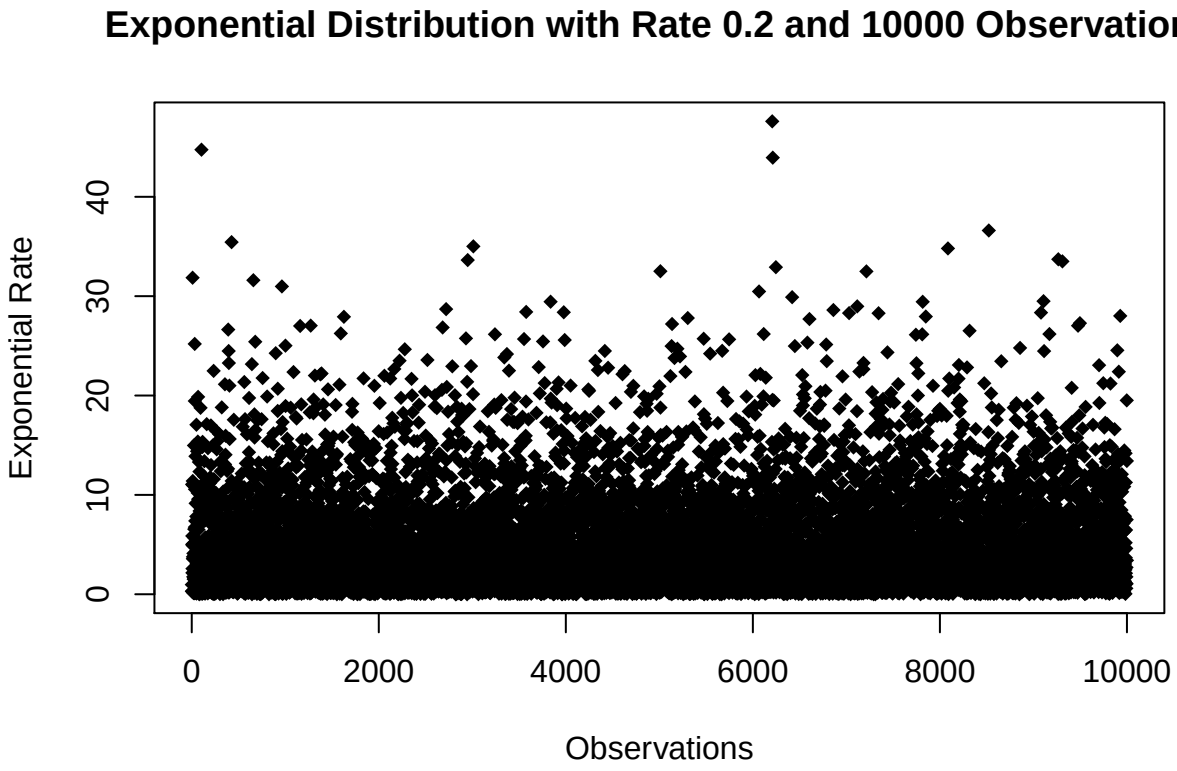
Simulations

We are given the values of sample size, lambda ,no.of simulations

```
set.seed(20)
sim <- 1000
n <- 40
lambda <- 0.2
```

Plotting a histogram of the Exponential Distribution, with lambda = 0.2

```
plot(rexp(10000,lambda),type="p",pch=18,main="Exponential Distribution with Rate 0.2 and 10000 Observati
```



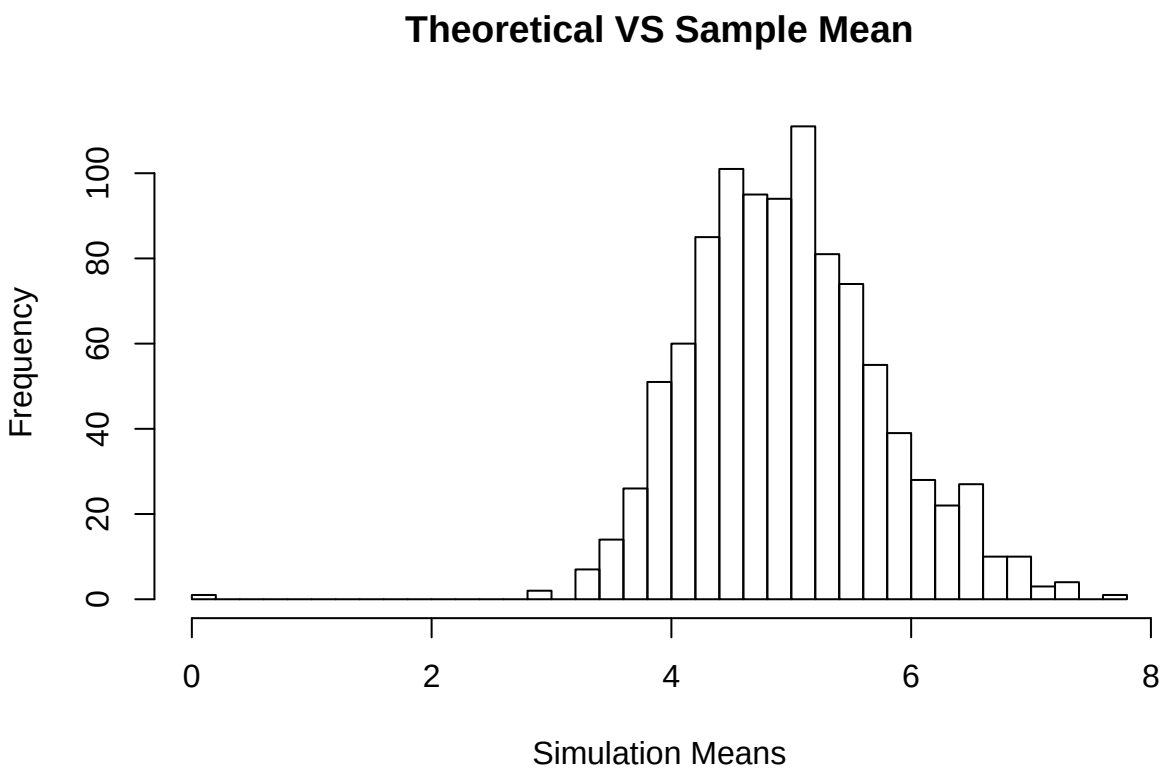
Distribution of Averages of 40 Means :

- Calculating the average of 40 means *

```
means = 0
for(i in 1 : sim)  means <- c(means,mean(rexp(n,lambda)))
```

Plotting the histogram of Distribution of 40 means:

```
hist(means,breaks = 30,xlab="Simulation Means", main="Theoretical VS Sample Mean")
```

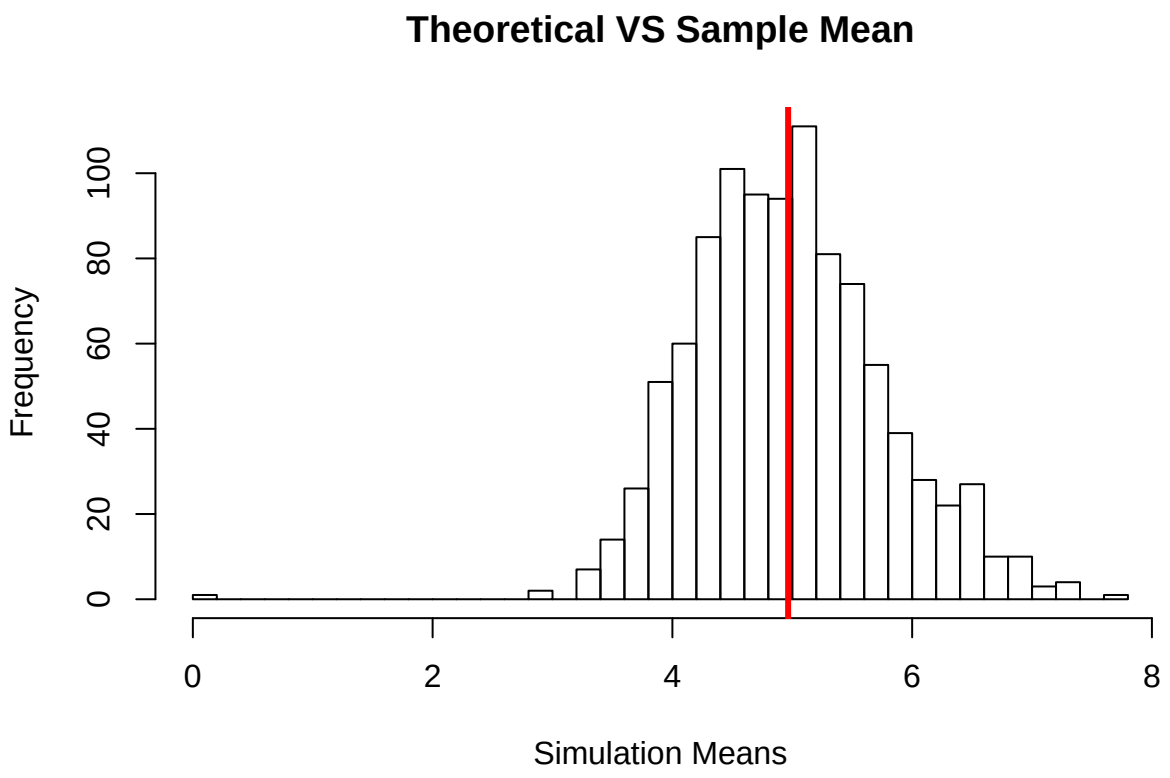


Sample Mean and Theoretical Mean :

By theory, mean = 1/?? , i.e in our case, theoretical mean = 1/0.2 =5

We will calculate the sample mean now:

```
hist(means,breaks = 30, xlab="Simulation Means", main="Theoretical VS Sample Mean")
abline(v=mean(means),lwd=3,col="red")
```




```
mn <- mean(means)
```

Thus the Sample mean is 4.9654483.

Hence the Theoretical Mean is 5 and the Sample Mean is 4.9654483

Sample Variance and Theoretical Variance

By theory, variance = mean/sqrt(n) , i.e

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

Thus, the theoretical variance is 0.7905694

We will find the sample mean :

```
s <- sd(means)
```

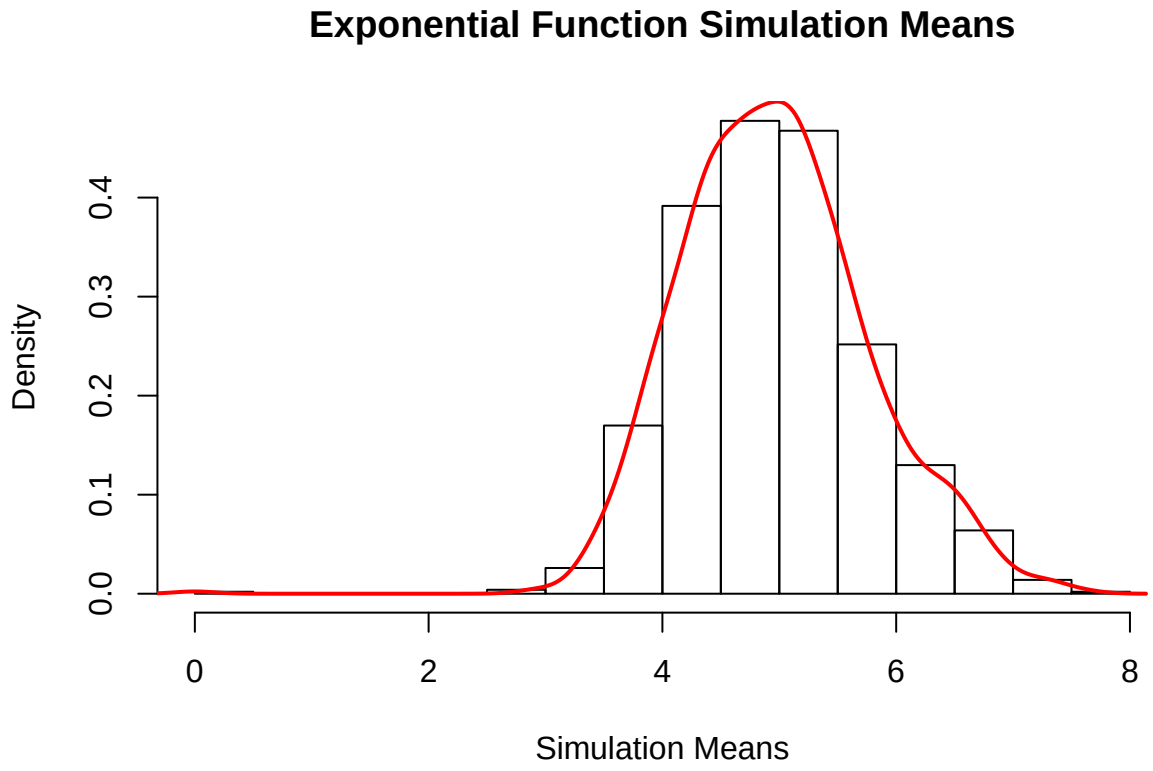
Thus the sample variance is 0.7966459

Hence the Theoretical Variance is 0.7905694 and the Sample variance is 0.7966459

To show that the distribution is normal

we'll investigate whether the exponential distribution is approximately normal. Due to the Central Limit Theorem, the means of the sample simulations should follow a normal distribution

```
hist(means,breaks = 20, prob=T, xlab = "Simulation Means", main = "Exponential Function Simulation Means
lines(density(means),lwd=2,col="red")
```



As the graph shows, the distribution of means of our sampled exponential distributions appear to follow a normal distribution, due to the Central Limit Theorem. If we increased our number of samples (currently 1000), the distribution would be even closer to the standard normal distribution.The dotted line above is a normal distribution curve and we can see that it is very close to our sampled curve, which is the red line above.