# Course Project Part 1

*Jeffrey Uslan*

*Monday, July 20, 2015*

```
library(ggplot2)
```

In this project I will explroe the central limit theorem. I will do so by comparing 1000 simulations of 40 random exponentials variables to the theoretical exponential distribution. ##Setting up simulations

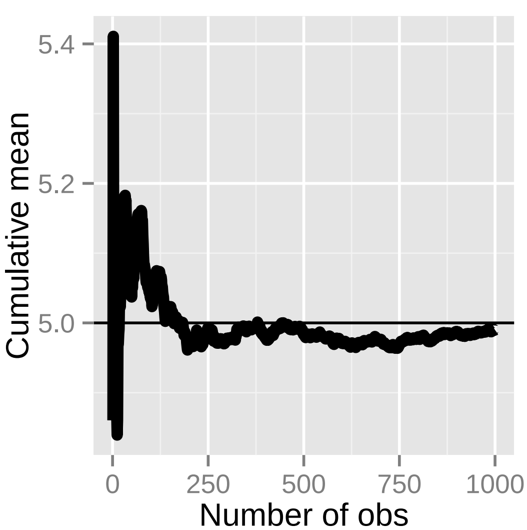
```
set.seed(1)
lambda=.2
n <- 1000
dists=NULL
for (i in 1:n){
  dists=cbind(dists,rexp(40,lambda)  )

}
```

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

```
means=apply(dists,2,mean)
cum_means <- cumsum(means)/(1:n)

g <- ggplot(data.frame(x = 1:n, y = cum_means), aes(x = x, y = y))
g <- g + geom_hline(yintercept = 1/lambda) + geom_line(size = 2)
g <- g + labs(x = "Number of obs", y = "Cumulative mean")
g
```




##Theoretical mean

```
1/lambda
```

```
## [1] 5
```

### Mean of simulation

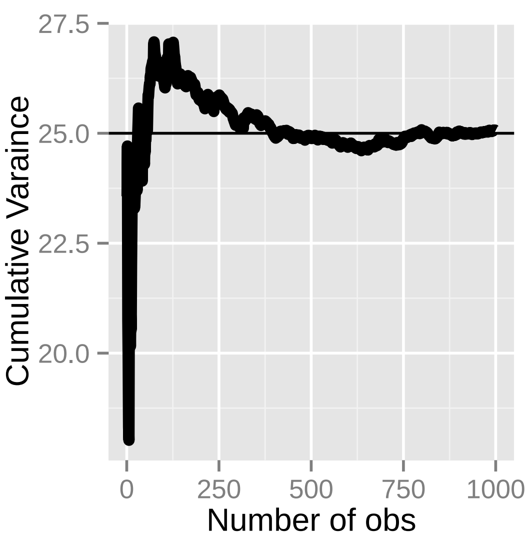
```
mean(means)
```

```
## [1] 4.990025
```

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

```
var=apply(dists,2,var)
cum_var <- cumsum(var)/(1:n)


g <- ggplot(data.frame(x = 1:n, y = cum_var), aes(x = x, y = y))
g <- g + geom_hline(yintercept = (1/lambda)^2) + geom_line(size = 2)
g <- g + labs(x = "Number of obs", y = "Cumulative Varaince")
g
```




##Theoretical variance

```
(1/lambda)^2
```

```
## [1] 25
```

### Mean of simulation

```
mean(cum_var)
```

```
## [1] 25.14339
```

## 3. Show that the distribution is approximately normal.

Creating a normal distrubtion using the borrowed parameters

```
normals=rnorm(n,mean=mean(means),sd=sd(means))
```

See how closely the normal ditribution lines up with the simulations

```
ggplot(data=data.frame(x = 1:n, mean = means,normals=normals))+theme_bw()+
  geom_density(aes(x=means),color="blue")+geom_density(aes(x=normals),color="red")
```