# Illustrate the Properties of the Distribution of the Mean of 40 Exponentials

*Aravind Sesetty*

*Saturday, January 24, 2015*

---

## Overview

This project will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution will 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.
We will investigate the distribution of averages of 40 exponentials. Note that we will do a 1000 simulations.

### Setting Global Option for knitr

- Set `echo=TRUE` so that the R code is displayed in the Report.
- Set `results='asis'` so that the result from R code is displayed in the Report.
- Set `fig.width=5` (Plot Width) & `fig.height=5` (Plot Height).

## Simulations

```
# Set seed for reproducibility
set.seed(01222015)

# Rate parameter for Exponential Distribution
lambda <- 0.2

# Sample Size
n <- 40

# Number of Simulations
num_of_simulations <- 1000

# 1000 Different Samples each of size 40
simulated_exponentials <- replicate(num_of_simulations, rexp(n, lambda))
```

## Sample Mean vs Theoretical Mean

```
# Calculate mean of every sample i.e. 1000
means_of_simulated_exponentials <- apply(simulated_exponentials, 2, mean)

sample_mean <- round(mean(means_of_simulated_exponentials), digits = 3)
message("Sample Mean : ", sample_mean)
```

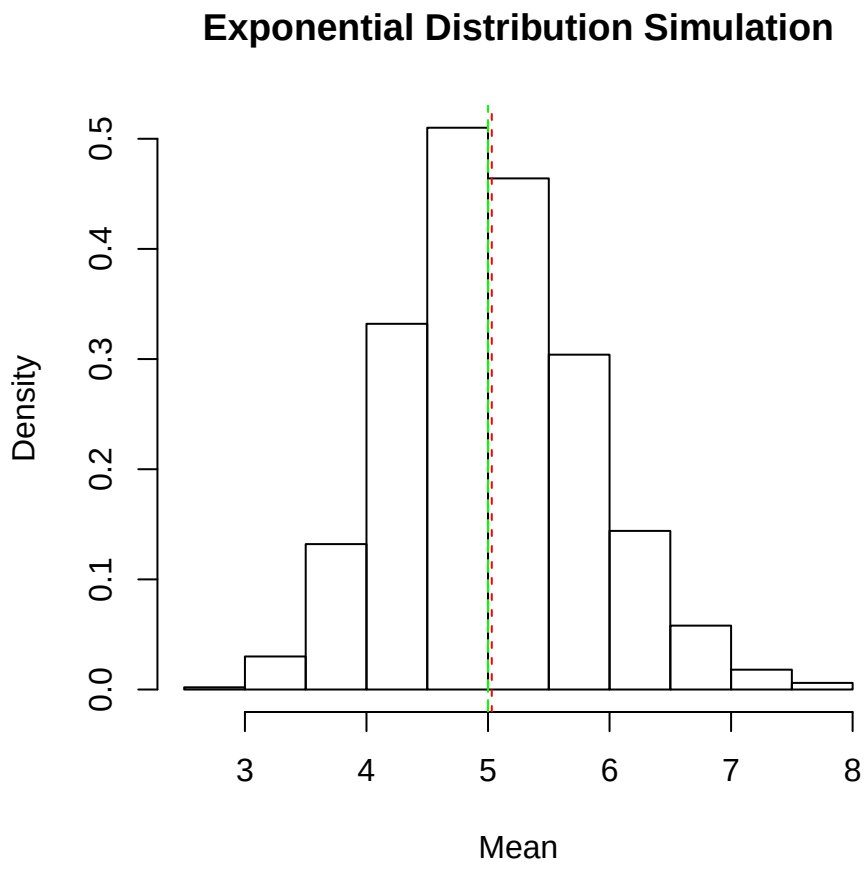
```
## Sample Mean : 5.031
```

```
theoretical_mean <- 1/lambda
message("Theoretical Mean : ", theoretical_mean)
```

```
## Theoretical Mean : 5
```

**1. Show the sample mean and compare it to the theoretical mean of the distribution.**
**A. Observation :** When compared **Sample Mean (5.031)** with **Theoretical Mean (5)**, the Sample Mean is almost near to Theoretical Mean with a difference of **0.031**.

```
# Plot Histogram
hist(means_of_simulated_exponentials,
     prob=T,
     main = "Exponential Distribution Simulation",
     xlab = "Mean")
abline(v=sample_mean, col = "RED", lty = 2)
abline(v=theoretical_mean, col = "GREEN", lty = 5)
```




**Observation :** The Histogram is plotted with Means of all Simulated Exponentials. The Theoretical Mean is displayed in **Green** and the Sample Mean is displayed in **Red**.
Therefore, the Center of Distribution of Means of 40 Exponentials is close to the Theoretical Center of Distribution.

## Sample Variance vs Theoretical Variance

```
sample_variance <- round(var(means_of_simulated_exponentials), digits = 3)
message("Sample Variance : ", sample_variance)
```

```
## Sample Variance : 0.614
```

```
theoretical_variance <- (1/lambda)^2/n
message("Theoretical Variance : ", theoretical_variance)
```

```
## Theoretical Variance : 0.625
```

**2.1. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.**
**A. Observation :** When compared **Sample Variance (0.614)** with **Theoretical Variance (0.625)**, the Sample Variance is almost near to Theoretical Variance with a difference of **0.011**.
Therefore, the Sample Variance in Distribution of Means of 40 Exponentials is close to the Theoretical Variance in the Distribution.

## Sample Standard Deviation vs Theoretical Standard Deviation

```
sample_standard_deviation <- round(sd(means_of_simulated_exponentials), digits = 3)
message("Sample Standard Deviation : ", sample_standard_deviation)
```

```
## Sample Standard Deviation : 0.783
```

```
theoretical_standard_deviation <- 1/lambda/sqrt(n)
message("Theoretical Standard Deviation : ", theoretical_standard_deviation)
```

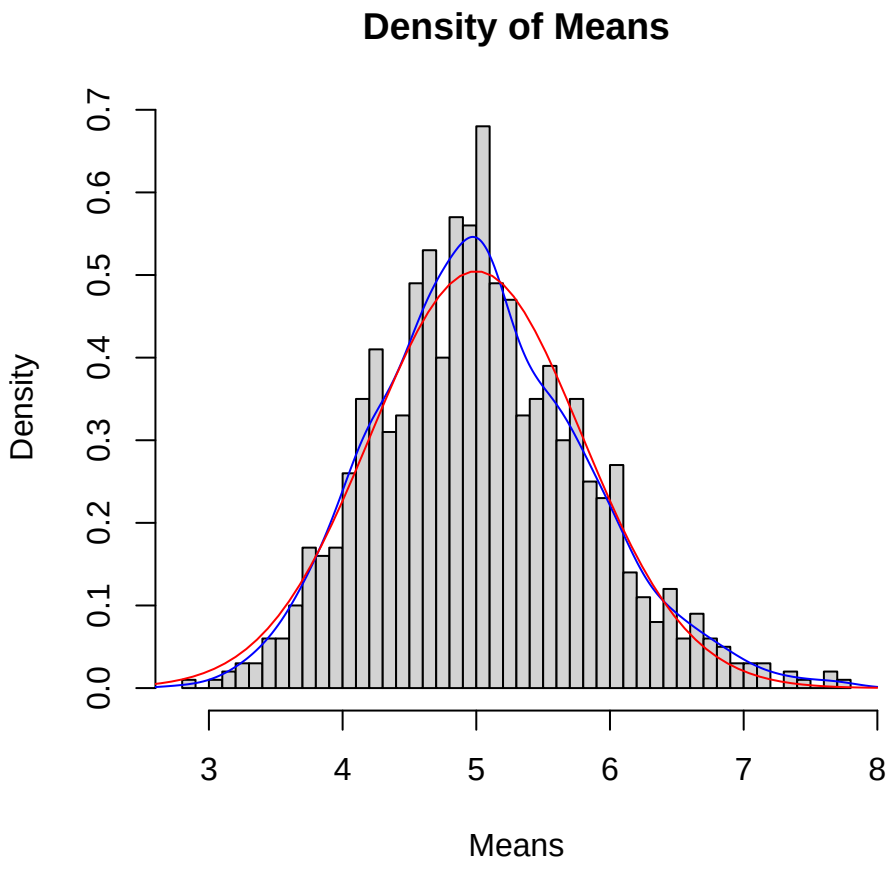
```
## Theoretical Standard Deviation : 0.790569415042095
```

**2.2. Show how variable the sample is (via Standard Deviation) and compare it to the theoretical Standard Deviation of the distribution.**
**A. Observation :** When compared **Sample Standard Deviation (0.783)** with **Theoretical Standard Deviation (0.790)**, the Sample Standard Deviation is almost near to Theoretical Standard Deviation with a difference of **0.007**.
Therefore, the Standard Deviation in Distribution of Means of 40 Exponentials is close to the Theoretical Standard Deviation in the Distribution.

## Distribution

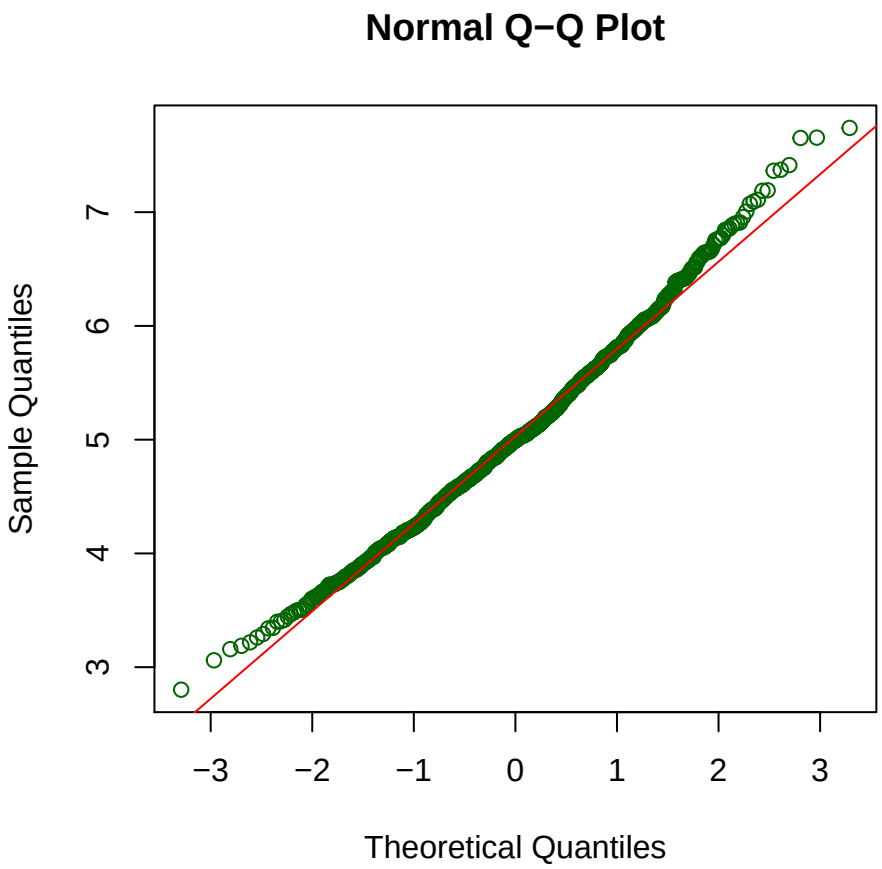
```
hist(means_of_simulated_exponentials,
     breaks = n,
     prob=T,
     col="LIGHT GRAY",
     main="Density of Means",
```

```
     xlab = "Means",
     ylab="Density")
lines(density(means_of_simulated_exponentials), col="BLUE")
curve(dnorm(x, theoretical_mean, theoretical_standard_deviation), 0, 8, col="RED", add=T)
```




**3. Show that the distribution is approximately normal.**
**A. Observation :** The Histogram is plotted with Means of all Simulated Exponentials. The Theoretical Distribution is displayed in **Red** and the Sample Distribution is displayed in **Blue**. Its clear that the Sample Distribution is very close to Theoretical Distribution.

```
qqnorm(means_of_simulated_exponentials, col="DARK GREEN");
qqline(means_of_simulated_exponentials, col="RED");
```




**Observation :** From the Q-Q plot, the Sample Distribution of Means of 40 Exponentials is very close to a Theoretical Distribution.