Exponential Distribution Simulation

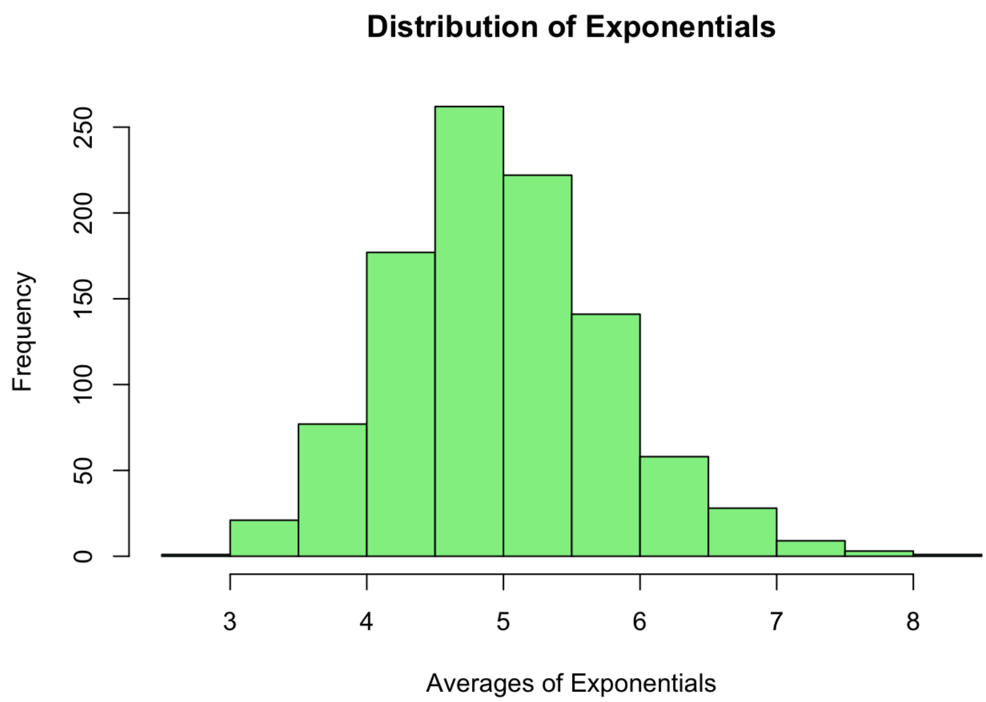
Overview

In this project, we will investigate the exponential distribution in R and compare it with the Central Limit Theorem. We will investigate the distribution of averages of 40 exponentials. The simulation will be repeated 1000 times and lambda = 0.2 will be used for all of the simulations.

Simulating the Exponential Distribution

We set a seed to ensure reproducibility and compute the averages of 40 exponentials. Its distribution is plotted as shown below:

```
set.seed(42)
exp_mns <- NULL
lambda <- 0.2
n <- 40
for (i in 1 : 1000) exp_mns <- c(exp_mns, mean(rexp(n, lambda)))
hist(exp_mns, col="lightgreen", xlab="Averages of Exponentials", main="Distribution of Exponentials")
```



1. Show the sample mean and compare it to the theoretical mean of the distribution.

We compute the mean of the averages of 40 exponentials and the theoretical mean of an exponential distribution, which is 1/lambda. The difference between the 2 means is computed below:

```
s_mean <- mean(exp_mns)
s_mean
```

```
## [1] 4.986508
```

```
t_mean <- 1 / lambda
t_mean
```

```
## [1] 5
```

```
paste("Difference between sample and theoretical mean is", abs(s_mean - t_mean))
```

```
## [1] "Difference between sample and theoretical mean is 0.0134916825454692"
```

We observe that the difference is very small, indicating that the sample mean is a good approximation of the theoretical mean.

2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

We compute the variance of the averages of 40 exponentials and the theoretical variance of an exponential distribution, which is the square of the standard deviation, 1/lambda, divided by the number of values. The difference between the 2 variances is computed below:

```
s_var <- var(exp_mns)
s_var
```

```
## [1] 0.6344405
```

```
t_var <- (1 / lambda)**2/n
t_var
```

```
## [1] 0.625
```

```
paste("Difference between sample and theoretical variance is", abs(s_var - t_var))
```

```
## [1] "Difference between sample and theoretical variance is 0.0094405206689484"
```

We observe that the difference is again very small, indicating that the sample variance is a good approximation of the theoretical variance.

3. Show that the distribution is approximately normal.

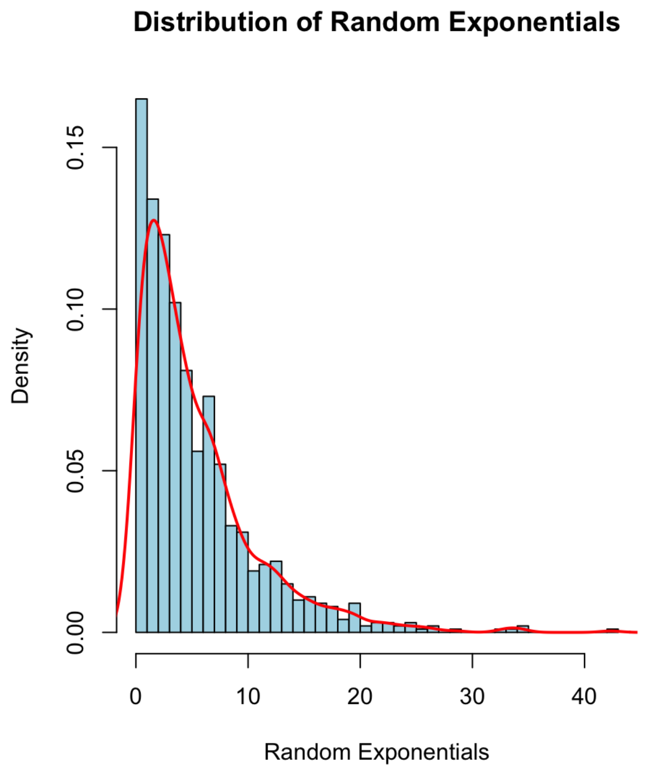
We want to determine the difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials.

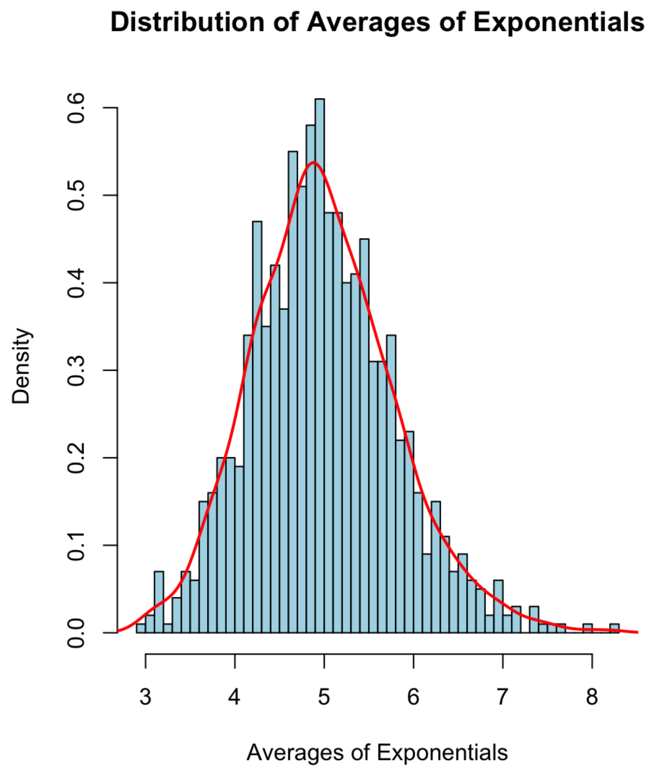
To find the distribution of a large collection of random exponentials, we simulate 1000 random exponentials, as shown below:

```
set.seed(42)
rand_exp <- rexp(1000, lambda)
```

Now, we plot these 2 distributions and their corresponding density plots.

```
par(mfrow=c(1,2))
hist(rand_exp, prob=TRUE, breaks = n, col="lightblue", xlab="Random Exponentials", ylab="Density", main="Dis
tribution of Random Exponentials")
lines(density(rand_exp), lwd=2, col="red")
hist(exp_mns, prob=TRUE, breaks = n, col="lightblue", xlab="Averages of Exponentials", ylab="Density", main=
"Distribution of Averages of Exponentials")
lines(density(exp_mns), lwd=2, col="red")
```





We observe that the distribution of a large collection of random exponentials is skewed to the right and is not normally distributed. On the other hand, the distribution of a large collection of averages of 40 exponentials has a bell-shaped curve and appears to be normally distributed.

This validates the Central Limit Theorem, which states that the distribution of averages of iid random variables becomes that of a standard normal as the sample size increases.