Part 1: Simulation Exercise

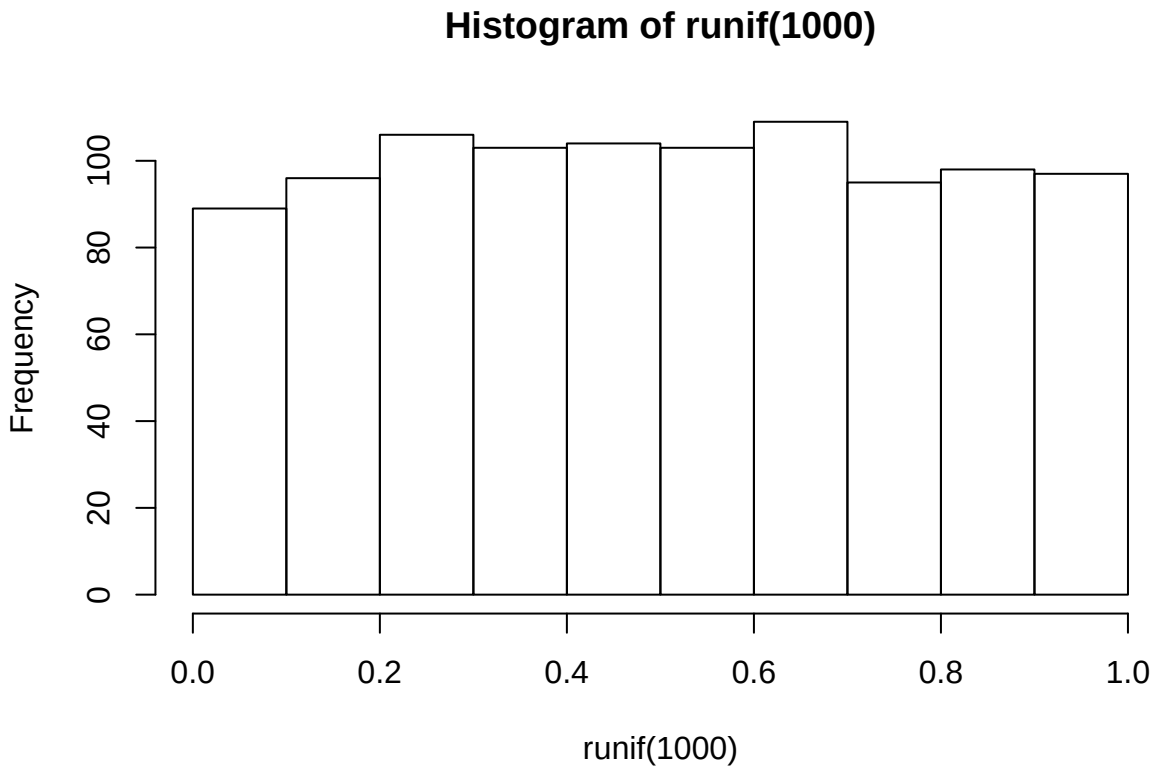
Sergio Paz

Overview

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.

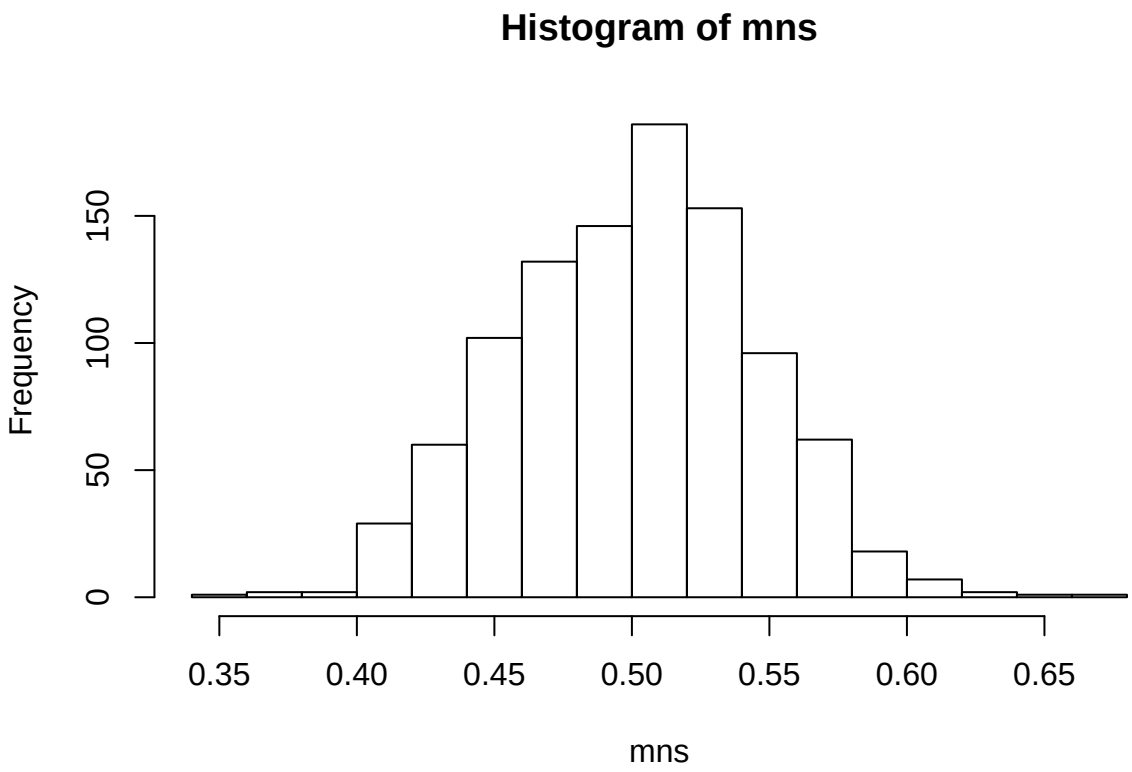
Here is the distribution of 1000 random uniform

```
hist(runif(1000))
```



and the distribution of 1000 averages of 40 random uniforms

```
mns = NULL
for(i in 1 : 1000) mns = c(mns, mean(runif(40)))
hist(mns)
```

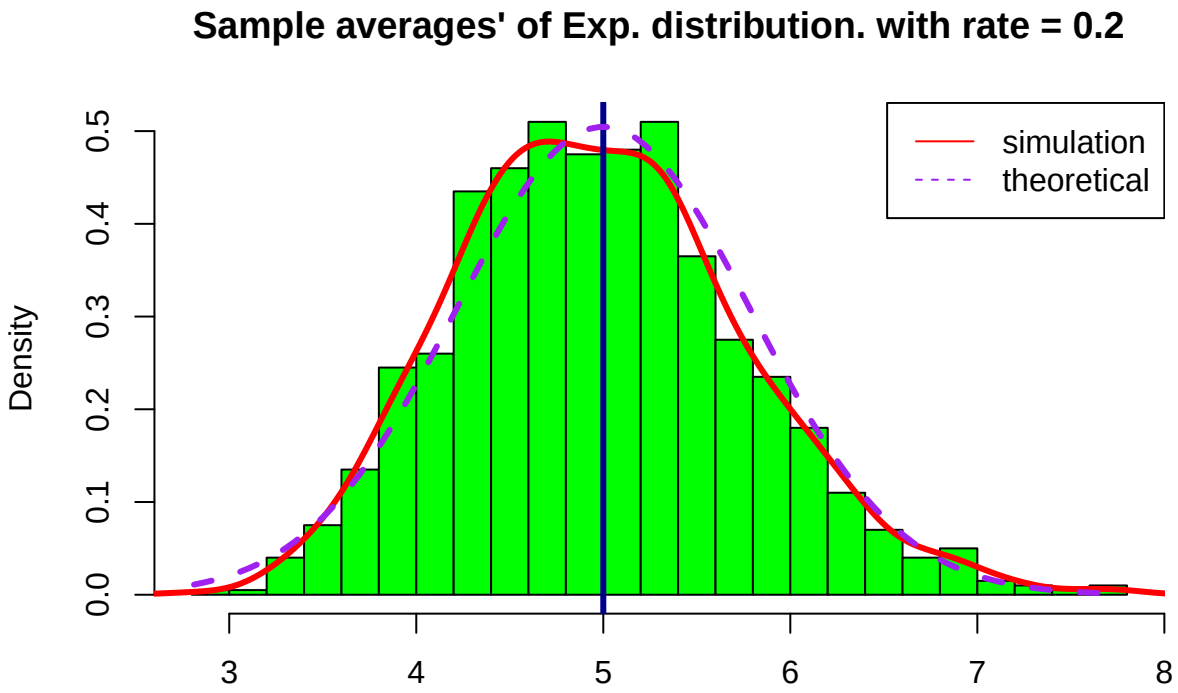


Simulation

```
set.seed(1234)       #set the seed
lambda <- .2         #lambda
n <- 40              #sample
simulations <- 1000 #number of test
d.mean <- 1/lambda
d.sd <- 1/lambda

exp.sim <- matrix(rexp(simulations*n, rate = lambda), simulations, n)
means <- rowMeans(exp.sim)
```

Graphic



Sample mean versus theoretical mean

```
sample.mean <- mean(means)
theoretical.mean <- 1/lambda
```

The deviation of the sample mean `4.9742388` from the theoretical mean `5` is approximately `0.5`%

Sample variance versus theoretical variance

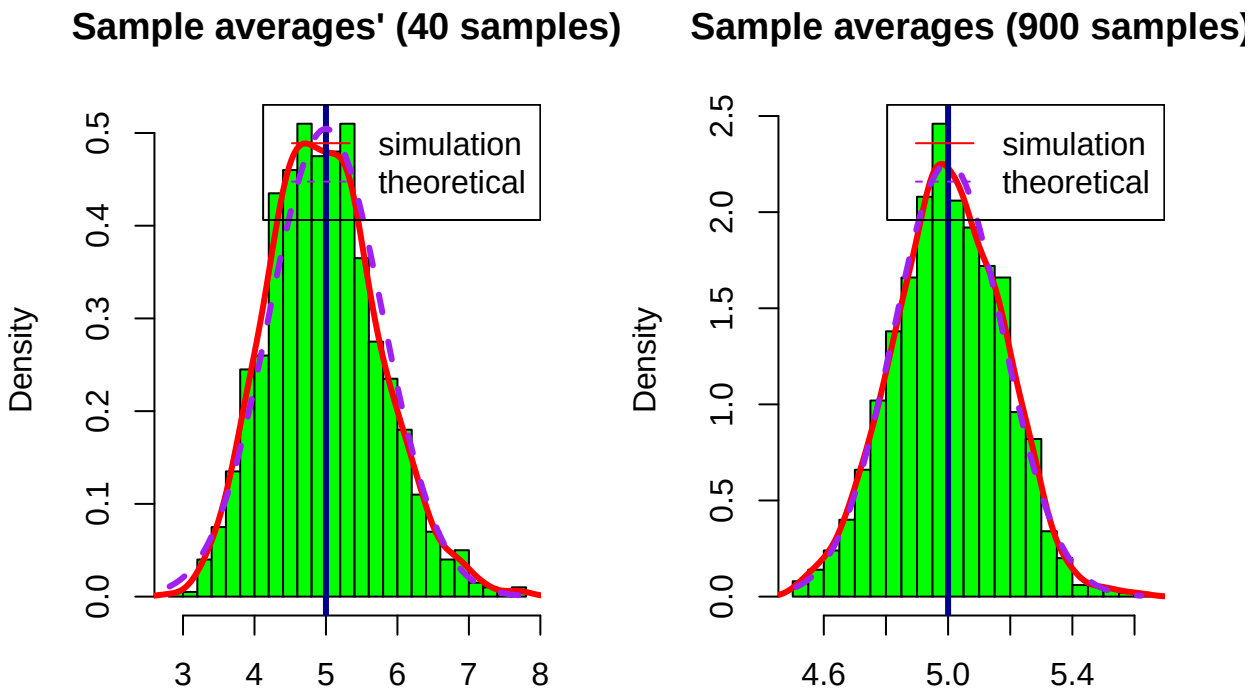
```
sample.sd <- sd(means)
theoretical.sd <- (d.sd)/sqrt(n)

sample.var <- var(means)
theoretical.var <- 1 / (lambda ** 2 * 40)
```

The sample variance `0.5949702` is different from the theoretical variance `0.625` by `4.8`% and the sample standard deviation `0.7713431` diffrantiates from the theoretical standard deviation `0.7905694` by `2.4`%

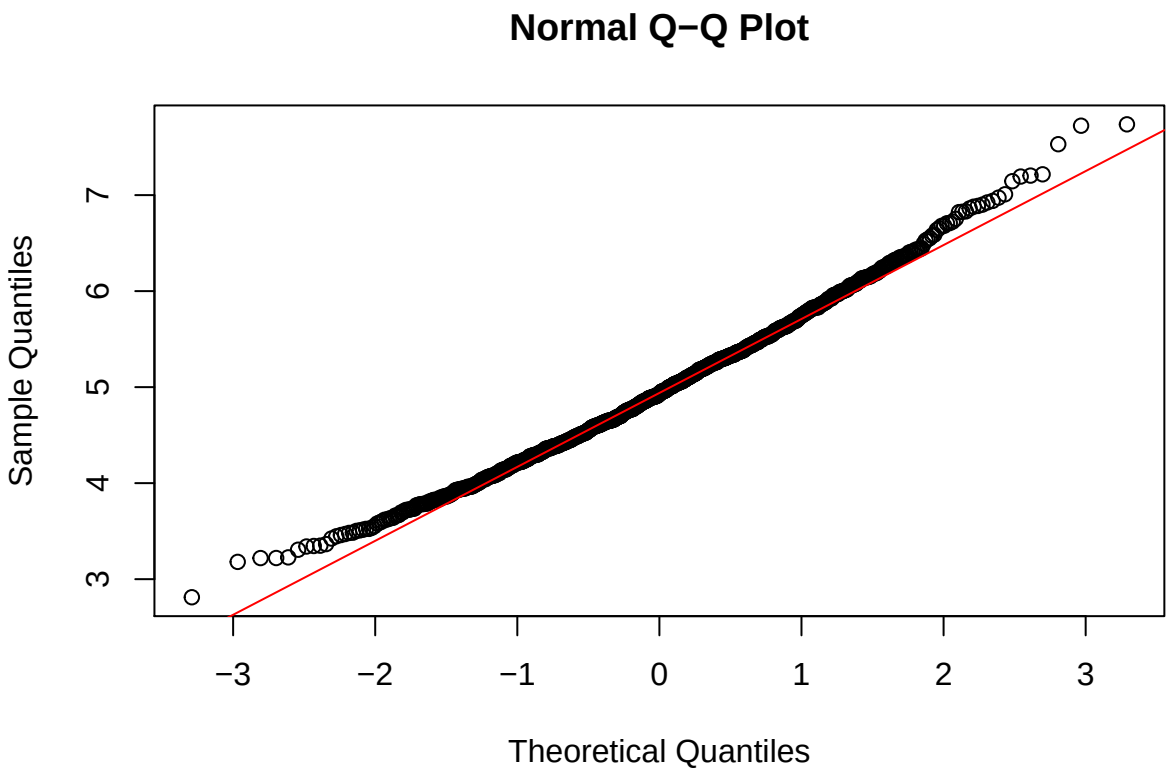
Distribution

```
mult = 20
par(mfrow=c(1,2))
means2 <- rowMeans(matrix(rexp(simulations*n*mult, rate=lambda),
                  simulations, n * mult))
plot.hist("Sample averages' (40 samples)", means, n)
plot.hist("Sample averages (900 samples)", means2, n * mult)
```



Graphic

```
qqnorm(means, col="black")
qqline(means, col="red")
```



A Q-Q plot indicates as well that the 40 exponentials are quite close to the normal distribution.