Simulation Exercise

Hatem Jasim Hatem

May 13, 2019

```
library(ggplot2)
```

Part 1: Simulation Exercise Instructions

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. Note that you will need to do a thousand simulations.

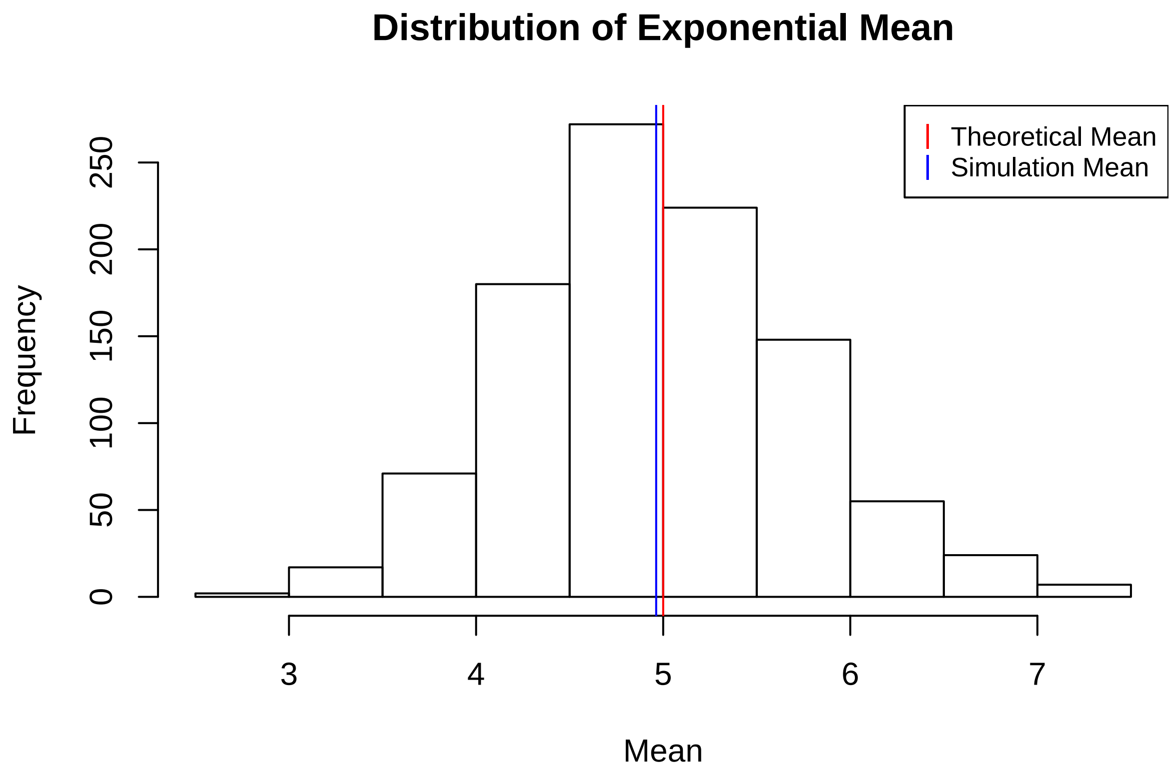
1. Show the sample mean and compare it to the theoretical mean of the distribution.

```
n <- 40
Lambda <- 0.2
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(n, Lambda)))
SampleMean<- mean(mns)
TheoreticalMean <- 1/Lambda
paste('Sample Mean', SampleMean, 'Theoretical Mean', TheoreticalMean, sep = ': ')
```

```
## [1] "Sample Mean: 4.9624590175826: Theoretical Mean: 5"
```

Sample mean 4.962459 approximately equal to theoretical mean 5

```
hist(mns, xlab = "Mean",
     main="Distribution of Exponential Mean")
abline(v = TheoreticalMean, col="red")
abline(v = SampleMean, col="blue")
legend("topright",c("Theoretical Mean", "Simulation Mean"),
       cex=.8,col=c("red","blue"), pch=c('|','|'))
```



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

```
# sample deviation & variance form sample
sample_std <- sd(mns)
sample_variance <- sample_std^2
paste('sample stander deviation', sample_std, 'sample variance',
      sample_variance, sep = ': ')
```

```
## [1] "sample stander deviation: 0.749321922712387: sample variance: 0.561483343857388"
```

Sample stander deviation can calculated theoretically according to this formula: $S = \frac{\sigma}{\sqrt{n}}$ and sample variance accourding to this formula $variance = S^2$

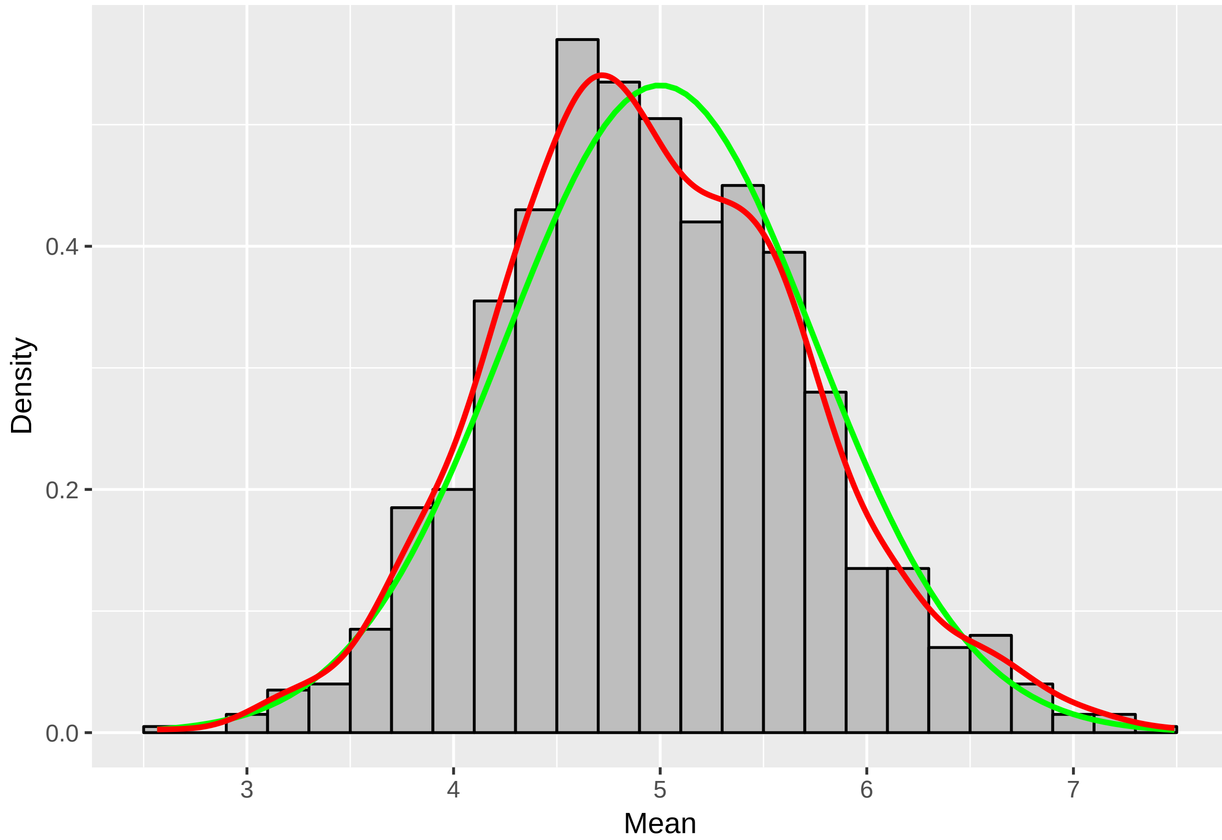
```
# theoretical deviation & variance form pupluation parapeter
theoretical_std <- (1/Lambda)/sqrt(n)
theoretical_variance  <- theoretical_std^2
paste('Theoretical stander deviation', theoretical_std,
      'Theoretical variance', theoretical_variance, sep = ': ')
```

```
## [1] "Theoretical stander deviation: 0.790569415042095: Theoretical variance: 0.625"
```

The simulation variance 0.5614833 approximately equal to Theoretical variance 0.625

3. Show that the distribution is approximately normal.

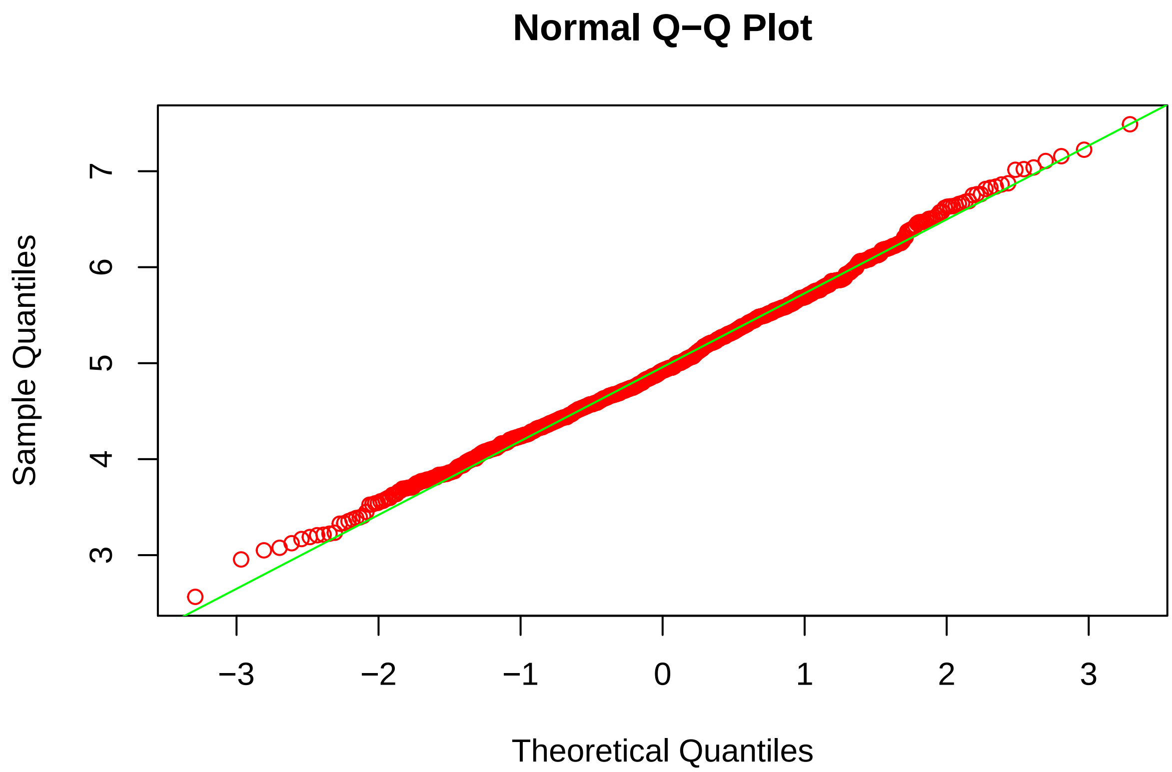
```
ds <- data.frame(mns)
ggplot(ds, aes(x = mns)) +
 geom_histogram(binwidth = .2, color="black", fill="gray" , aes(y=..density..))+
           stat_function(fun=dnorm, args=list(mean=TheoreticalMean, sd=sd(mns)),
                        color="green", size =1) +
           stat_density(geom = "line", color = "red", size =1)  +
           labs(x="Mean", y= "Density")
```



The above figure show that sample distribution curve is approximately similar to normal distribution curve.

For further explaint used Q-Q Normal Plot to compare between sample quantiles and theortical quantiles

```
qqnorm(mns, col = 'red')
qqline(mns, col = "green")
```


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