SI_Project_P1

Rahul Jain

30/01/2022

Overview

The main purpose of this project is to investigate the exponential distribution and then compare it to the CLT. For this analysis we will use lambda as 0.2 and simulate the distribution of averages of 40 exponentials over 1000 simulations.

Library Import

```
library(ggplot2)
```

Simulations

Simulate the data from exponential distribution

```
set.seed(42)
lambda = 0.2
nosim = 1000
count = 40

expData <- matrix(sample(rexp(nosim*count,lambda)),nosim,count)
expMns <- apply(expData,1,mean)
```

Sample Mean Comparison with theoretical Mean

```
paste("Mean of sample means from simulation = ",mean(expMns))
```

Sample Mean Calculation

```
## [1] "Mean of sample means from simulation =  4.98650831745453"
```

```
paste("Theoretical Mean = ",1/lambda)
```

Theoretical Mean Calculation

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

Comparison There is only a small difference between theoretical mean and sample mean

```
1/lambda - mean(expMns)
```

```
## [1] 0.01349168
```

Sample Variance Comparison with theoretical variance

```
paste("Variance of Sample means from simulation = ",var(expMns))
```

Sample Variance Calculation

```
## [1] "Variance of Sample means from simulation =  0.666698060983617"
```

```
paste("Theoretical Variance = ",1/((lambda*sqrt(count))^2))
```

Theoretical Variance Calculation

```
## [1] "Theoretical Variance =  0.625"
```

Comparison There is very little difference between theoretical variance and variance of simulated means

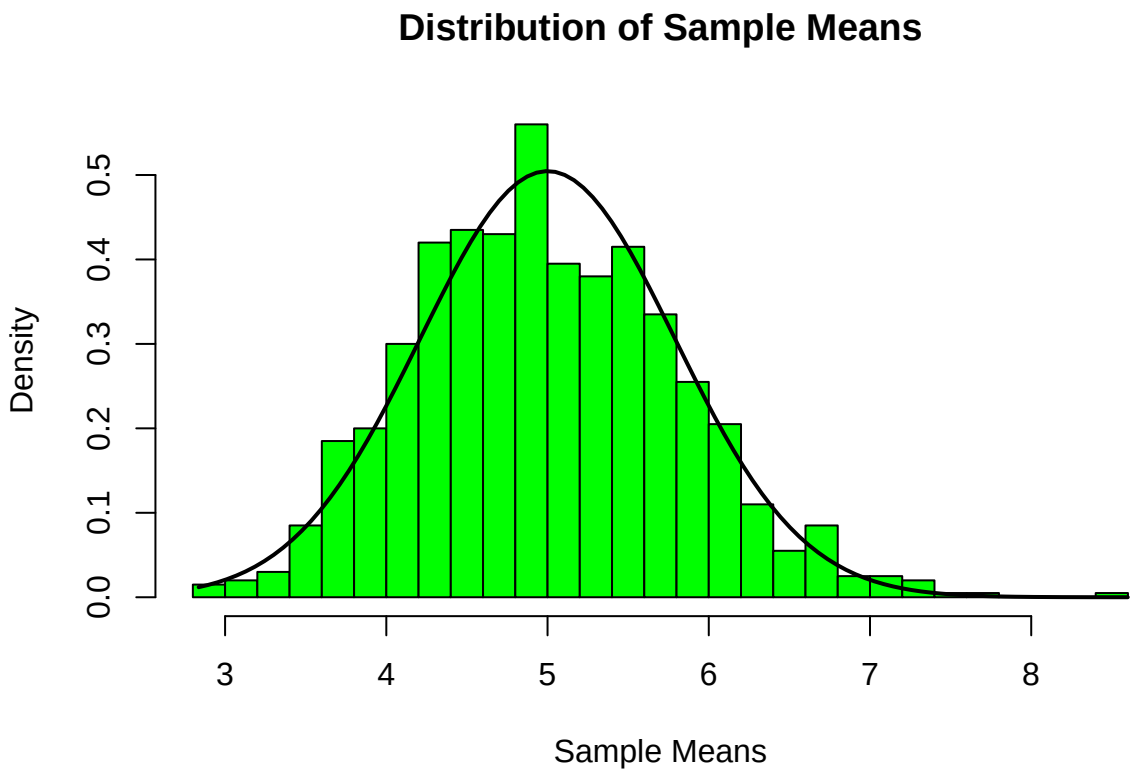
```
1/((lambda*sqrt(count))^2) - var(expMns)
```

```
## [1] -0.04169806
```

Distribution

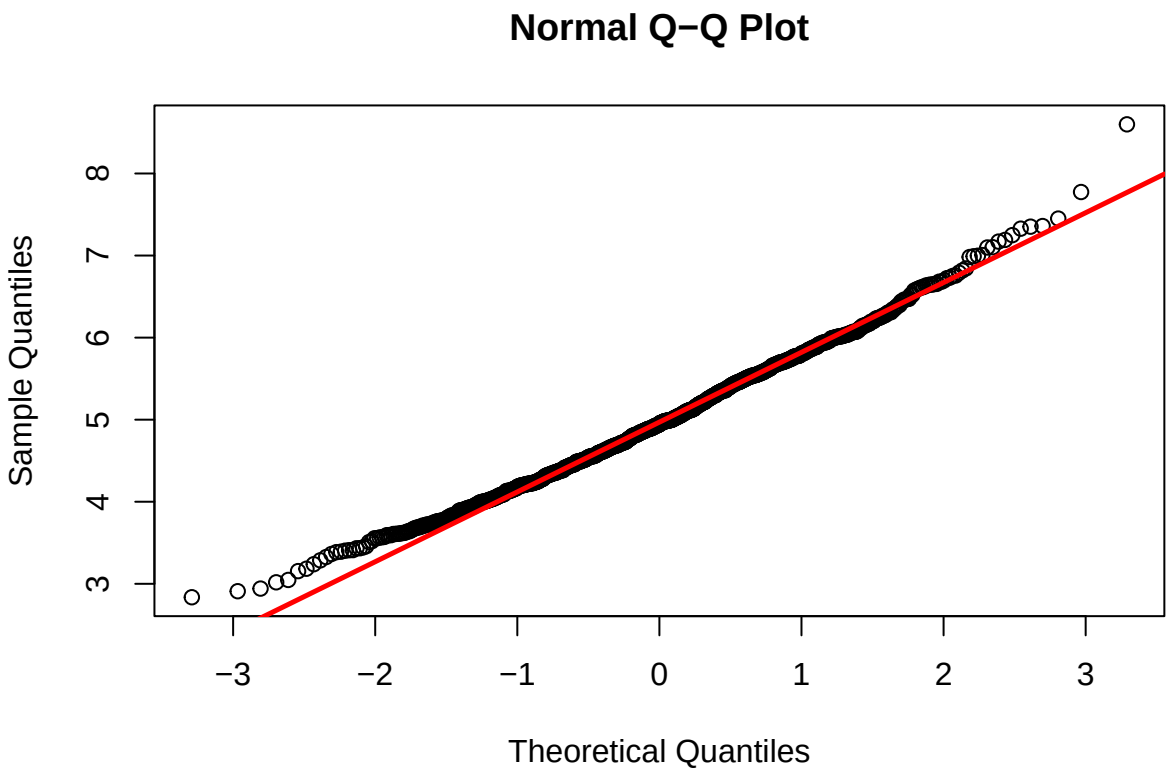
Distribution of Sample Means from exponential distribution

```
xfit <- seq(min(expMns), max(expMns), length=100)
yfit <- dnorm(xfit, mean=1/lambda, sd=(1/lambda/sqrt(count)))
hist(expMns,breaks=count,prob=T,col = "green",xlab = "Sample Means",main = "Distribution of Sample Means
lines(xfit,yfit,col="black",lwd=2)
```



Comparison with normal distribution Quantile plots show that the distribution of Sample Means from exponential distribution is approximately normal

```
qqnorm(expMns)
qqline(expMns,col="red",lwd=2.5)
```


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
3
4
5
6
7
8
−3
−2
−1
0
1
2
3