Exponential Ditribution Simulation

Shang Ju

January 13, 2017

Overview

In this project we will investigate the exponential distribution in R and compare it with the Central Limit Theorem. We will simulate the exponential distribution using R with rexp(n, lambda) where lambda is the rate parameter. We set lambda = 0.2 for all of the simulations, and will investigate the distribution of averages of 40 exponentials, and do a thousand simulations.

Simulations

Simulate 40 samples from the exponential distribution with `lambda = 0.2`. Calculate the mean. Repeat the simulation for 1000 times and record the 1000 means in a vector **mns**.

```
# set seed
set.seed(123)

# prepare vector of means
mns <- NULL

# simulation
for (i in 1:1000) mns <- c(mns,mean(rexp(40, 0.2)))
```

Sample Mean versus Theoretical Mean

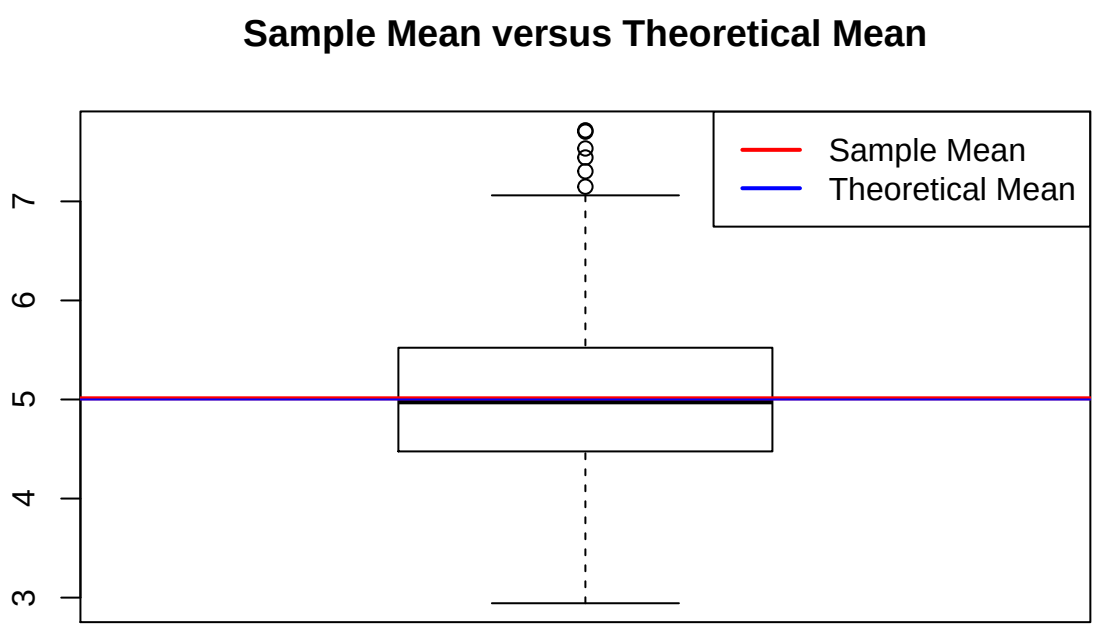
```
# Sample mean of the 1000 means
mean(mns)
```

```
## [1] 5.011911
```

```
# Theoretical mean of the distribution: lambda^(-1)
1/0.2
```

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

```
# figure
boxplot(mns)
abline(h=mean(mns),col='red',lwd=2)
abline(h=1/0.2,col='blue',lwd=1)
title(main = 'Sample Mean versus Theoretical Mean')
legend('topright',legend=c('Sample Mean','Theoretical Mean'),lty=c(1,1),col=c('red','blue'),lwd=2)
```



Sample mean 5.012 is very close to the theoretical mean 5.

Sample Variance versus Theoretical Variance

```
# Sample variance
var(mns)
```

```
## [1] 0.6004928
```

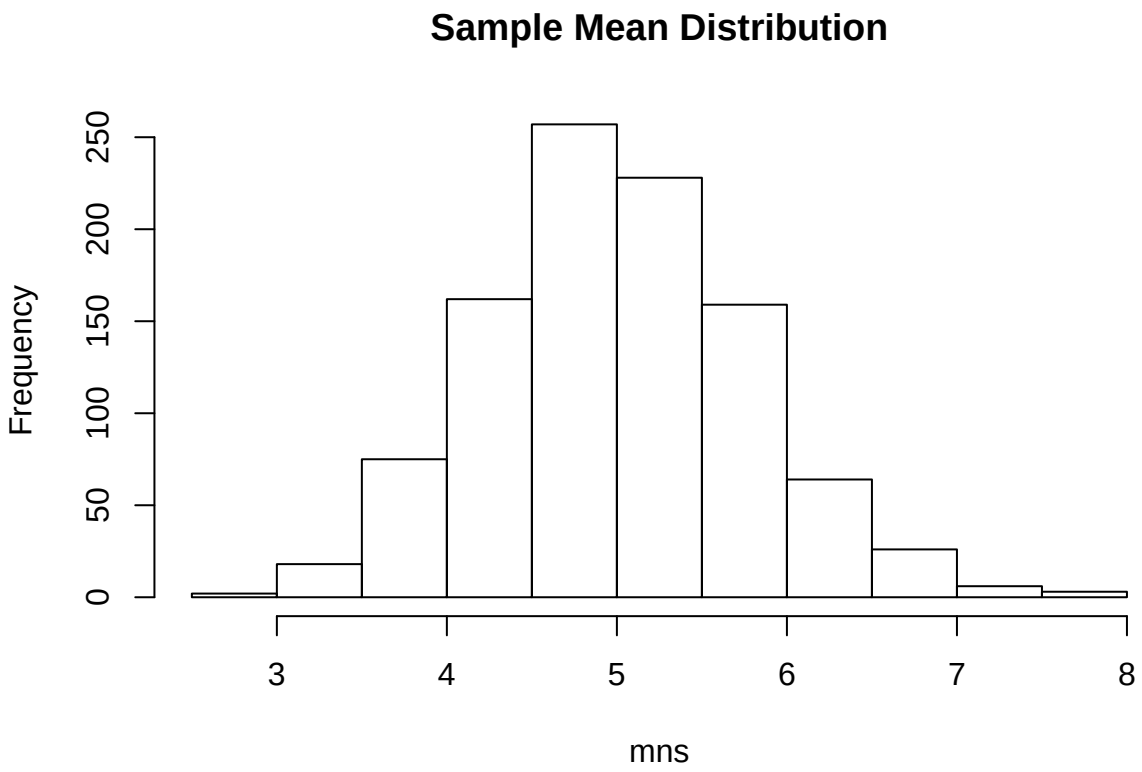
```
# Theoretical variance of the distribution: lambda^(-2)/n
(1/0.2)^2/40
```

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

The sample variance 0.600 is smaller than the theoretical variance of 0.625. This is likely due to noise which could be alleviated by increasing simulation times.

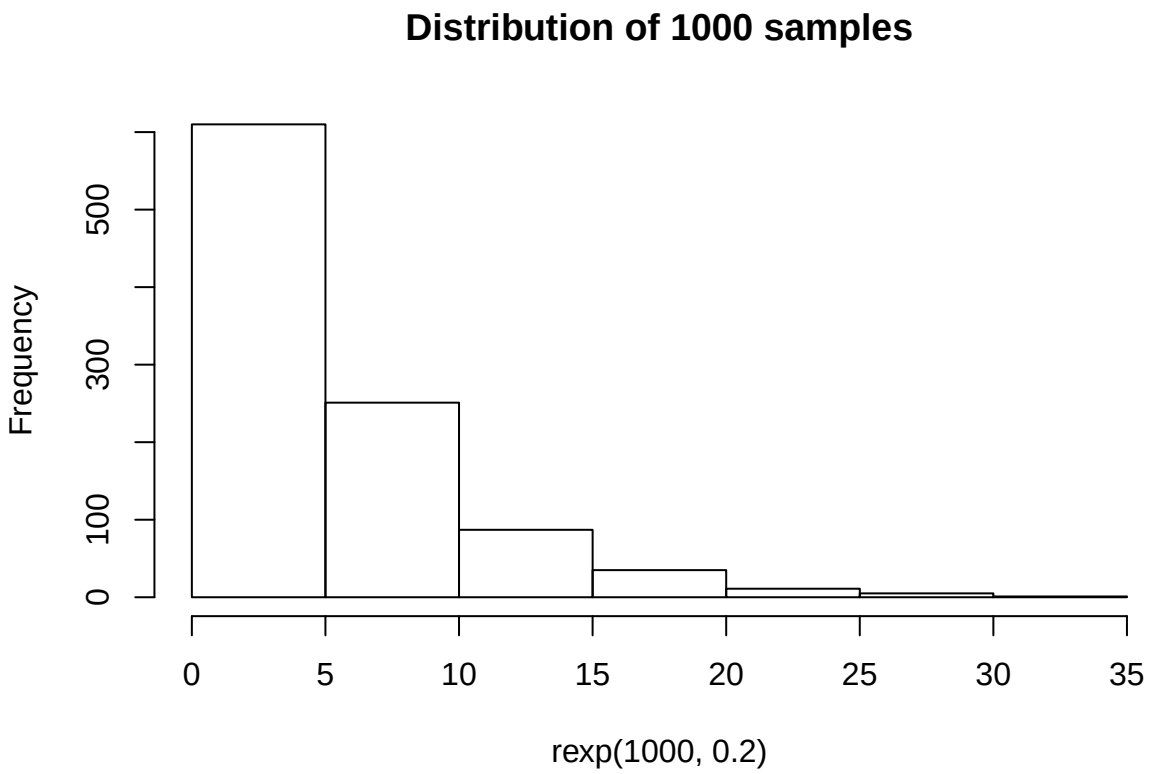
Distribution

```
# The distribution of the sample means is approximately normal.
hist(mns,main = 'Sample Mean Distribution')
```



The bell-like curve suggests that the distribution of means is approximately normal. On the other hand, the distribution of 1000 samples of the exponential distribution is not normal:

```
# Versus the distribution of the 1000 samples.
hist(rexp(1000, 0.2),main = 'Distribution of 1000 samples')
```


Distribution of 1000 samples
Frequency
0
100
300
500
0
5
10
15
20
25
30
35
rexp(1000, 0.2)