Exponential Distribution and Central Limit Theorem

Jenhau Chen

February 15, 2016

# Overview

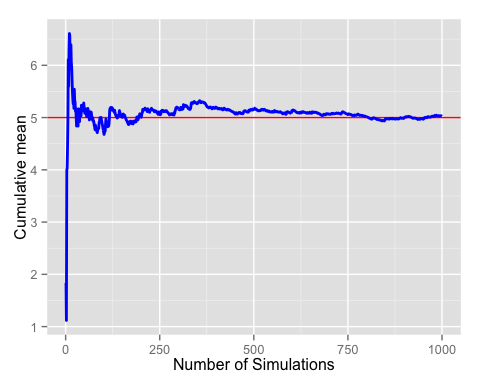
This is a simulation to demonstrate Central Limit Theorem by using Exponential Distribution, which has the following properties:

* **PDF**
* **CDF**
* **Mean**
* **Variance**
* **Standard Deviation**

# Simulations

I will be setting with observations in each simulation. simulations will be executed. According to the Law of Large Numbers, or LLN, the mean should converge to the theorectical mean.

# simulation parameters  
# number of observations  
n = 40  
# number of simulations  
sim\_n = 1000  
# rate  
lambda = 0.2  
  
# Theoretical mean  
mu = lambda^(-1)  
  
# Theoretical variance  
sigma\_sq = lambda^(-2)  
  
# Theoretical standard deviation  
sigma = lambda^(-1)  
  
# Standard Error  
se = lambda^(-1)/sqrt(n)  
  
# set the seed so our random generation is reproducible  
set.seed(40)  
  
mean\_distribution = NULL  
var\_distribution = NULL  
sd\_distribution = NULL  
clt\_distribution = NULL

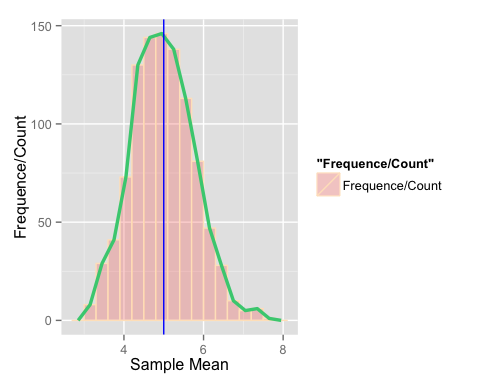


Now start the simulation for mean and variance. Since I have controlled the seed, **rexp** will always generate the same set of random variables and the results are comparable. If the seed isn't controlled, each **rexp** call will gerenrate its own random variables and the **mean** and **variance** will have nothing to do with each other.

for(i in 1: sim\_n) {  
 mean\_distribution = c(mean\_distribution, mean(rexp(n, lambda)))  
}  
  
mean\_df = data.frame(x = mean\_distribution)  
  
for(i in 1: sim\_n) {  
 var\_distribution = c(var\_distribution, var(rexp(n, lambda)))  
}  
  
var\_df = data.frame(x = var\_distribution)  
  
for(i in 1: sim\_n) {  
 sd\_distribution = c(sd\_distribution, sd(rexp(n, lambda)))  
}  
  
sd\_df = data.frame(x = sd\_distribution)

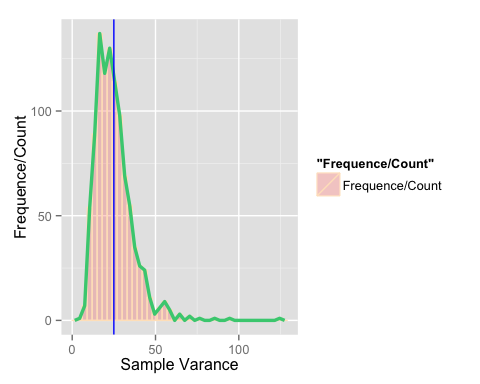
# Sample Mean vs Theoretical Mean

When comparing with the theoretical mean, , the following plot shows the sample mean converging to the theoretical mean:

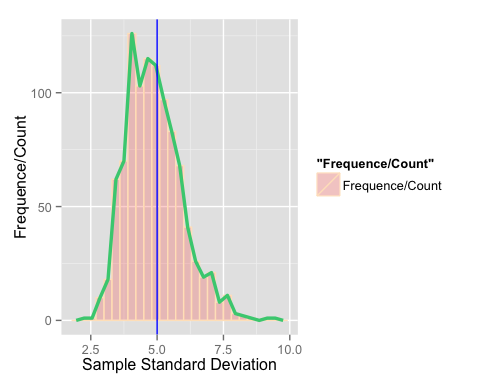


# Sample Variance vs Theoretical Variance

Similiarly when comparing with the theoretical variance, , the following plot shows the sample variance converging to the theoretical variance:



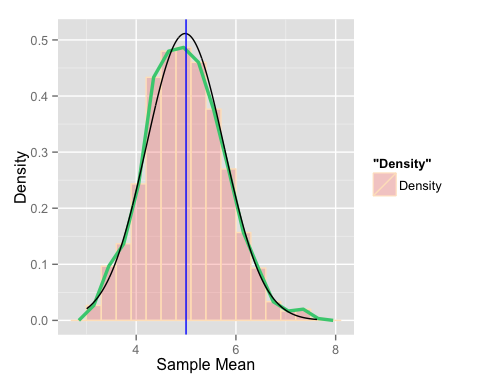
However notice the long tail on the right. I believe it is due to how variance is calculated. If I use standard deviation instead, the plot becomes more Gaussian:



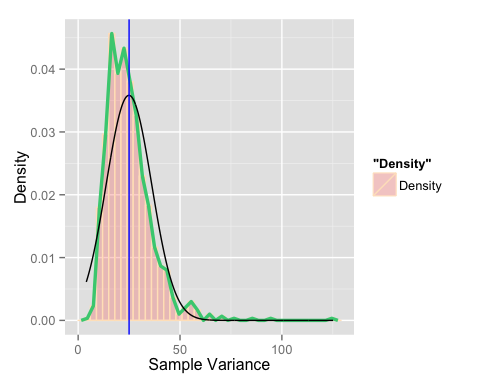
# Distribution

According to Central Limit Theorem, CLT, I am expecting the reuslt closely match with a standard normal distribution. I demonstrated this point using a density plot:

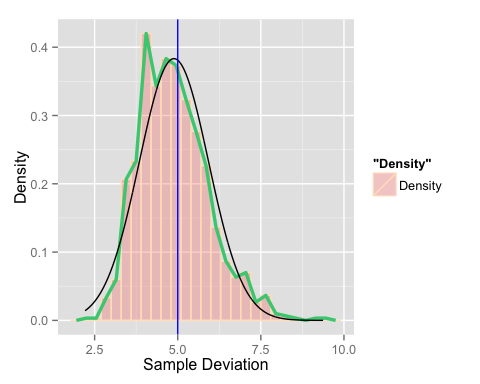
## Sample Mean vs. Normal Density



## Sample Variance vs. Normal Density



## Sample Standard Deviation vs. Normal Density

Again I don't like seeing the long tail so I am using standard deviation distribution to validate CLT: 

## Normalized Distribution

Lastly according to CLT, I am expecting

has a distribution like that of a standard normal for large .

Playing with the similuation data again, I have the following plot which validates this point:

for(i in 1: sim\_n) {  
 clt\_distribution = c(clt\_distribution, (mean\_distribution[i] - mu)/se)  
}  
clt\_df = data.frame(x = clt\_distribution)

