CLT vs ExpDist

Cy

11/1/2020

Part 1

Overview

This report studies the exponential distribution in R programming. 40 sets of 1000 exponential random numbers are generated and their mean is compared to the theoretical mean.

Random Number Generator

40000 random exponential numbers are generated and arranged in 1000x40 matrix.

```
set.seed(222)
set<-1000
n<-40
lmd<-0.2
simulation<-replicate(set,rexp(n,rate = 0.2))
average<-apply(simulation,2,mean)
```

Comparison of Exponential Distrution with Central Limit Theorem

Mean

Compute the center of simulated exponential distribution by looking for its *mean* and compare it to theoretical mean

$$TheoreticalMean = \frac{1}{\lambda}$$

```
simulated_mean<-mean(average)
theoretical_mean<-1/lmd
```

The mean of simulated mean is 4.9792395,theoretical mean is 5.

```
hist(average,xlab = "sample data",main = "Exponential Distribution")
abline(v= simulated_mean,col='red')
abline(v= theoretical_mean,col='blue')
```

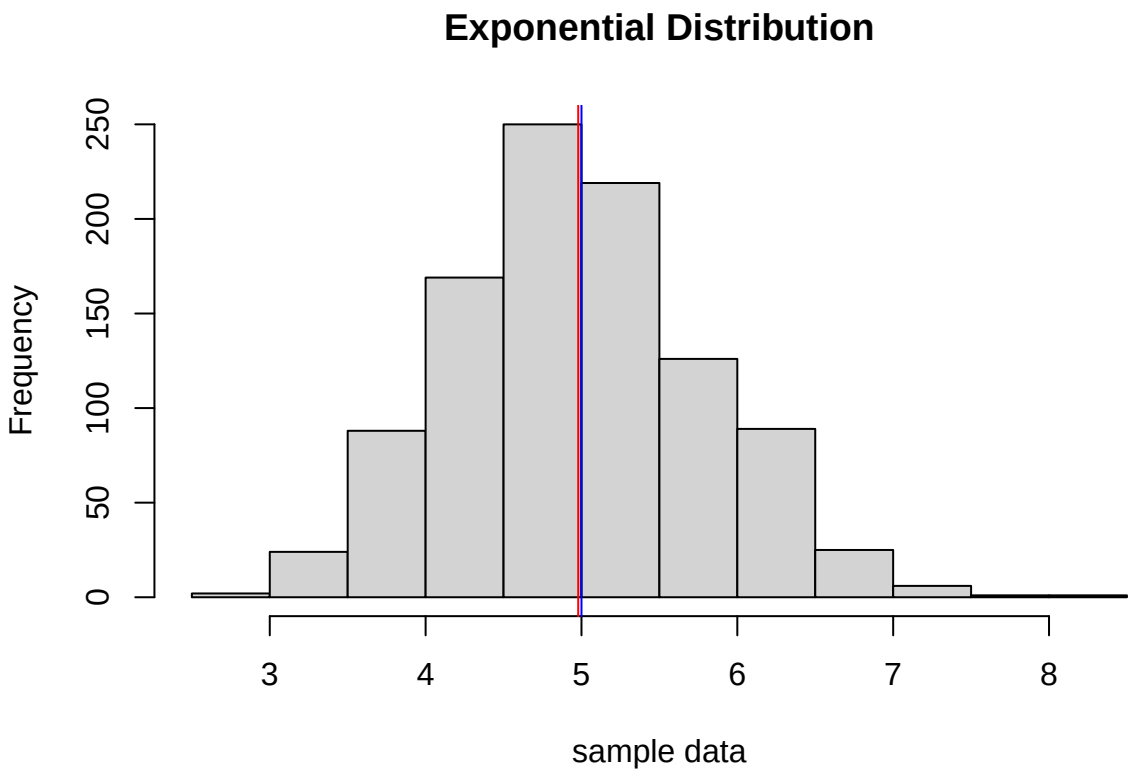

Figure above show the visualization in histogram form, the simulated mean is in red and theoretical mean is in blue. Therefore, it can be concluded the mean of simulation exponential function has almost similar mean with normal distribution

---

Variance

Compute the variance of simulated exponential distribution by looking for its *standard deviation* then squared and compare it to theoretical variance

$$TheoreticalVariance = (\frac{\frac{1}{\lambda}}{\sqrt{n}})^2$$

```
simusd<-sd(average)
theosd<-1/lmd/sqrt(n)
simuvar<-simusd^2
theovar<-theosd^2
```

The variance of simulated samples is 0.6542521 while theoretical variance is 0.625. Both values are close.

Comparison of curves

The simulated data is plotted and compare to the normal distribution curve.

```r
xnormal<-seq(min(average),max(average),length= 100) # get scale of x-axis
ynormal<-dnorm(xnormal,mean=1/lmd, sd=(1/lmd/sqrt(n))) # find density using boundaries

hist(average,breaks=n,prob=T,col="green",xlab = "Means",main="Density",ylab="Density")
lines(xnormal, ynormal, col="red", lty=5)
```

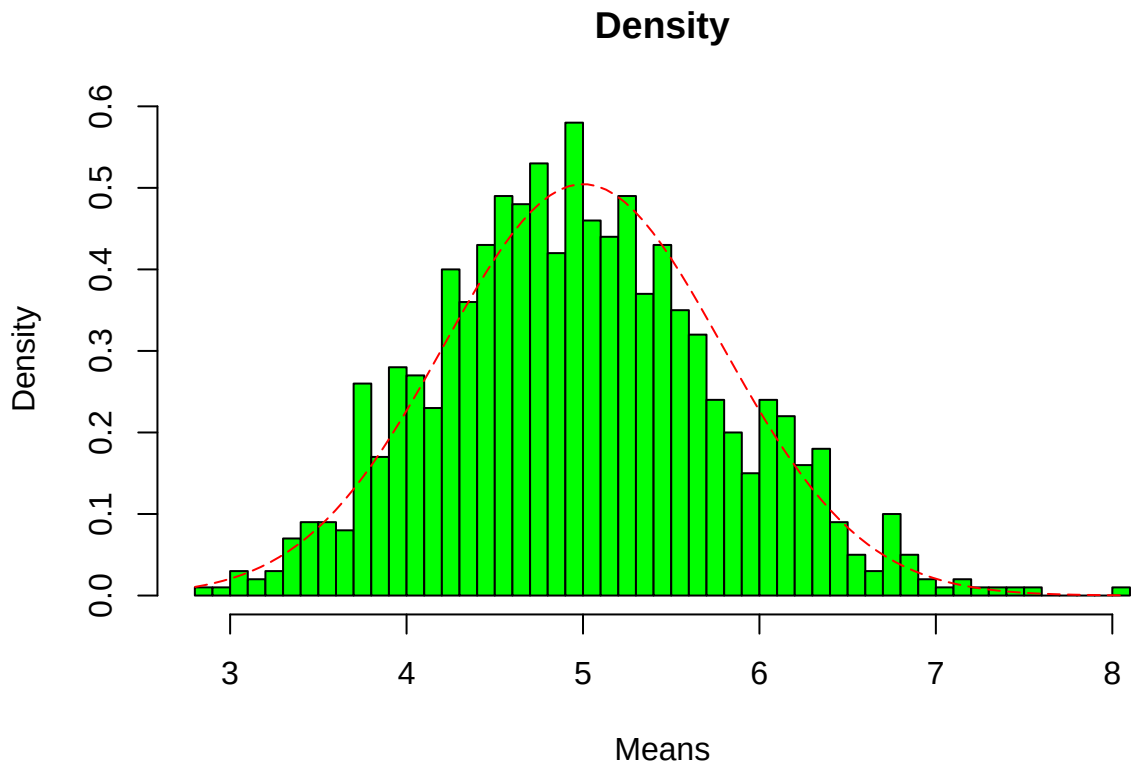

Figure above shows that the Exponential Dist is close to normal Dist funtion curve

```r
qqnorm(average,main= "Normal Q-Q Plot")
qqline(average,col='red')
```

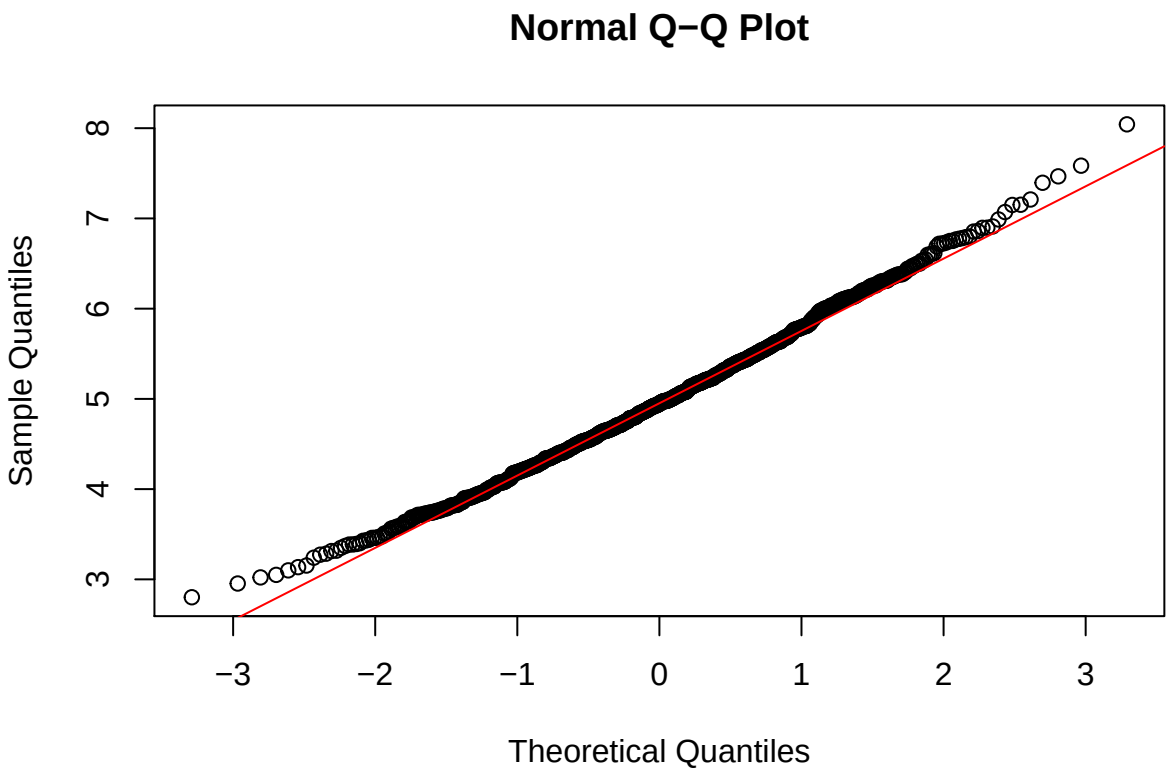

Figure above shows the deviation of Exponential Dist to Normal Dist

Conclusion

The simulation data has very similar trend compared to Gaussian Distribution by looking at their means, variance, sample distribution curve, and qqline.