Part 1: Simulation Exercise

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

First, let's set our lambda value and calculate the theoretical mean, given by 1/lambda.

```
# Lambda
l = 0.2

# Theoretical mean
1/l
```

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

Next, let's generate a a distribution of 1000 exponential samples, where each sample contains 40 exponentials.

```
# Set the random seed for reproducible results
set.seed(12345)

# Set N
n <- 40

means <- NULL
# Generate 1000 distributions of 40 exponentials,
# taking the mean of each one and putting it into
# a collection of means
for (i in 1:1000){
  means = c(means, mean(rexp(n, l)))
}
```

The mean of this distribution is a close approximation of the theoretical mean. This is explained by the Law of Large Numbers: as sample size increases, the sample mean converges on the theoretical / population mean.

```
# Mean of means
mean(means)
```

```
## [1] 4.971972
```

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

The theoretical standard deviation is given by 1/lambda. To get the variance we square this value (because SD is the square root of the variance). Finally, to get the theoretical variance for the distribution of averages, we must take this and divide it by the sample size.

```
# Theoretical variance
(1/l)^2 / n
```

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

Using the distribution of 1000 exponentials that we generated before, we can calculate the variance. We see that the actual variance for this distribution of averages is very close to the theoretical variance.

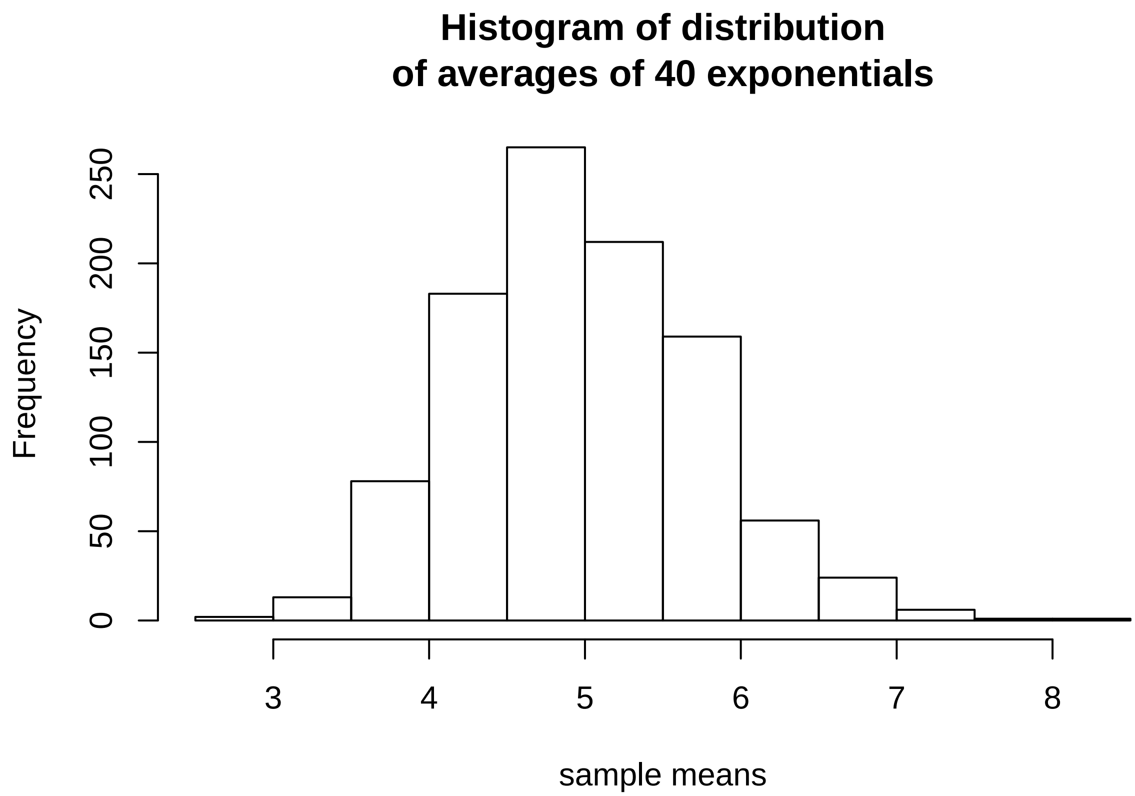
```
# Variance of the distribution of averages, rounded to 5 decimals
round(var(means), 3)
```

```
## [1] 0.595
```

Show that the distribution is approximately normal

Next we can look at the distribution of averages of 40 exponentials. We see that it is approximately normal. We know this because of the symmetrical, bell-curved shape.

```
# Histogram
hist(means,
     main = "Histogram of distribution\nof averages of 40 exponentials",
     xlab = "sample means")
```

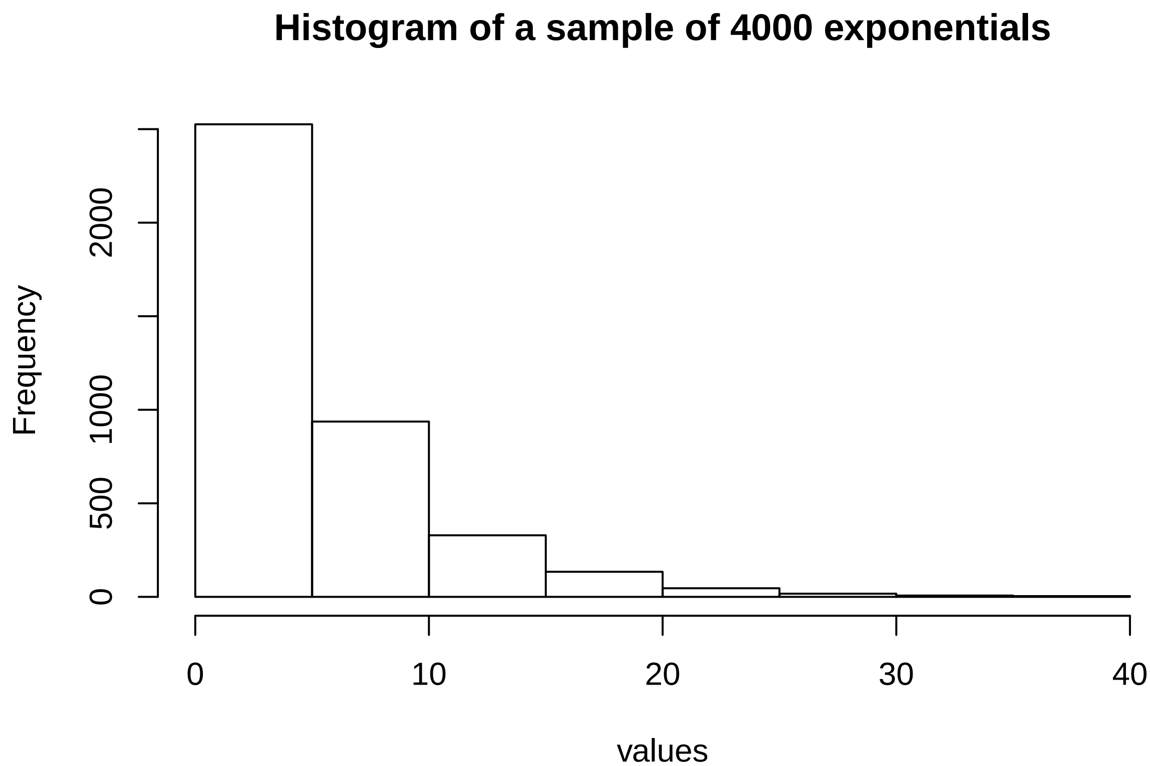


If we compare this to a large collection of exponentials we see that it is non-normal. We know it is non-normal because it does not have a symmetrical, bell-curved shape. It is skewed towards 0.

```
# Set the random seed for reproducible results
set.seed(12345)

# Set N to something large
n = 4000

# Histogram of the generated sample
hist(rexp(n, 1),
     main = "Histogram of a sample of 4000 exponentials",
     xlab = "values")
```


Histogram of a sample of 4000 exponentials
Frequency
0
500
1000
2000
0
10
20
30
40
values