# Analysis of Asymptotic Behavior of Exponential Distribution

*E Drake*

### Overview:

This is an investigation into the exponential distribution. We will compare it to the normal distribution using the Central Limit Theorem.

The simulation will find the average of 40 exponentials. Each exponential will have $\lambda = 0.2$, in other words the mean and variance will be 5.

By carrying out the simulation a total of 1000 times I will be able to analyse this new distribution and compare the experimental results to the behavior expected by the theory.

### Simulations:

The simulation will calculate the mean of 40 exponential random variables with lambda = 0.2.

The R programming language makes it easy to create 40 exponentials and find their mean using the following code:

```
lambda = 0.2
n = 40
set.seed(123456)
mean(rexp(40, lambda))
```

```
## [1] 4.671228
```

The simulation will run 1000 times. In R, this can be acheived using a for loop:

```
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(n, lambda)))
```

The list mns contains 1000 means, each mean is the average of 40 exponentials.

### Sample Mean versus Theoretical Mean:

Using the Law of Large Numbers, we know that sample mean is an unbiased estimator of the distributional mean.

The underlying exponential distribution has mean 5 (that is $1/\lambda$) so we would expect the mean of our simulation to also be 5.

Let’s check the sample mean using the following R code:

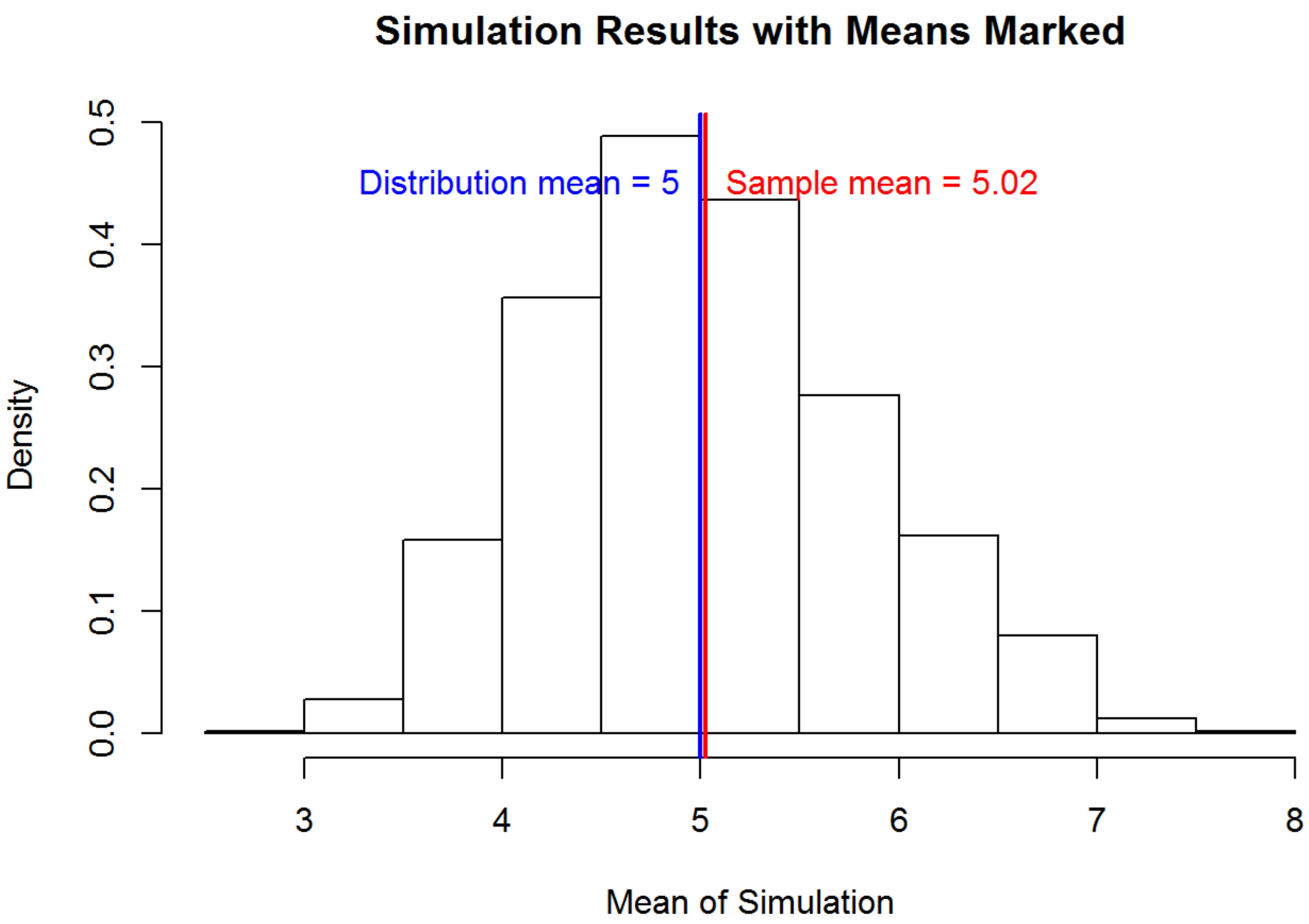
```
print(mean(mns))
```

```
## [1] 5.02251
```

As expected, the sample mean is very close to the distributional mean.

We can plot these means on the histogram using the following code:

```
hist(mns,freq = FALSE, main = "Simulation Results with Means Marked",
     xlab = "Mean of Simulation")
abline(v = mean(mns), lwd = 2, col = "RED")
text(x = mean(mns), y = 0.45, pos = 4,
     labels = paste("Sample mean =", round(mean(mns),2)), col = "RED")
abline(v = 5, lwd = 2, col = "BLUE")
text(x = 5, y = 0.45, pos = 2,
     labels = "Distribution mean = 5", col = "BLUE")
```




## Sample Variance versus Theoretical Variance:

Since the experiments are independent, our theory states that the variance will be $\sigma^2/n$.

We know that the distribution standard deviation, $\sigma$, is 5 and n is 40 so the theoretical variance will be 0.625.

Using R we can find the sample variance of the simulation very easily:

```
print(var(mns))
```

```
## [1] 0.6574878
```

As expected, the sample variance is very close to the theoretical value.

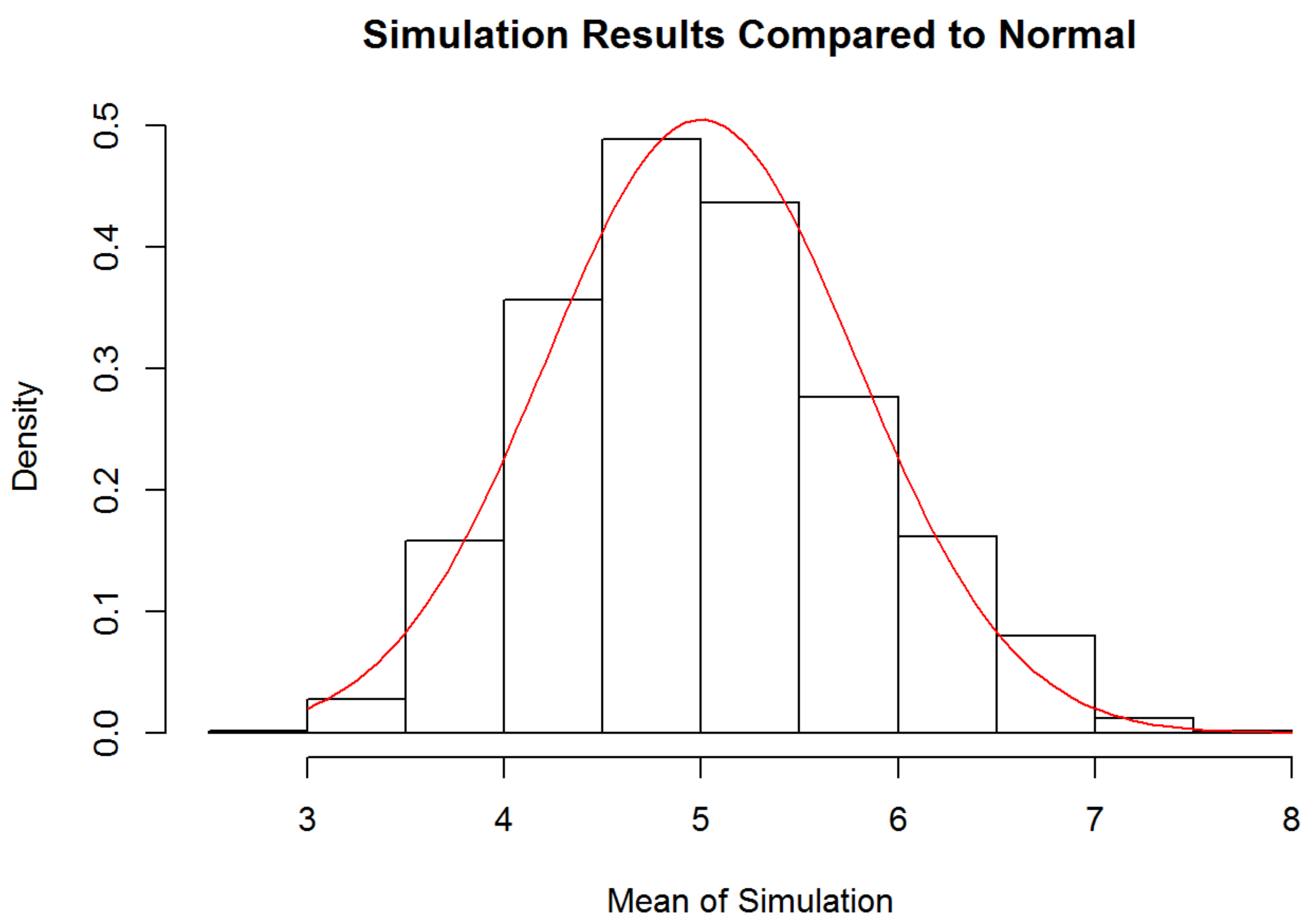
## Central Limit Theorem:

The Central Limit Theorem says that the distribution of averages can be approximated by the normal distribution and that the approximation gets better for larger sample sizes.

We know that the theoretical mean of the distribution is 5 and the theoretical variance is 0.625.

Let's compare our histogram to the density curve for the normal distribution with these properties:

```
hist(mns,freq = FALSE, main = "Simulation Results Compared to Normal",
     xlab = "Mean of Simulation")
x <- seq(3, 8, length=100)
y <- dnorm(x, mean = 5, sd = sqrt(0.625))
lines(x,y, col = "RED")
```




We can see that the normal curve is a good, but not perfect, approximation of the simulation results. If we increased the sample size (for example from 40 up to 100) then we would expect an even better approximation.