Analysis of the Exponential Distribution

Synopsis

The purpose of this report is to analyze the exponential distribution. Both the mean and standard deviation of the exponential distribution are 1/lambda, where lambda is the rate parameter. In this report we are using a lambda value of 0.2.

Data Simulation

For our simulation, we are going to sample 40 values from the exponential distribution in 1000 simulation runs. The code to run the simulation is as follows:

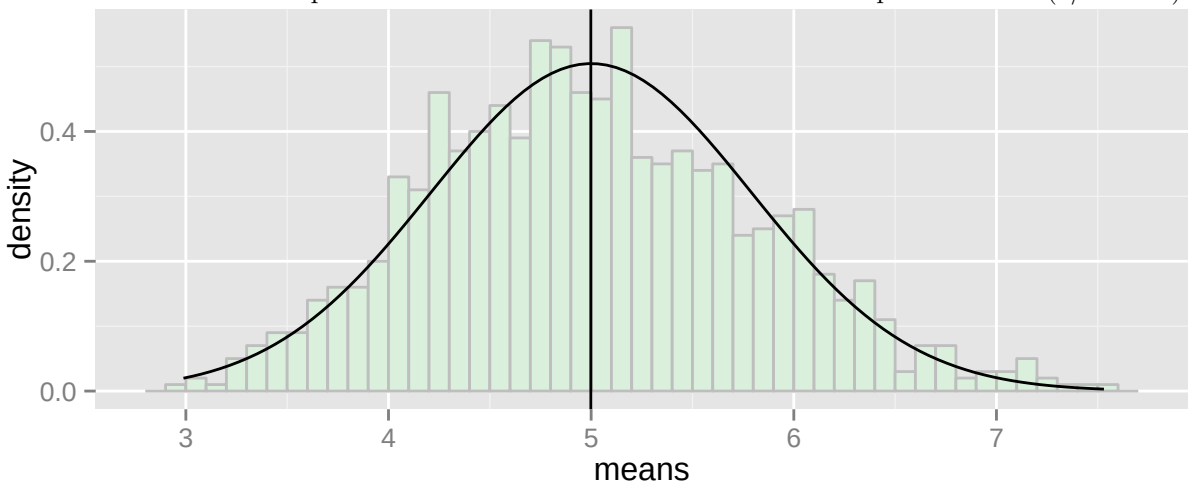
```
library(ggplot2)

set.seed(25)
lambda <- 0.2
expectedMean <- expectedSd <- 1/lambda
numSamples <- 40
numSims <- 1000
sampleData <- matrix(rexp(numSamples * numSims, lambda), ncol = numSamples)
means <- apply(sampleData, 1, mean)
variances <- apply(sampleData, 1, var)
```

The means vector now contains 1000 values that are the means of 40 exponential distribution samples in each simulation run. The variances vector contains the variance of each simulation run.

Distribution of Sample Means

Let's first look at the distribution of our sample means. If we plot a histogram of the sample means we see that they are centered around the theoretical mean of 1/lambda (5). We also plot a normal distribution curve with a mean of 1/lambda and a standard deviation of (1/lambda)/sqrt(40) and see that the distribution of our sample means is approximately normal. We also print out the variance of our sample means and see that it is close to its expected value (1/labmda)^2/40.

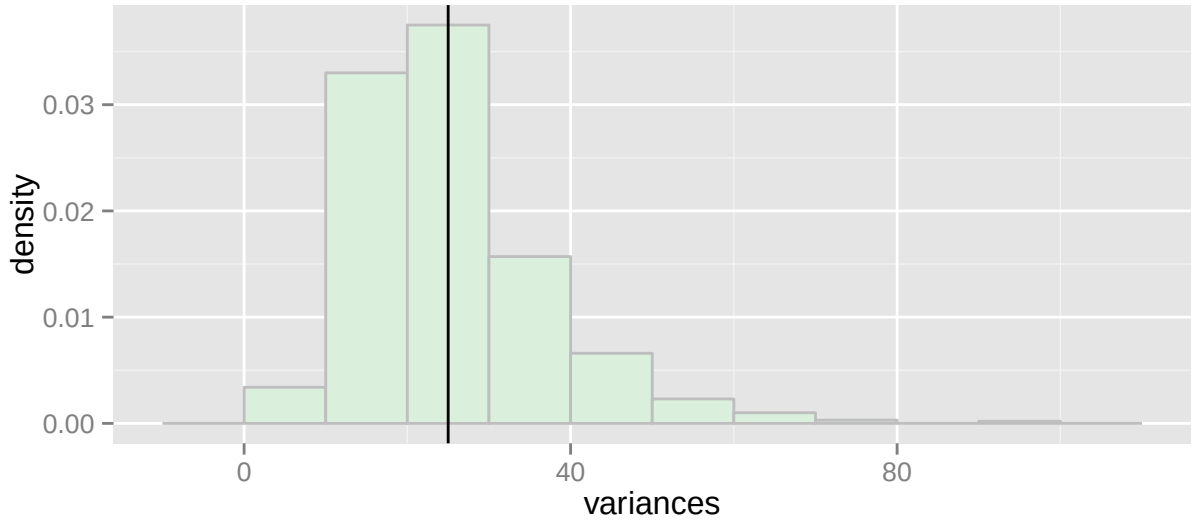


```
## [1] "Our sample variance =  0.661985767399141"
```

```
## [1] "Our expected sample variance = 0.625"
```

Distribution of Sample Variances

Now let's look at the distribution of sample variances. A plot of these variances shows that they are centered around (1/lambda)^2 (25) as expected.



Confidence Interval

Finally we calculate the 95% confidence interval for our sample means and show that the interval covers over 95% of our samples..

```
interval <- meanOfMeans + c(-1, 1) * qt(.975, numSamples - 1) * (1/lambda) / sqrt(numSamples)
covered <- sapply(means, function(mean)
{
      interval[1] < mean & interval[2] > mean
})

print(paste("Our confidence interval =", toString(round(interval, 4))))
```

```
## [1] "Our confidence interval = 3.3994, 6.5976"
```

```
print(paste(sum(covered) / numSims * 100, "% of our sample means are covered by this interval.", sep=""))
```

```
## [1] "95.2% of our sample means are covered by this interval."
```

Complete Rmd File

The complete Rmd file for this report can be found at the following url: exponential_analysis.Rmd