Statistical Inference Project Part 1

Jacob Townson

July 20, 2015

Overview

In this project I will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Note, I did not make an appendix section because I merged my code with the explanations, and still did not go over 3 pages.

Part 1

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

Well, we know that the theoretical mean is $\frac{1}{\lambda}$. Well, since $\lambda = .2$, the theoretical mean is 5. Now I just need the simulation to compare it to. In order to do our simulation, I will run this code:

```
set.seed(2200)
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40, .2)))
```

Note in the above code I set the seed that I will be working with. This is in order to make sure my simulations don't vary from code chunk to code chunk.

Now, if I take the mean of the means, it can be compared to the theoretical mean of 5:

```
mean(mns)
```

```
## [1] 5.00045
```

One can see that the mean of all of the simulated means is 5.00045. This is extremely close to our theoretical mean.

Part 2

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

We know that the theoretical variance is $\frac{1}{\lambda^2}$. Just like the theoretical mean, because $\lambda = .2$, so we know that the theoretical variance is going to be 25, and the theoretical standard deviation to be 5. To do the simulation, I will use a similar method as I did for the mean.

```
set.seed(2200)
vars = NULL
for (i in 1 : 1000) vars = c(vars, var(rexp(40, .2)))
mean(vars)
```

```
## [1] 24.77243
```

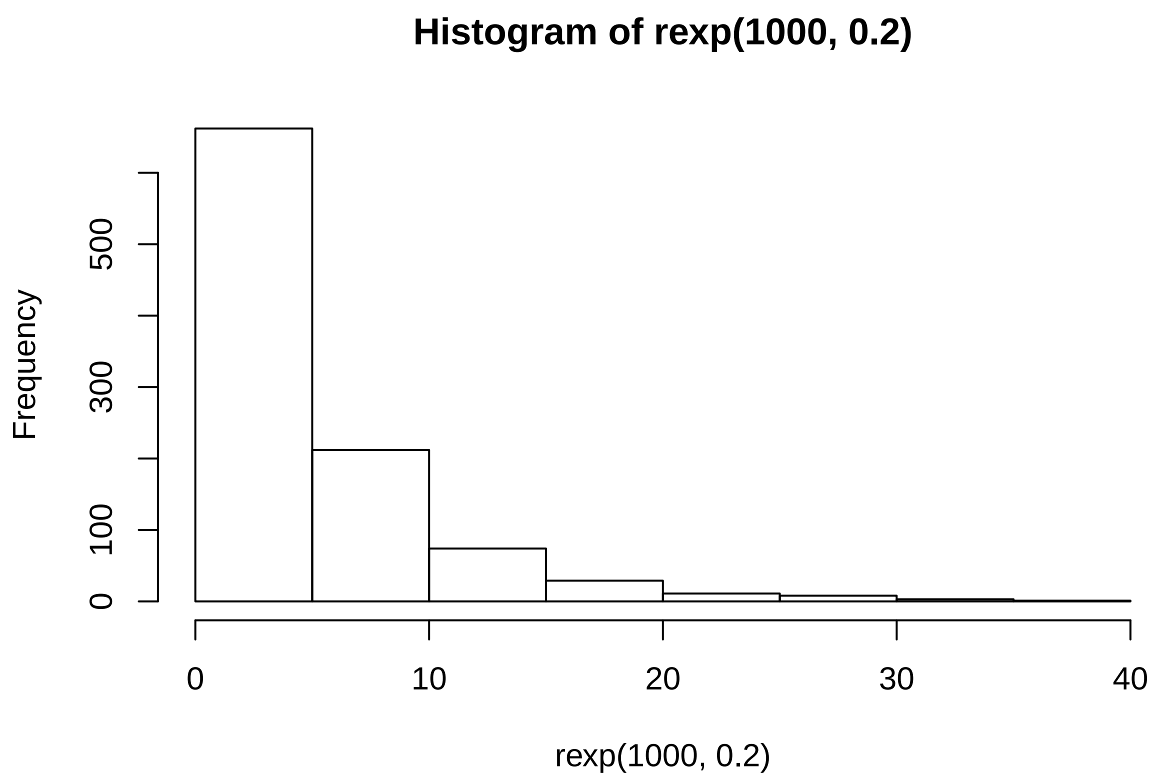
After taking the mean of all of the variances, we can see that the samples all have around a variance of 25 and standard deviation of 5. Thus our sample and theoretical are extremely close once again.

Part 3

Show that the distribution is approximately normal.

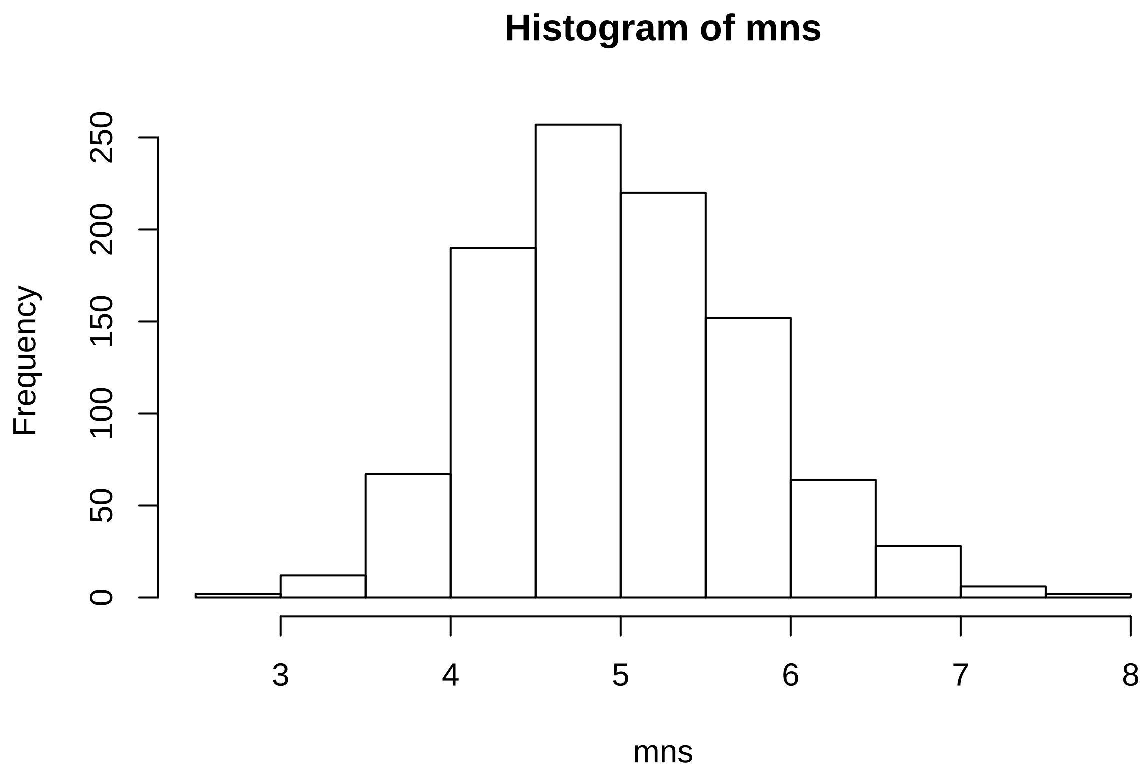
First, let's take a look at a simple histogram of 1000 random exponentials with $\lambda = .2$.

```
set.seed(2200)
hist(rexp(1000, .2))
```



This doesn't look very normal at all. However, if we take the mean of a large number of simulations using the code below, we find something different.

```
set.seed(2200)
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40, .2)))
hist(mns)
```



In this histogram, we get something that looks much more normal centered at our theoretical mean of 5. So we can see through these examples how the Central Limit Theorem works for an exponential distribution.