

# Statistical Inference Course Project Part 1

## Overview

In this project we will investigate the exponential distribution in R and compare it with the Central Limit Theorem. In this report we will use simulation to illustrate the properties of the distribution of the mean of 40 exponentials.

We will show the sample mean and compare it to the theoretical mean of the distribution. Then we will show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. In the end we will show that the distribution is approximately normal.

## Simulations

The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` is the rate parameter. We set `lambda = 0.2` for all of the simulations. We will investigate the distribution of averages of 40 exponentials and will do a thousand simulations.

```
lambda <- 0.2
nr.exp <- 40
sim <- 1000

# we will calculate and store the means in a data frame
df.means <- data.frame(means=numeric(sim))

for (i in 1:sim) {
  sims <- rexp(nr.exp,lambda)
  df.means[i,1] <- mean(sims)
}
```

## Sample Mean vs Theoretical Mean

The mean of exponential distribution is  $1/\lambda$  and the standard deviation is also  $1/\lambda$ . We will compare the sample mean to the theoretical mean.

Here we calculate the theoretical mean:

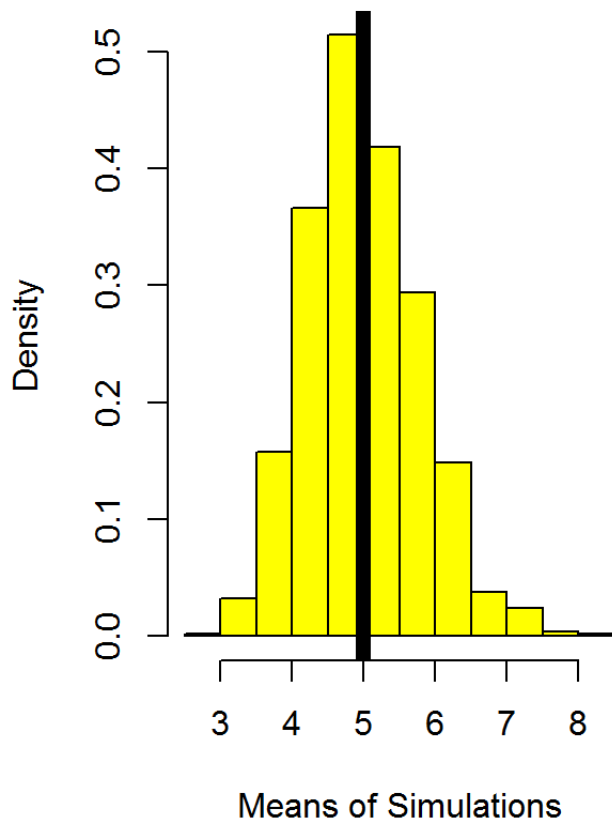
```
theo.mean <- 1/lambda # the theoretical mean
```

And the mean of the simulated means:

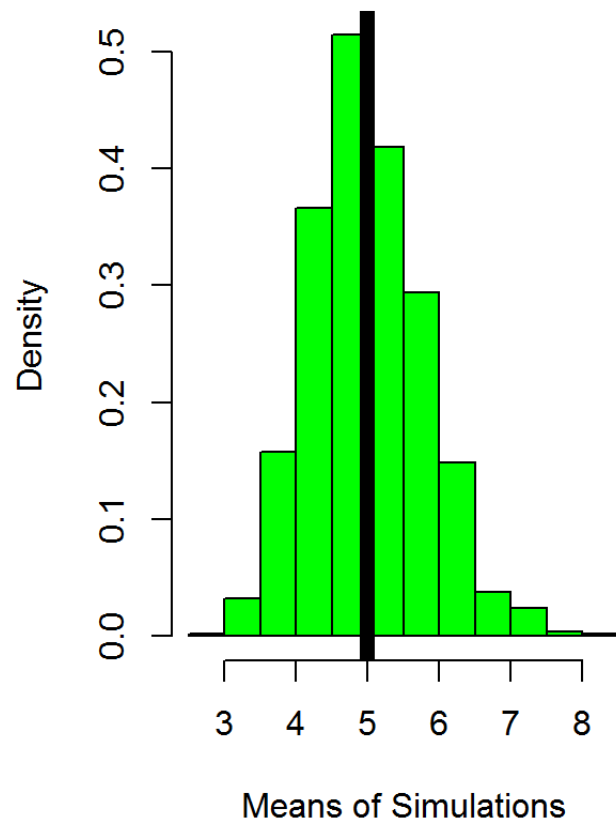
```
real.mean <- mean(df.means$means)
```

The actual simulated sample mean is very close to the theoretical mean of 5. In the plot below we also notice the histograms are almost identical.

**Theoretical Mean of 5**



**Actual Mean of 4.993347**



## Sample Variance vs Theoretical Variance

The theoretical variance can be defined as below:

```
theo.var <- ((1/lambda)^2)/nr.exp
```

Then we determine the actual variance of the simulations means:

```
real.var <- var(df.means$means)
```

We notice the theoretical variance of 0.625 is very close to the real one of 0.6260397.

## The Distribution

The plot below shows that the distribution is approximately normal, the curve of a normal distribution and the actual distribution curve are similar.

## Comparison with a Normal Distribution

