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:

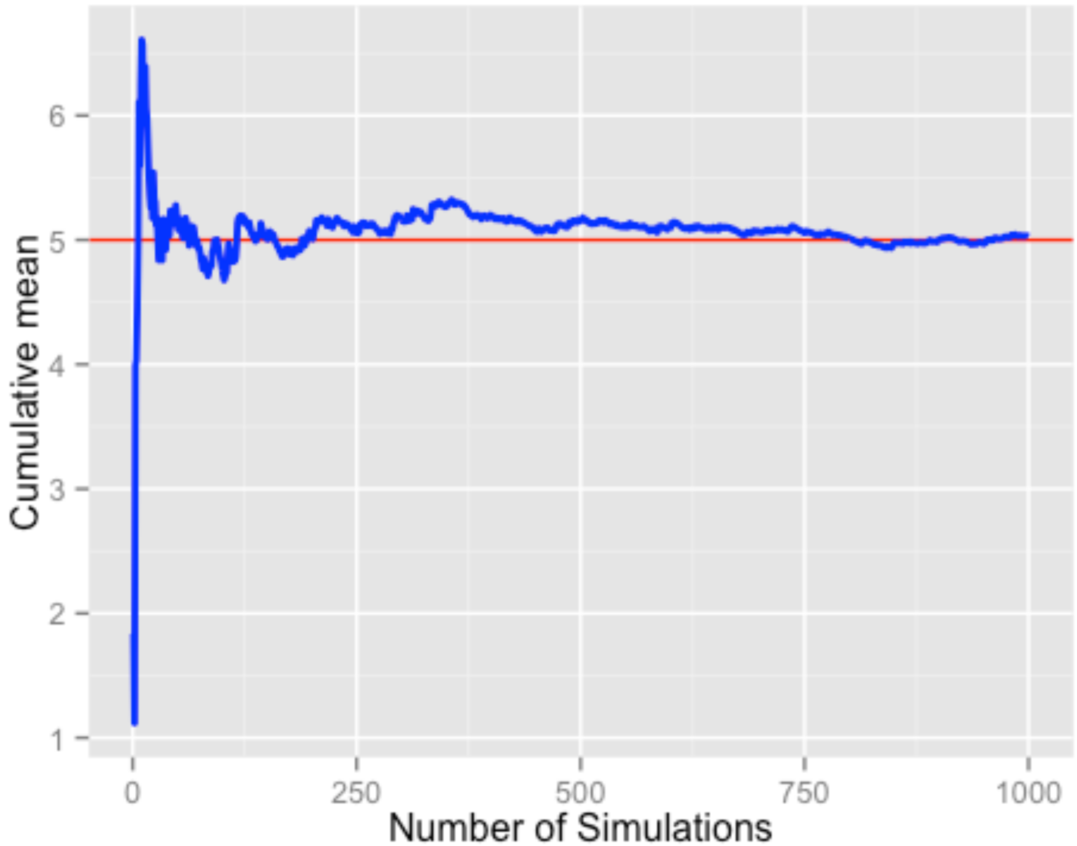
- **PDF** $\lambda e^{-\lambda x}$
- **CDF** $1 - e^{-\lambda x}$
- **Mean** λ^{-1}
- **Variance** λ^{-2}
- **Standard Deviation** λ^{-1}

Simulations

I will be setting $\lambda = 0.2$ with 40 observations in each simulation. 1000 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
```



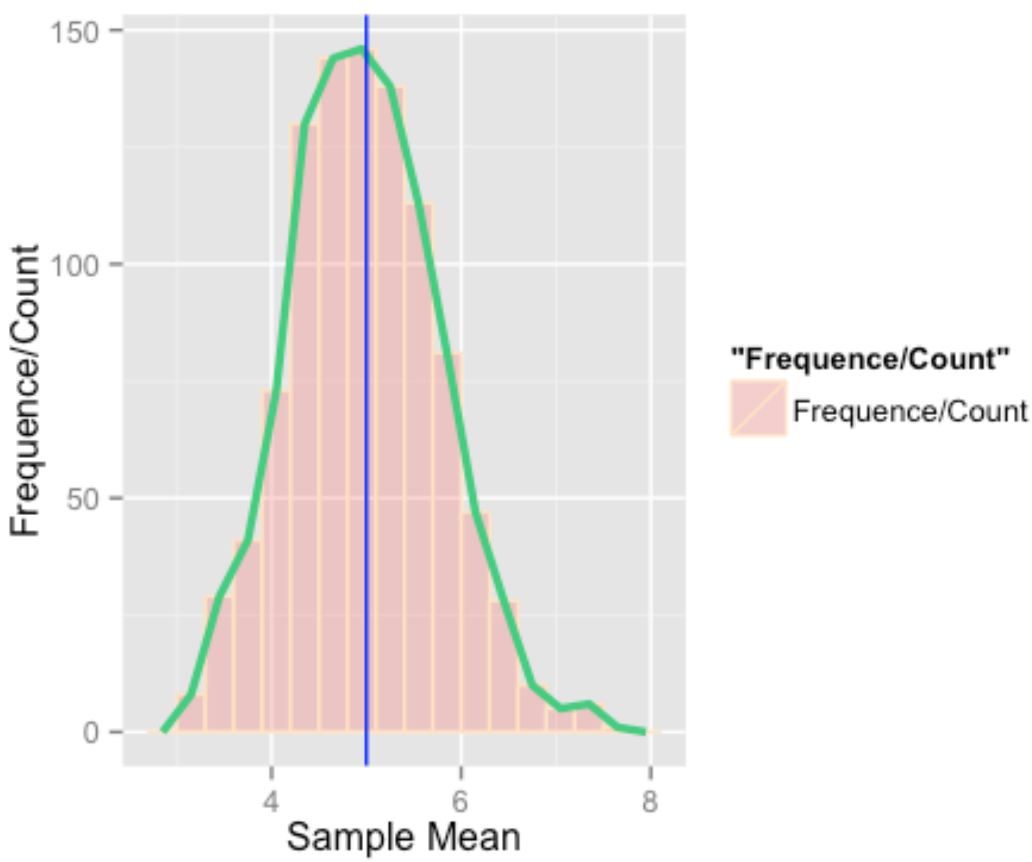
Now start the simulation for mean and variance. Since I have controlled the seed, **rexp** will always generate the same set of random variables and the results are comparable. If the seed isn't controlled, each **rexp** call will gerenrate its own random variables and the **mean** and **variance** will 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)
```

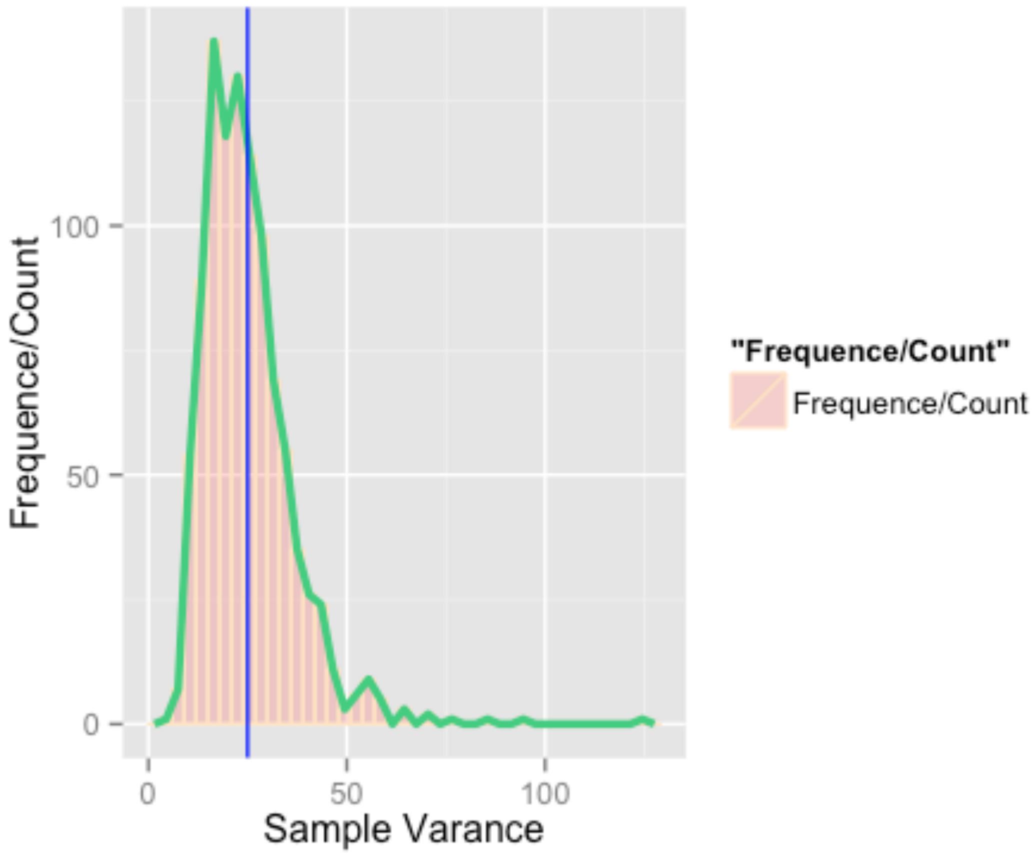
Sample Mean vs Theoretical Mean

When comparing with the theoretical mean, $\lambda^{-1} = 1/0.2 = 5$, the following plot shows the sample mean converging to the theoretical mean:

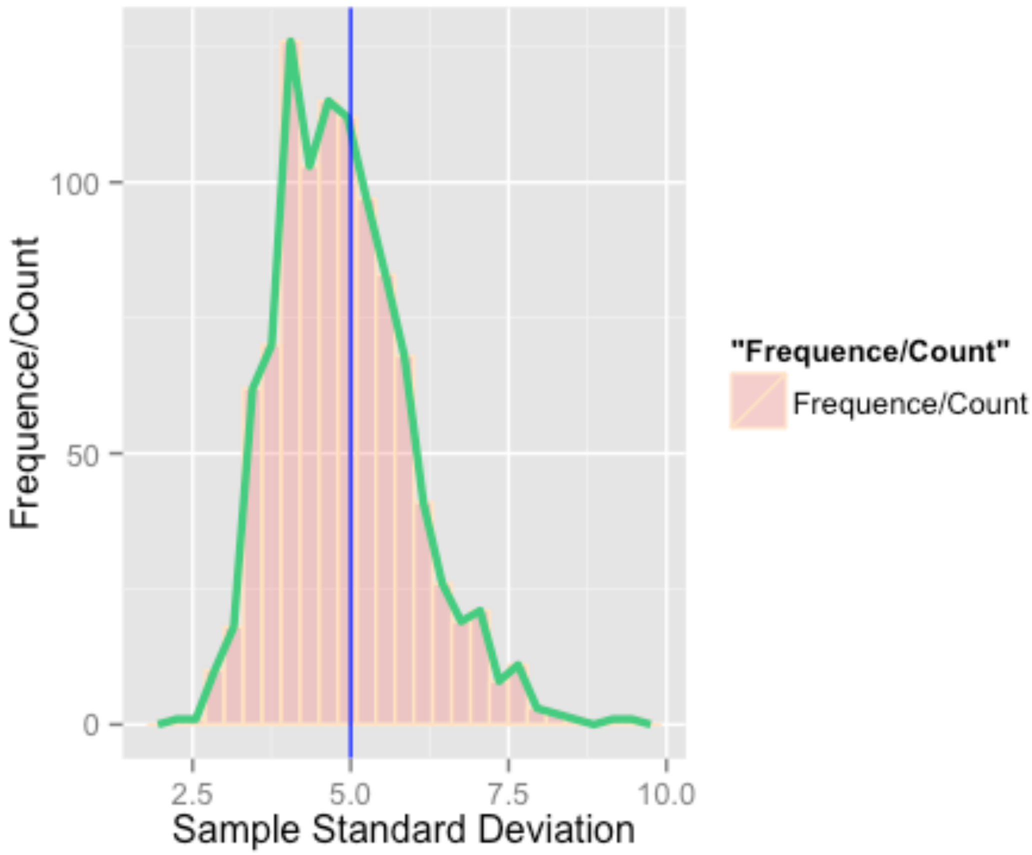


Sample Variance vs Theoretical Variance

Similiarly when comparing with the theoretical variance, $\lambda^{-2} = (1/0.2)^2 = 25$, the following plot shows the sample variance converging to the theoretical variance:



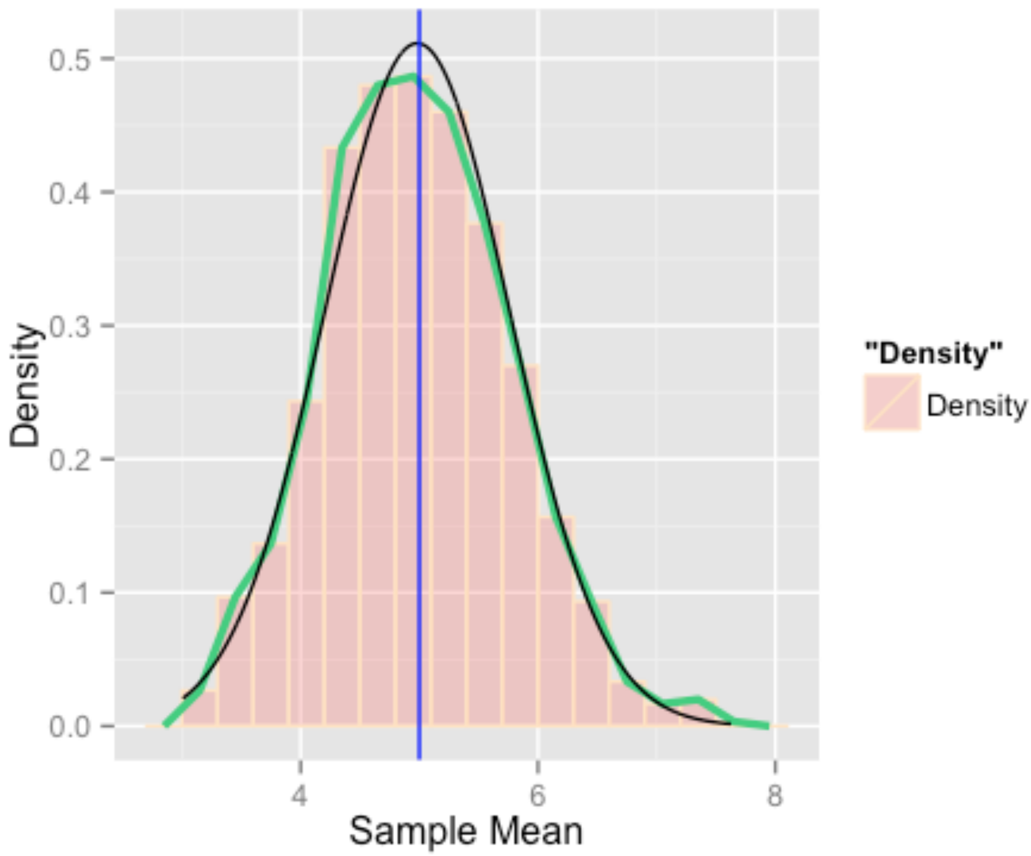
However notice the long tail on the right. I believe it is due to how variance is calculated. If I use standard deviation instead, the plot becomes more Gaussian:



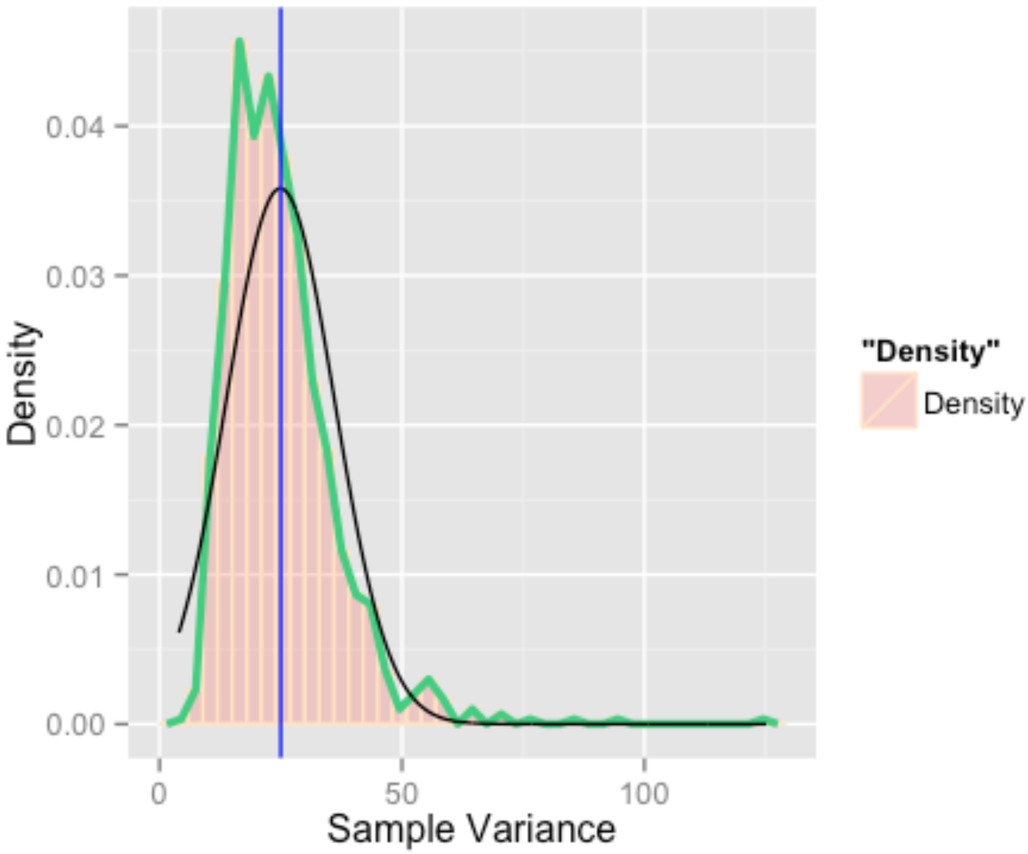
Distribution

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

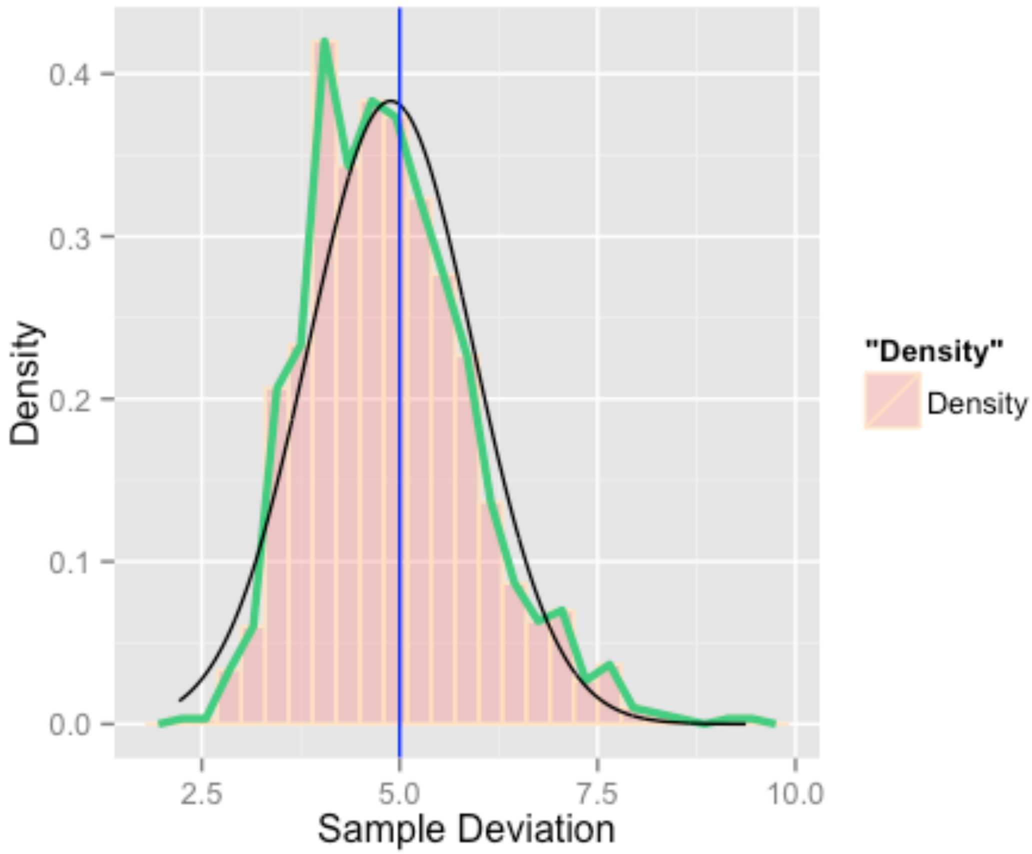


Sample Variance vs. Normal Density



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

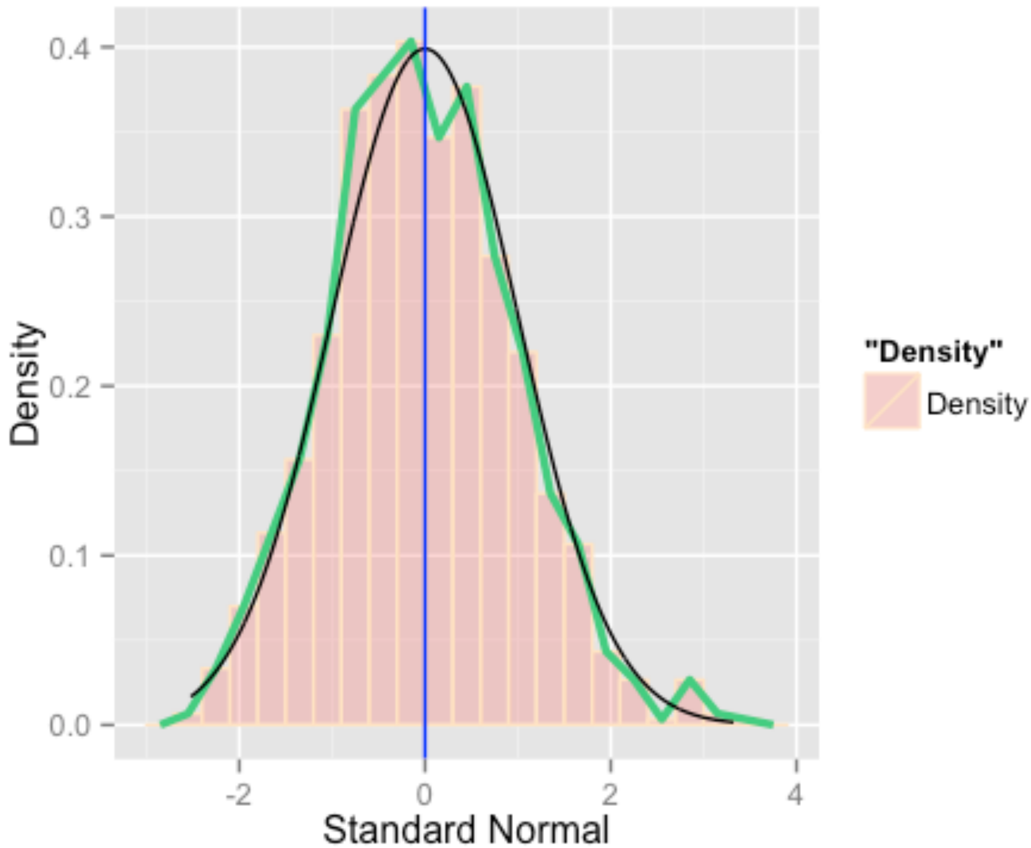
Lastly according to CLT, I am expecting

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{\text{Estimate} - \text{Meanofestimate}}{\text{Std. Err. ofestimate}}$$

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

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)
```



0.4
0.3
0.2
0.1
0.0
Density
-2
0
2
4
Standard Normal
"Density"
Density