Project 1

Simulation

```
n=40
lambda=0.2
simulation=1000
set.seed(77)
data=matrix(rexp(n*simulation,lambda),simulation)
row_means=apply(data,1,mean)
act_mean=mean(row_means)
act_sd=sd(row_means)
act_var=var(row_means)
exp_mean=1/lambda
exp_sd=((1/lambda)*(1/sqrt(n)))
exp_var=exp_sd^2
act_mean
## [1] 4.996833
exp_mean
## [1] 5
act_sd
## [1] 0.7941483
exp_sd
## [1] 0.7905694
act_var
## [1] 0.6306716
exp_var
## [1] 0.625
```

Question 1 - Show where the distribution is centered at and compare it to the theoretical center of the distribution

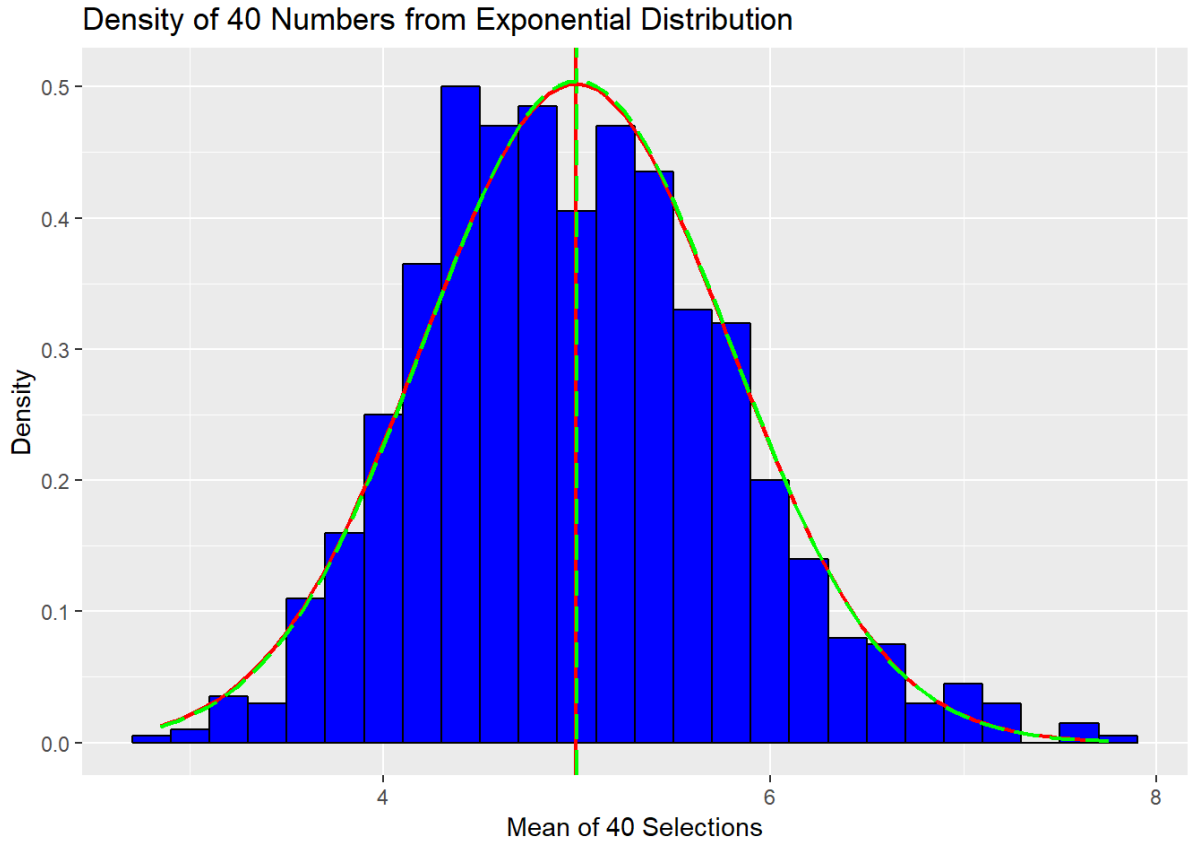
From the simulation result, we get the distribution is centered at 4.996833, while the theoretical distribuion is centered at 5.

Question 2 - Show how variable it is and compare it to the theoretical variance of the distribution

From the simulation result, we get the variance is 0.6306716, while the variance of the theoretical distribution is 0.625.

Question 3 - Show that the distribution is approximately normal

```
library(ggplot2)

data_means=data.frame(row_means)

g=ggplot(data_means, aes(x=row_means))

g+geom_histogram(binwidth=lambda, fill="blue", color="black", aes(y=..density..))+

    labs(title="Density of 40 Numbers from Exponential Distribution",

         x="Mean of 40 Selections",y="Density")+

    geom_vline(xintercept=act_mean, size=0.8,color="red")+

    geom_vline(xintercept=exp_mean, size=0.8,color="green",linetype="longdash")+

    stat_function(fun=dnorm,args=list(mean=act_mean,sd=act_sd),color="red",size=0.8)+

    stat_function(fun=dnorm,args=list(mean=exp_mean,sd=exp_sd),color="green",size=0.8,

                  linetype="longdash")
```



The plot shows that the distribution is approximately normal. The red line is actual distribution and the green line is theoretical distribution. And both centers of the distributions seems no difference.Therefore, the Central Limit Theory is approved.