Statistical Inference Course project : Simulation Excercise - Neeraj Ahire

Overview

This project is about simulating the exponential distribution and using it to ascertain the validity of the central limit theorem. So, accordingly we will create a sample means distribution with sample size 40 and 1000 such means and check for the values of mean, variance for this distribution and compare them with the theoretical values. We will also check the shape of the distribution and see if it is Gaussian.

Comparing Mean

So lets simulate sample means distribution with sample size 40 and 1000 iterations. Lambda = 0,2.

```
set.seed(1)
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40, 0.2)))
```

According to Central limit theorem the sample means distribution should approximately center around actual population mean. In our case since Lambda = 0.2, Theoretical mean = 5. Lets the check the mean and median of the sample means distribution.

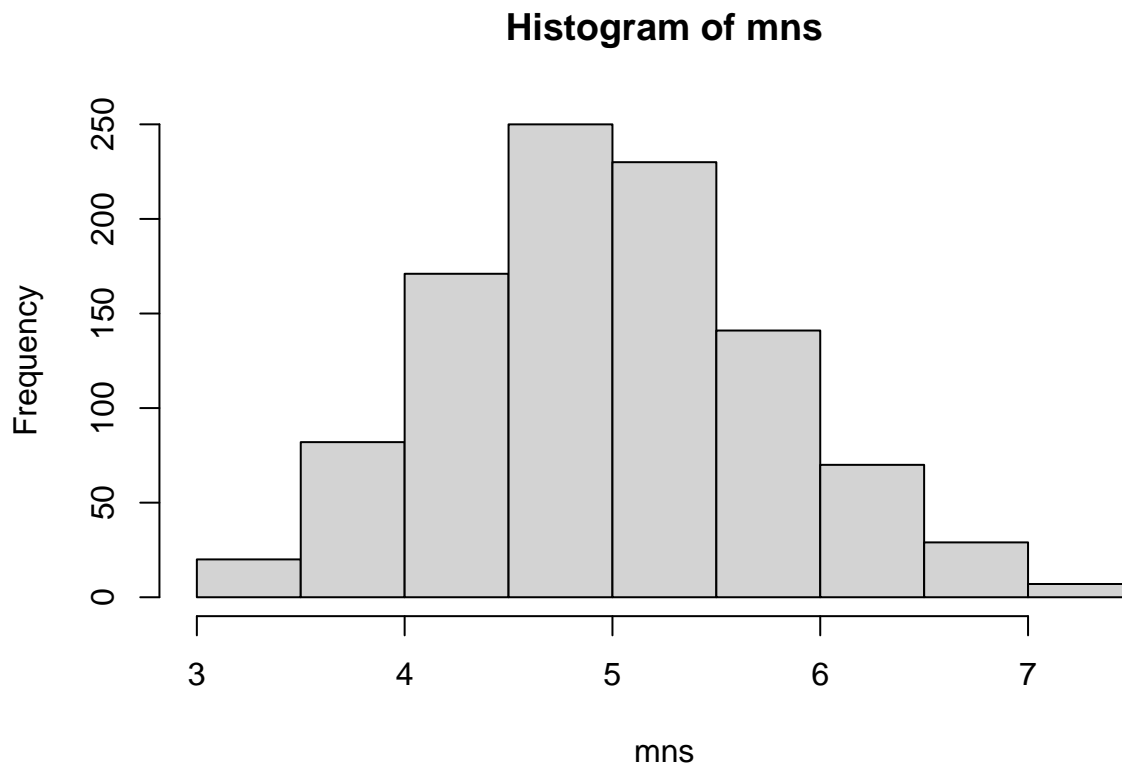```
mean(mns)
```

```
## [1] 4.990025
```

```
median(mns)
```

```
## [1] 4.950169
```

As we can see this value is quite close the the theoretical value of 5.

We can also see the histogram of mns by plotting it as below.

```
hist(mns)
```

Histogram of mns

We can see how the plot is roughly centered around 5 as expected from CLT.

Comparing variance

Now lets compare variance of of our distribution with the theoretical value. From CLT, Theoretical value of variance will be (sigma)^2/n. Now since, Lambda = 0.2, and n = 40, the theoretical variance = 0.625. Lets check the actual value.
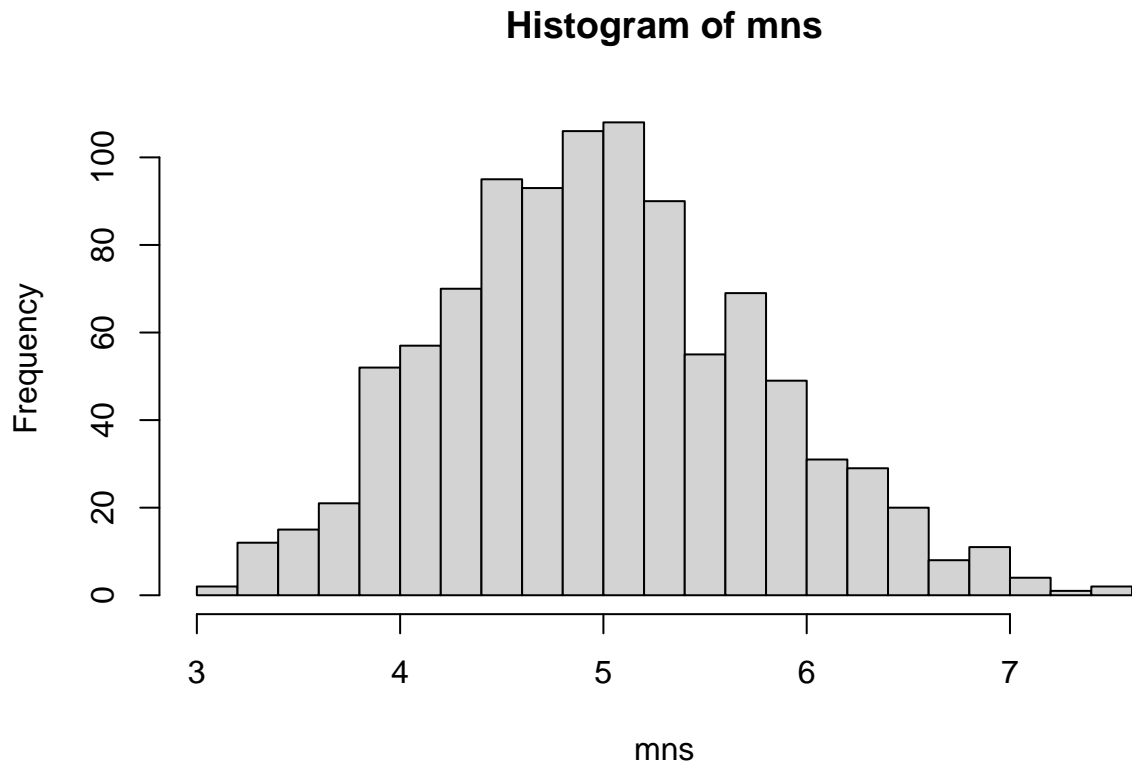
```
var(mns)
```

```
## [1] 0.6111165
```

As we can see this value is quite close to the theoretical as expected from the CLT.

checking the distrubtion shape

Lets now again look at our distribution

```
hist(mns, breaks = 20)
```

Histogram of mns

As we can see the distribution is approximately Gaussian as expected from CLT. We can also compare the it by plotting the actual function distribution and not the means distribution to validate that population distribution can be anything but means distribution for large samples sizes approximates to Gaussian.

```
hist(rexp(1000, 0.2))
```

Histogram of rexp(1000, 0.2)
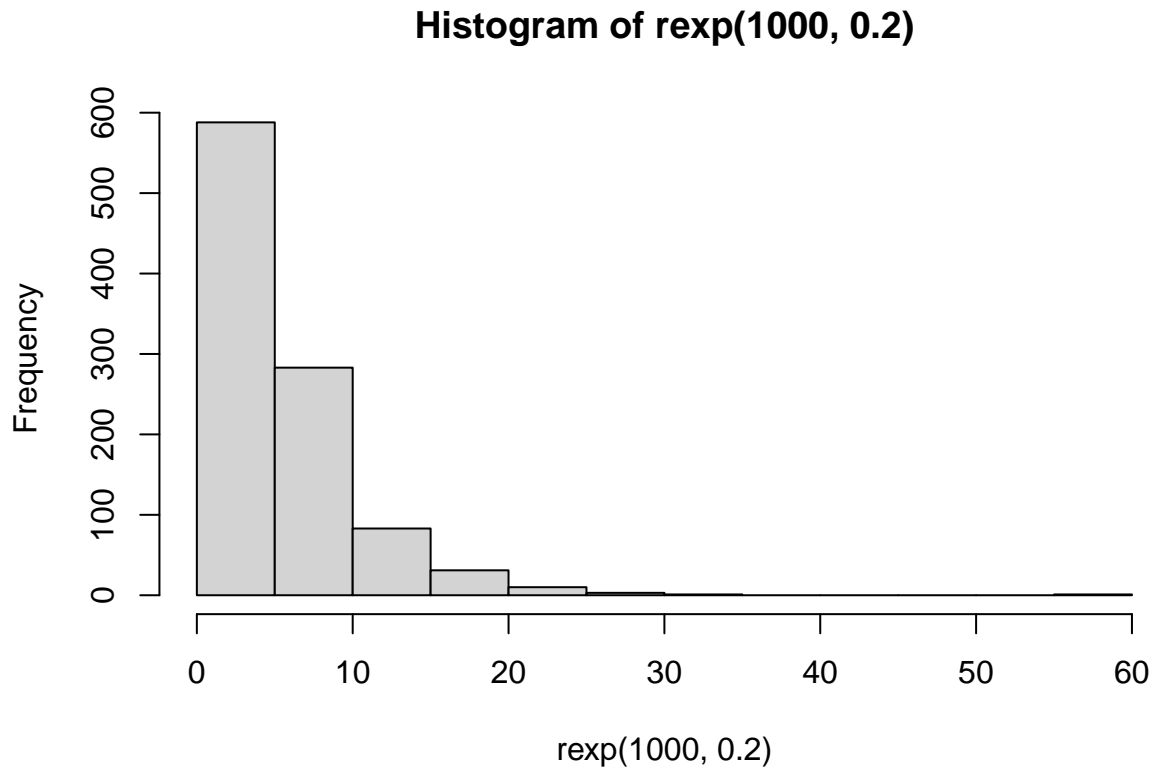
As we can see, actual distribution by simulating 1000 randoms from the exponential distribution is not normal but the sample means distribution shown earlier is approximately Gaussian.

Conclusion

In this assignment we successfully tested and verified the Central Limit theorem by using exponential distribution and comparing theoretical and actual values of mean and variance and checking if distribution is Gaussian.