# Statistical Inference Project Part 1 - A Simulation Exercise

P.Y.Fong

12/24/2020

## Overview:

Part 1 of this project investigates the exponential distribution in R and compares it with the Central Limit Theorem.

### Simulation:

The exponential distribution was simulated in R with *rexp(n, lambda)* where lambda is the rate parameter and $n$ is the number of observations.
The distribution of the averages of 40 exponentials was investigated across 1000 simulations.
The mean of the exponential distribution is $\frac{1}{\lambda}$ and the standard deviation is also $\frac{1}{\lambda}$.
lambda was set to 0.2 for all of the simulations.

R Script for the simulation:

```
# Initialising simulation control variables and reproducible seed.
numsim <- 1000
sampsize <- 40
lambda <- 0.2
set.seed(100)

# Store exp dist sim results in matrix with 1000 rows (simulations) and 40 columns (samples).
matsim <- matrix(rexp(n = numsim * sampsize, rate = lambda), numsim, sampsize)

# Calculate the mean of each row of the sim. matrix.
meansim <- rowMeans(matsim)
```

Visualisation of the simulation mean distribution:

```
# Histogram
hist(meansim, breaks=30, prob=T, col='steel blue', xlim=c(2,8), ylim= c(0,.6), main="", xlab="")

# Distribution Averages Outline
lines(density(meansim),col= 'dark blue',lwd= 2)

# Theoretical mean
abline(v=1/lambda, col="red", lwd= 1)
abline(v=mean(meansim), col='black', lwd= 1)
text(x=5,y=0.6, " Theoretical Mean = 5 ", col = "red", adj = c(0, .5))
text(x=5,y=0.6, paste("Sample Mean=", round(mean(meansim),2)," "), col = "black", adj = c(1, .5))
```

```r
# Theoretical density.
xfit <- seq(min(meansim), max(meansim), length=100)
yfit <- dnorm(xfit, mean=1/lambda, sd=(1/lambda/sqrt(sampsize)))
lines(xfit, yfit, pch=22, col="red", lty=2, lwd= 2)

legend('right', c("Simulation", "Theoretical"), lty=c(1,2), col=c("black", "red"))
```

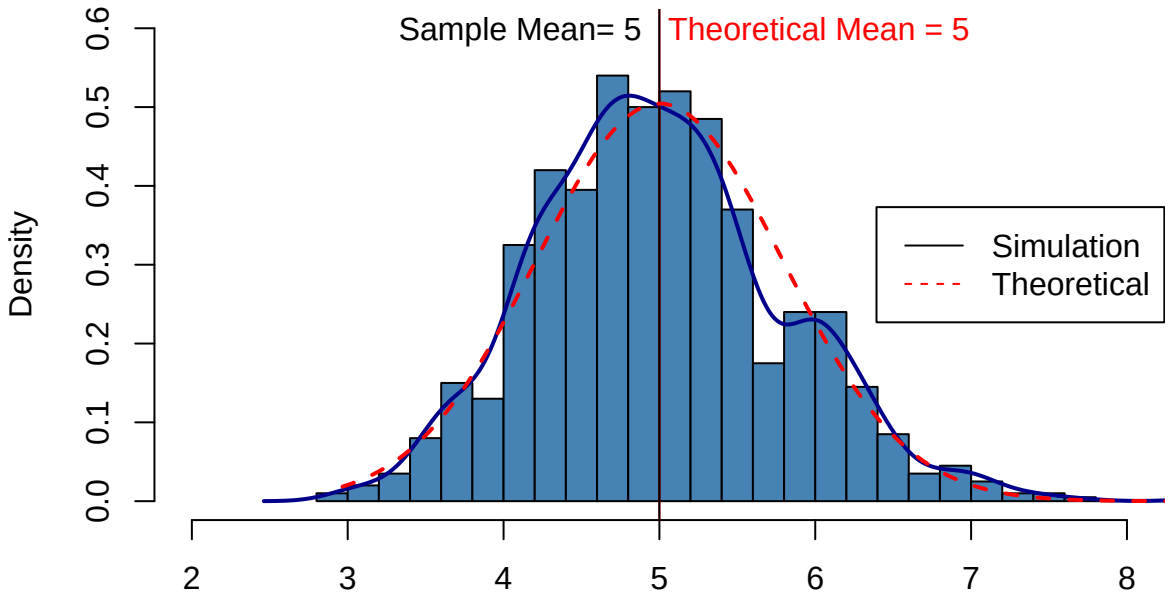



Figure 1: Averages for Exponential Distribution (lambda= 0.2)

**Sample Versus Theoretical :**

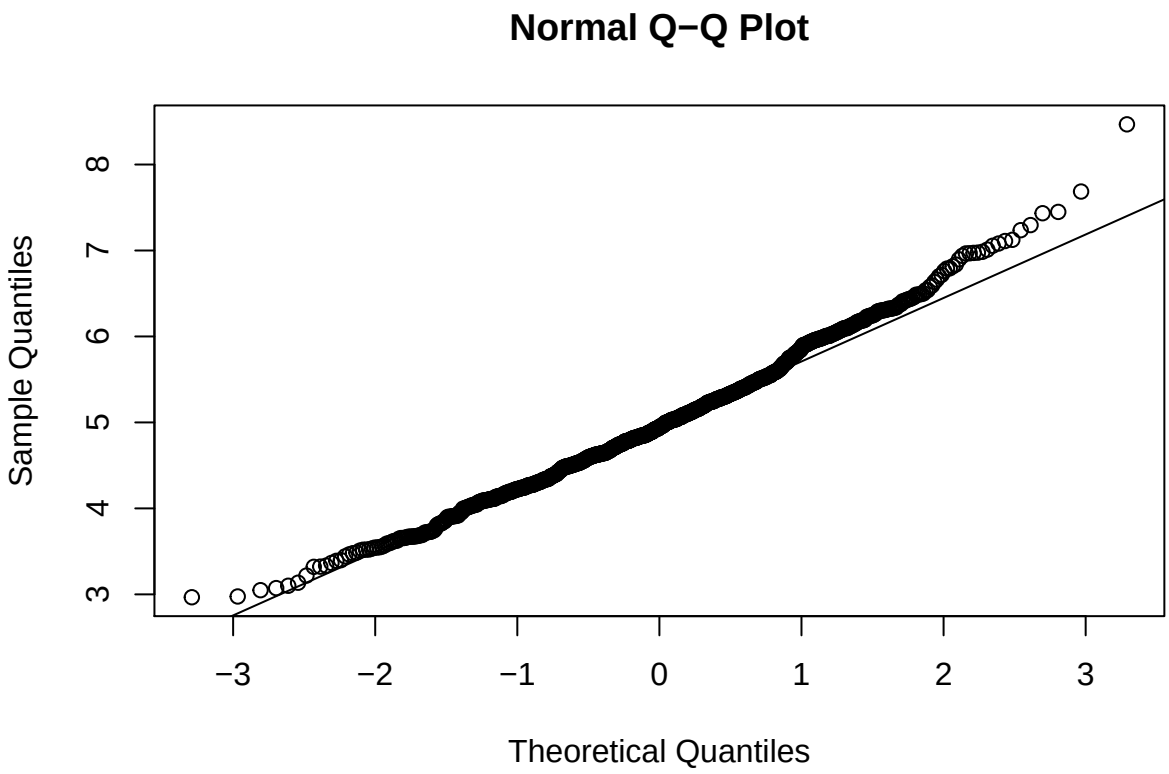
The sample mean is 5, and sample variance is 0.634,
while the theoretical mean is $\frac{1}{\lambda} = 5$ and theoretical variance is $\frac{1}{(n*\sigma^2)} = 0.625$.

## Does the simulation approximate the Normal Distribution?:

From figure.1 it can be seen that the simulation distribution approximates a Gaussian distribution.

A quantile to quantile (q-q) plot also suggests a normal distribution.

```r
qqnorm(meansim); qqline(meansim)
```


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