Statistical Inference: Exponential Distribution & CLT

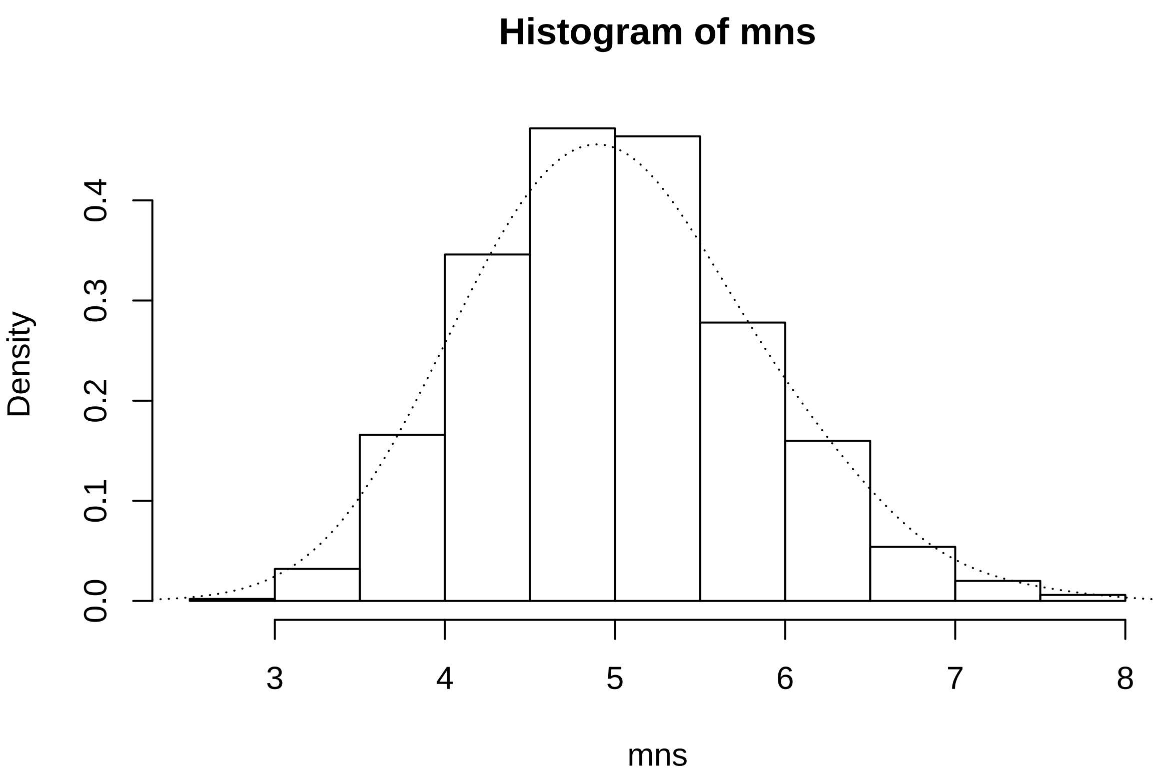
Jonney Liu

In this paper we will investigate the exponential distribution and compare it with the Central Limit Theorem. In particular we will run exponential distribution simulations in R with parameters of n = 40 and lambda = 0.2 to show that the averages of the means will tend to normal distribution with mean 1/lambda and variance (1/lambda)^2/n.

```
#Set variables for exponential distribution
lambda = 0.2
n = 40
```

In the first plot we will run a simulation using a for loop on the means of the exponential distribution 1000 times and plot the means in a probability density function and draw a fitted density line to show that this looks on the surface approximately like a normal distribution.

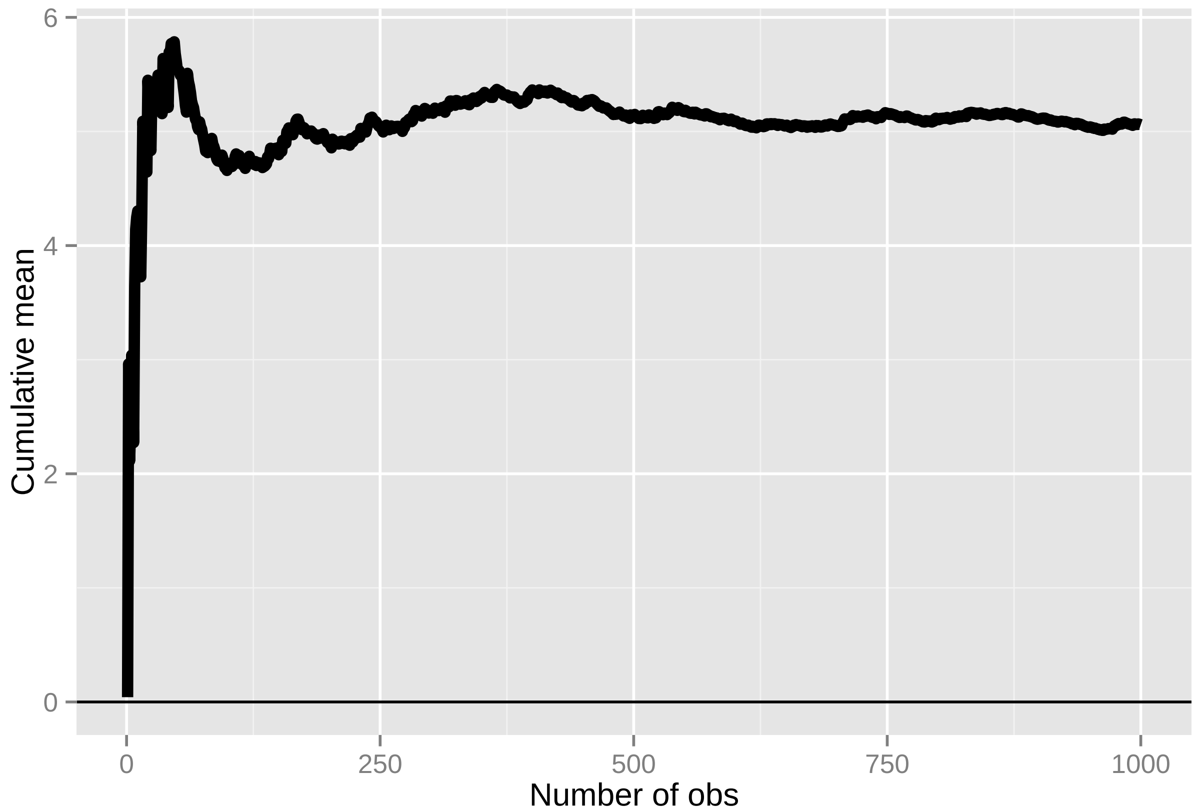
```
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(n,lambda)))
hist(mns,prob=TRUE)
lines(density(mns, adjust=2), lty="dotted")
```



As we can see above that by running the simulation on the means of an exponential distribution and plotting the resulting histogram density, the figure visually becomes much like a normal distribution as we would expect.

If the Law of large numbers apply that means the sample mean from the simulation should approximate to the theoretical mean of the exponential distribution where theoretical mean = 1/lambda. We can show this with the following graph which shows that the sample mean tends to the theoretical mean (in this case 1/lambda = 5) as number of observations increases.

```
## Warning: package 'ggplot2' was built under R version 3.1.2
```




```r
theo_mean <- 1/lambda ##lambda = 0.2
theo_mean
```

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

```r
mean(mns)
```

```
## [1] 5.010076
```

Furthermore, the variance on the simulation by Central Limit Theorem will approximate to the theoretical variance divided by the sample size n = 40, where theoretical variance = (1/lambda)^2

```r
theo_var <- (1/lambda)^2
theo_sam_var <- theo_var/n #n = 40 (the number of samples)
theo_sam_var
```

```
## [1] 0.625
```

Compare this to the variance of the simulation

```r
var(mns)
```

```
## [1] 0.65048
```

So far the sample mean and variance on the simulation have come out very close to our perdicted values with n = 40 and lambda = 0.2 for the exponential distribution. We can show with the final graph below that the probability density function of our simulation should be well fitted by a normal distribution with mean = 1/lambda (5) and variance = (1/lambda)^2 divided by n.

This time we will run the simulation 10000 to get a better picture

```
mns = NULL
for (i in 1 : 10000) mns = c(mns, mean(rexp(n,lambda)))
hist(mns,prob=TRUE,ylim=c(0, 0.7))
m <- 1/lambda
ssdev <- 1/lambda/sqrt(n)
curve(dnorm(x, mean=m, sd=ssdev),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```