Exploring the Properties of the Exponential Distribution

Matt Moocarme

July 24, 2015

Overview

This report will use simulations to explore the properties of the exponential distribution, given by the formula $\lambda e^{-\lambda x}$. Both the theoretical mean and standard deviation for this distribution is $1/\lambda$ when the number of samples approaces infinity.

Process

Here, I run 1000 simulations, each with 40 samples of an exponential distribution. The mean and variance of of those 40 samples will be calculated and compared to the theoretical estimates. Central limit theorem states that as the number in the sample increases or the number of simulations increase, the sample mean and variance should more accurately represent the theoretical estimate.

First I set the seed so the data can be reproduced and set the control parameters, λ is the rate parameter and n is the number of samples in each simulation

```
set.seed(1)
lambda <- 0.2
n  <- 40
```

Following, a for loop runs over 1000 simulations, for each simulation the mean and the varianceis computed and appended to an array.

```
mns = NULL
vrs= NULL
for (i in 1 : 1000) {
        mySample <- rexp(n,lambda)
        mns <- c(mns, mean(mySample))
        vrs <- c(vrs, var(mySample))
}
```

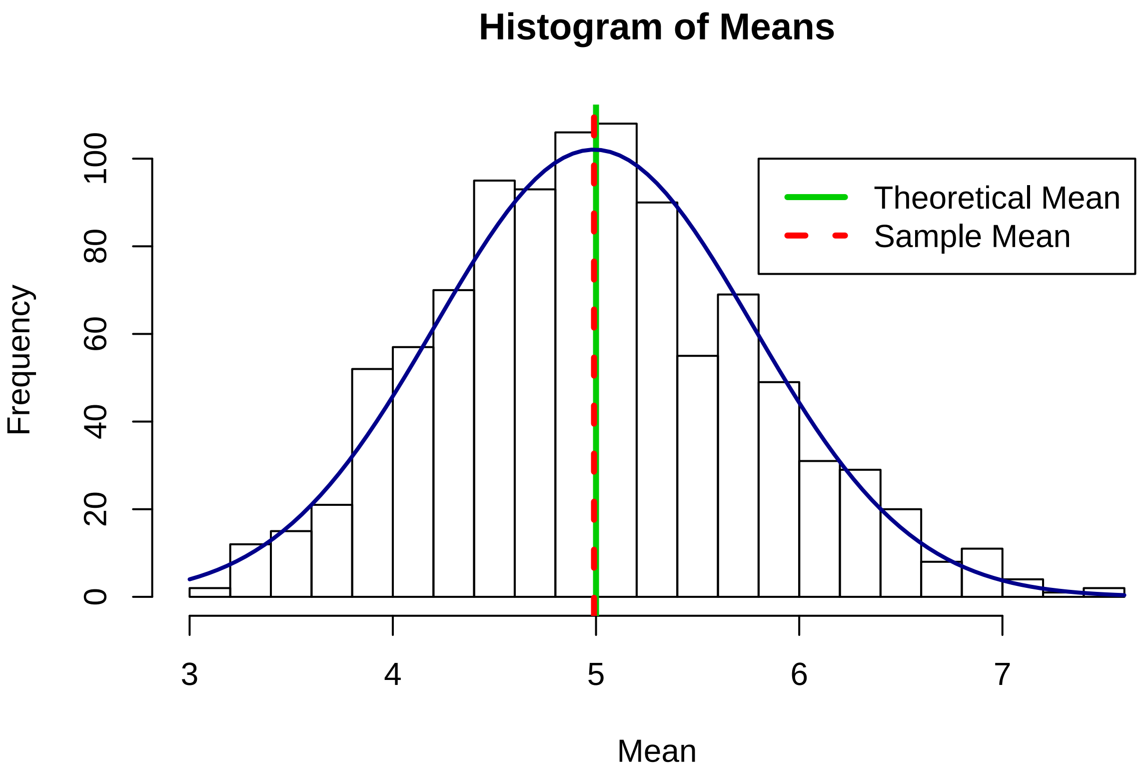
Next, I compute the average of the mean and the variance across all 1000 simulations

```
myMean <- mean(mns)
theoMean <- 1/lambda
myVar <- mean(vrs)
theoVar <- (1/lambda)^2
```

In these simulations the average of the sample means is 4.9900252, and the theoretical mean is 5. The average of the sample variances is 25.064586, and the theoretical variance is 25. We can see that they are very close.

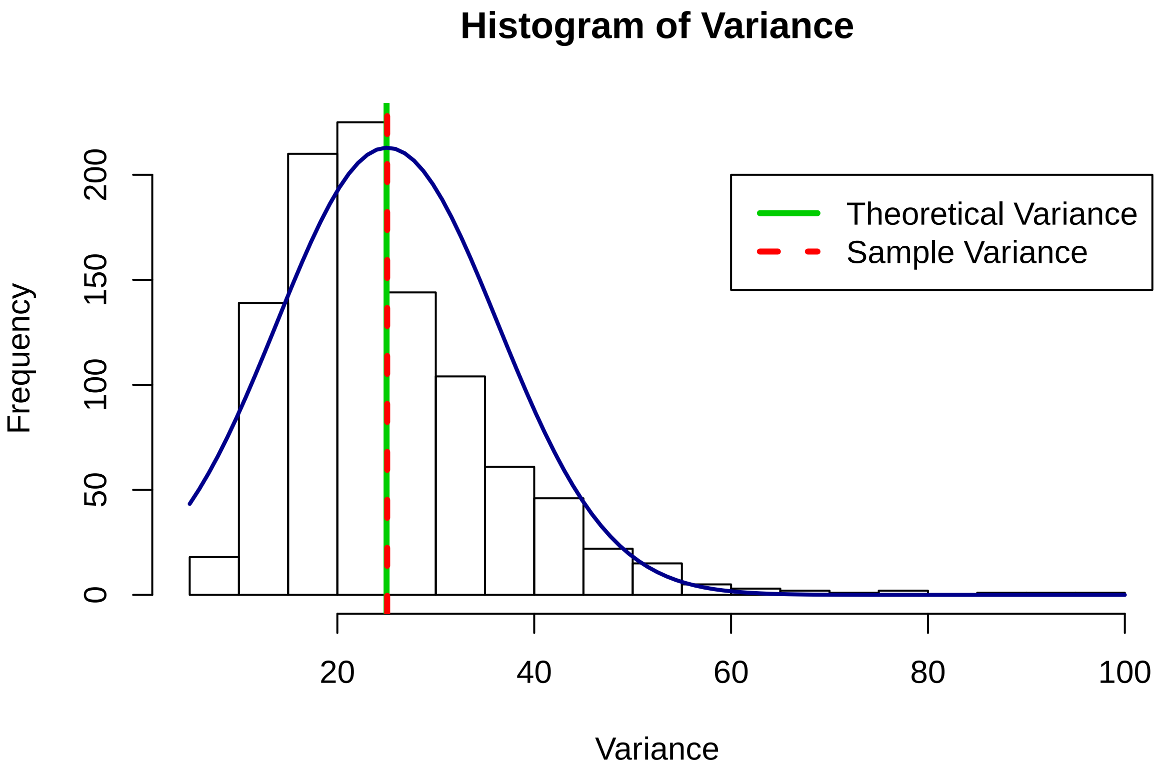
We can plot this as a histogram and show that it is approximately normal.

```r
hist(mns, breaks = 20, xlab = 'Mean', main = 'Histogram of Means')
abline(v = theoMean, col = 3, lwd = 3)
abline(v = myMean, col = 2, lwd = 3, lty = 2)
legend(5.8,100, c('Theoretical Mean', 'Sample Mean')
       , lwd = c(3,3), lty = c(1,2), col = c(3,2))
curve(dnorm(x, mean=myMean, sd=sd(mns))*200, col="darkblue", lwd=2, add=TRUE)
```



With the normal curve overlayed we can see that the mean resembles a normal distribution. We can do the same for the variance.

```r
hist(vrs, breaks = 14, xlab = 'Variance', main = 'Histogram of Variance')
abline(v = theoVar, col = 3, lwd = 3)
abline(v = myVar, col = 2, lwd = 3, lty = 2)
legend(60, 200, c('Theoretical Variance', 'Sample Variance'),
       lwd = c(3,3), lty = c(1,2), col = c(3,2))
curve(dnorm(x, mean=myVar, sd=sd(vrs))*6000, col="darkblue", lwd=2, add=TRUE)
```



The variance does not resemble a normal distribution like the mean.

Conclusion

In conclusion I have shown various properties of the exponential distribution, including the mean and the variance. In accordance with the central limit theorem the mean of a sufficiently large sample size has a normal distribution across the thetheortical mean. This was shown with 1000 simulations, each with a sample size of 40 random variables.