Comparison of the Exponential Distribution and the Central Limit Theorem

Saul Lugo

October 20, 2015

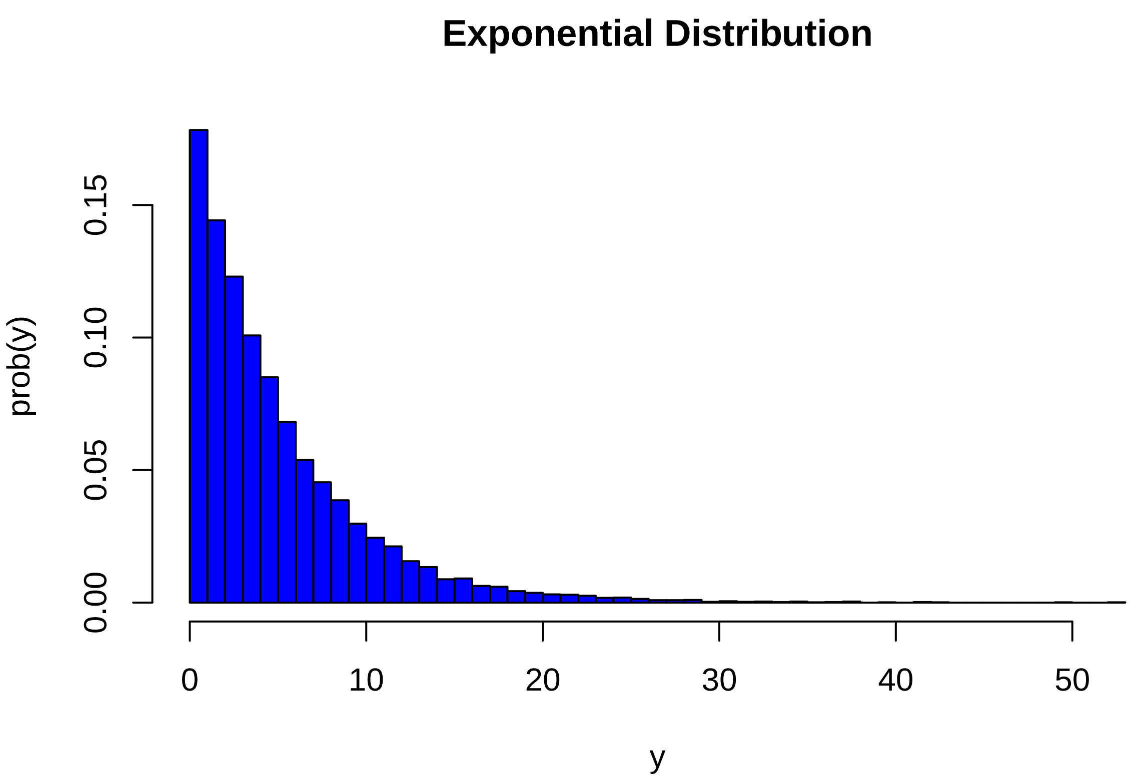
Overview

The purpose of this report is to compare the **Exponential Distribution** to the **Central Limit Theorem**. Using a Simulation, it will be shown that the distribution of averages of n number of random exponential variables becomes that of an standard normal as n sample size increases. This is exactly what the Central Limit Theorem states.

Simulations

To study the exponential distribution, let's first compare the distribution of **10000** random exponentials with the distribution of 40 averages of random exponentials:

```
set.seed(5)
lambda = 0.2
n = 10000
y <- rexp(n, rate = lambda)
hist(y,prob=TRUE,ylab="prob(y)",col="blue",breaks=50,main="Exponential Distribution")
```



```
mean_y <- round(mean(y),2)
mean_y
```

```
## [1] 5.02
```

```
sd_y <- round(sd(y),2)
sd_y
```

```
## [1] 5.05
```

So, it is known that the theoretical mean of the exponential distribution is equal to $\frac{1}{\lambda}$:

- In this example, lambda $= 0.2$, so $\frac{1}{\lambda}$ $\mathbf{= 5}$
- As it is shown in the previous R code chunk, the mean of these **10^{4}** exponentials is **5.02**

Also, the theoretical standard deviation of the exponential distribution is equal to $\frac{1}{\lambda}$:

- As it is shown in the previous R code chunk, the standard deviation of these **10^{4}** exponentials is **5.05**

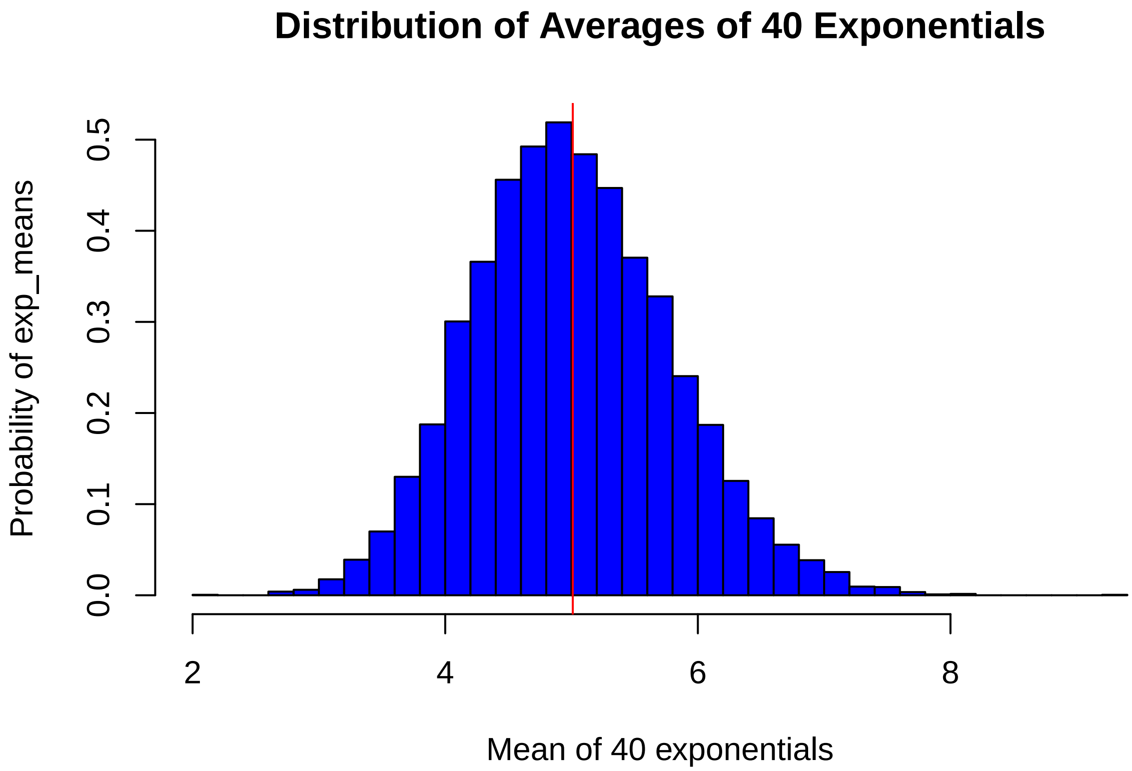
Now let's simulate the random distribution resulted from the calculation of **10^{4}** averages of 40 exponentials:

```
exp_means = NULL

for(i in 1 : n) exp_means = c(exp_means,mean(rexp(40, rate = lambda)))
sample_mean = round(mean(exp_means),2)
sample_mean
```

```
## [1] 5.01
```

```
hist(exp_means,prob=TRUE,xlab="Mean of 40 exponentials",ylab="Probability of exp_means",col="blue",breal
abline(v=sample_mean,col="red")
```




```
# a red line is added to the histogram at the value of the sample_mean
```

Sample Mean vs Theoretical Mean in the Exponential Distribution

So, as it is shown in the previous R code chunck:

- The sample mean (the mean of the vector of averages of 40 exponentials) is equal to **5.01**
- The theoretical mean of the Exponential Distribution is $\frac{1}{\lambda}$, for lambda = 0.2, the theoretical mean is equal to **5**
- Also, in the R code chuck **"exp_dist"** it is shown that the mean of **10^{4}** exponentials is **5.02**

Sample Variance vs Theoretical

The theoretical variance of the Sample is equal to $\frac{\sigma^2}{n}$. In the **Exponential Distribution** $\sigma = \frac{1}{\lambda}$, so we have that the **theoretical variance of the sample is:**

$$var(sample) = \frac{\frac{1}{\lambda}^2}{n} = \frac{1}{n\lambda^2} = \frac{1}{40 * 0.2^2} = 0.625$$

Now, we can also calculate the variance of the variable **exp_means** that is the vector of means of 40 exponentials that we simulated in the R code chuck **"exp_dist_simulation"**:

```
var_exp_means = round(var(exp_means),4)
var_exp_means
```

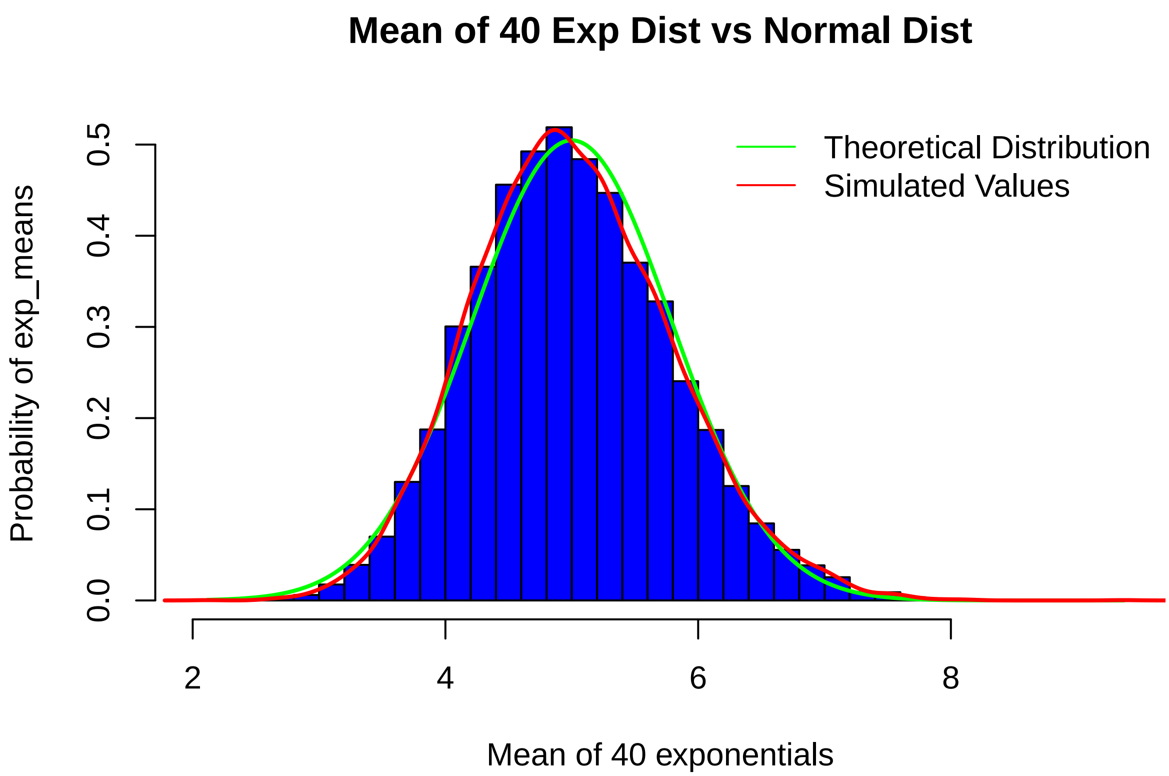
```
## [1] 0.6196
```

So, as shown in the previous R code chunk, the variance of the means of 40 exponentials is equal to **0.6196**, that is pretty close to the theoretical variance that we got of **0.625**.

Verifying if the Distribution of the Sample of Averages of 40 Exponential is Approximately a Normal Distribution

In order to verify if the distribution of the vector or averages of 40 exponentials is close to a **Normal Distribution**, let's first calculate and graph a normal distribution with the same mean and standard deviation of that of the vector of averages of exponentials:

```
x_norm = seq(min(exp_means),max(exp_means),length=n)
y_norm = dnorm(x_norm, mean = 1/lambda, sd = 1/lambda/sqrt(40))
hist(exp_means,prob=TRUE,xlab="Mean of 40 exponentials",ylab="Probability of exp_means",col="blue",brea
lines(x_norm,y_norm,pch=22,col="green",lwd=2)
lines(density(exp_means), pch=22,col="red",lwd=2)
legend('topright',c("Theoretical Distribution","Simulated Values"),
       bty="n",lty=c(1),col=c("green","red"))
```



So, as we can see in the graph, the distribution of the averages of exponentials converged to the normal distribution. In other words, **the distribution of averages of exponentials is approximately a Normal Distribution.**