Statistical inference Course Project

Reshu

1 June 2017

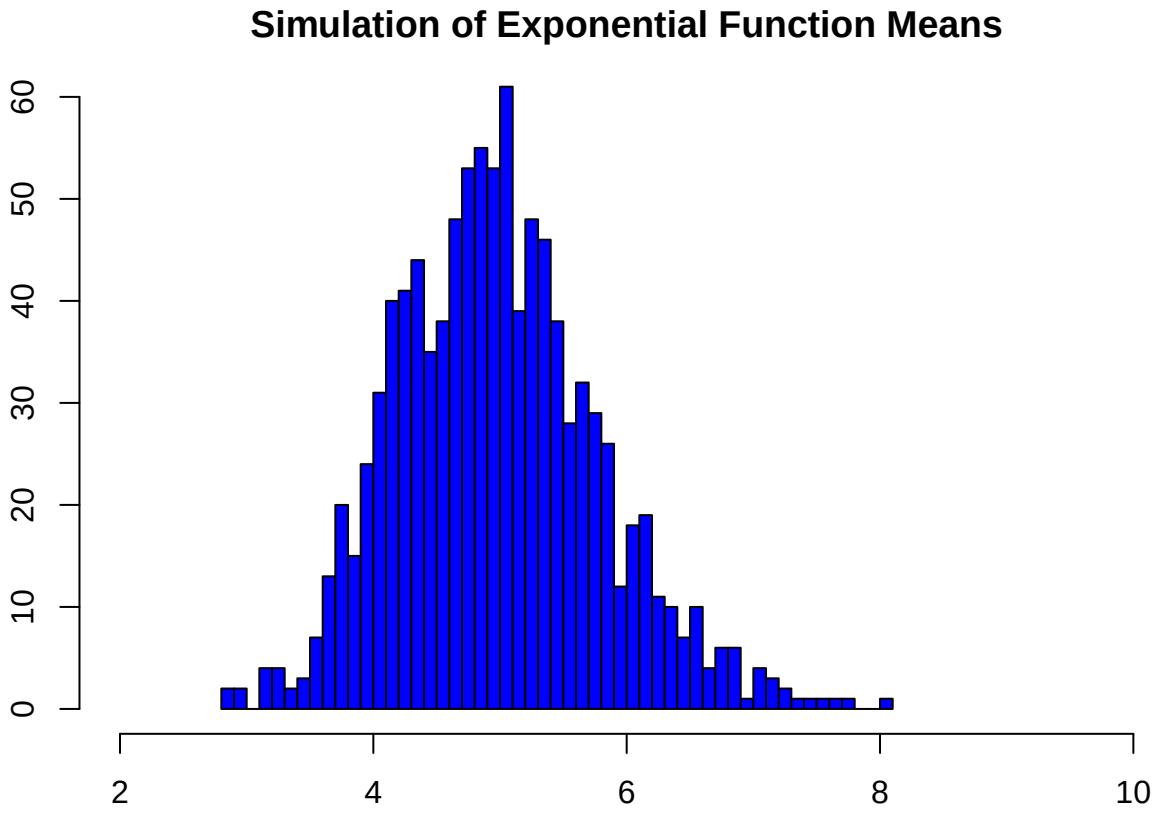
Overview

In this project we will investigate Exponential distribution in R and compare it with Central Limit Theorem.We will simulate the disribution using rexp(n,lambda) function (where lambda is rate parameter) and will illustrate the properties of the distribution of the mean of 40 exponentials.

Simulations

We will simulate the exponential distribution in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda.We consider the following guidelines for simulation 1-lambda = 0.2 for all of the simulations 2-We will investigate the distribution of averages of 40 exponentials 3-We will do a 1000 simulations. According to the above guidelines, we set parameters in R

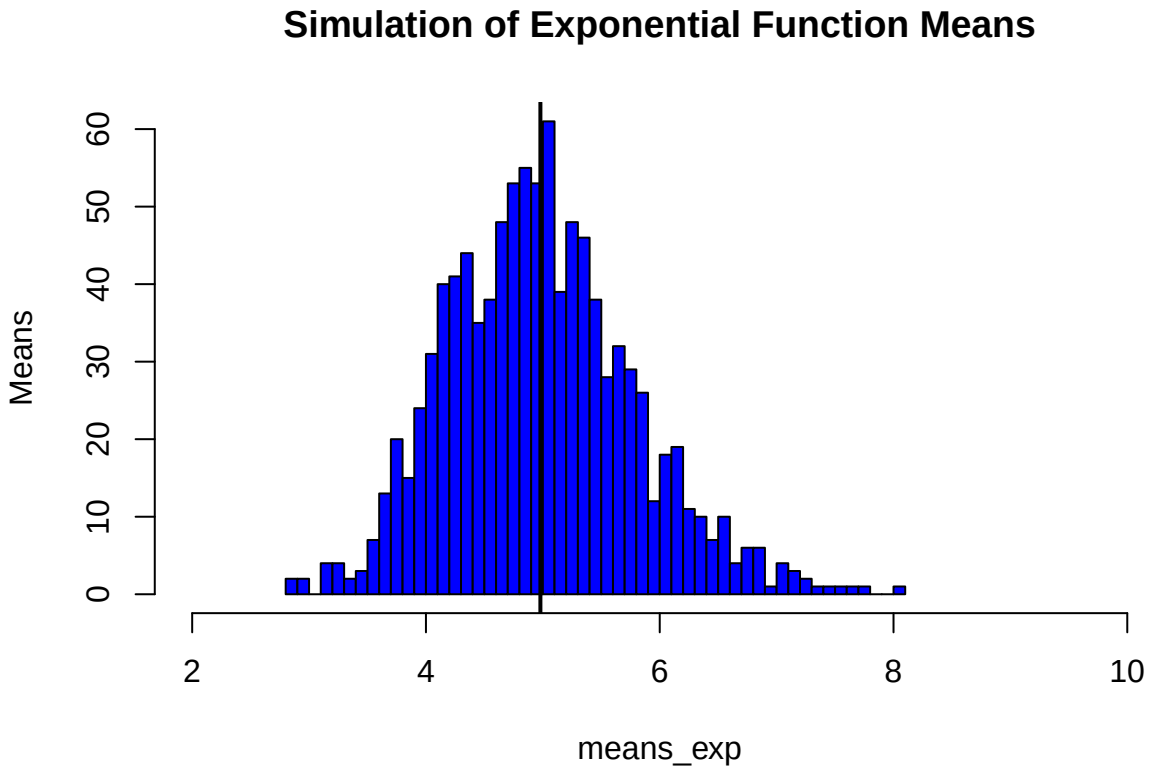
```
setwd("~/Documents/Stats_inference_project")
set.seed(2016) #Set seed for making the results reproducible
#Set parameters as given in the instructions
lambda <-0.2 #lambda
n<-40        #no. of exponentials
nosim<-1000  #no. of simulations
coverage<-replicate(nosim,rexp(n,lambda)) #Create matrix with simulated data
means_exp<-apply(coverage,2,mean) #Calculate means of exponential simulations
x11(width=9, height=6, pointsize=12)
par(mar=c(3,2,2,2))
hist(means_exp,breaks=60,xlim=c(2,10),col='blue',main='Simulation of Exponential Function Means',ylab='
```



Question 1-Comparison of sample mean with theoretical mean of distribution

Theoretical Mean-Mean is 1/Lambda.Since lamda is 0.2 ,theoretical mean should be 5(1/lambda=1/0.2).Sample Mean-We draw a histogram showing means of sample and draw vertical line through the mean of the sample as below

```
hist(means_exp,breaks=60,xlim=c(2,10),col='blue',main='Simulation of Exponential Function Means',ylab='
abline(v=mean(means_exp),lwd=2) #Draw vertical line through mean
```




```
mean(means_exp) #Following is the mean
```

```
## [1] 4.979186
```

The code above shows that sample mean is 4.979 which is close to theoretical mean 5

Question 2- Comparison of sample variance with theoretical variance of the distribution

Theoretical variance-Standard deviation of distribution is 1/lambda .Standard distribution of sample is ((1/lambda)/sqrt (40)) .Variance is ((1/lambda)/sqrt (40))^2.

```
#theoretical variance
var_theory<-((1/0.2)/sqrt(40))^2
print(paste("Theoretical standard deviation ",round(var_theory,4)))
```

```
## [1] "Theoretical standard deviation  0.625"
```

```
var_sample<-var(means_exp) #Sample variance
print(paste("Sample standard deviation ",round(var_sample,4)))
```

```
## [1] "Sample standard deviation  0.6385"
```

```
var_sample
```

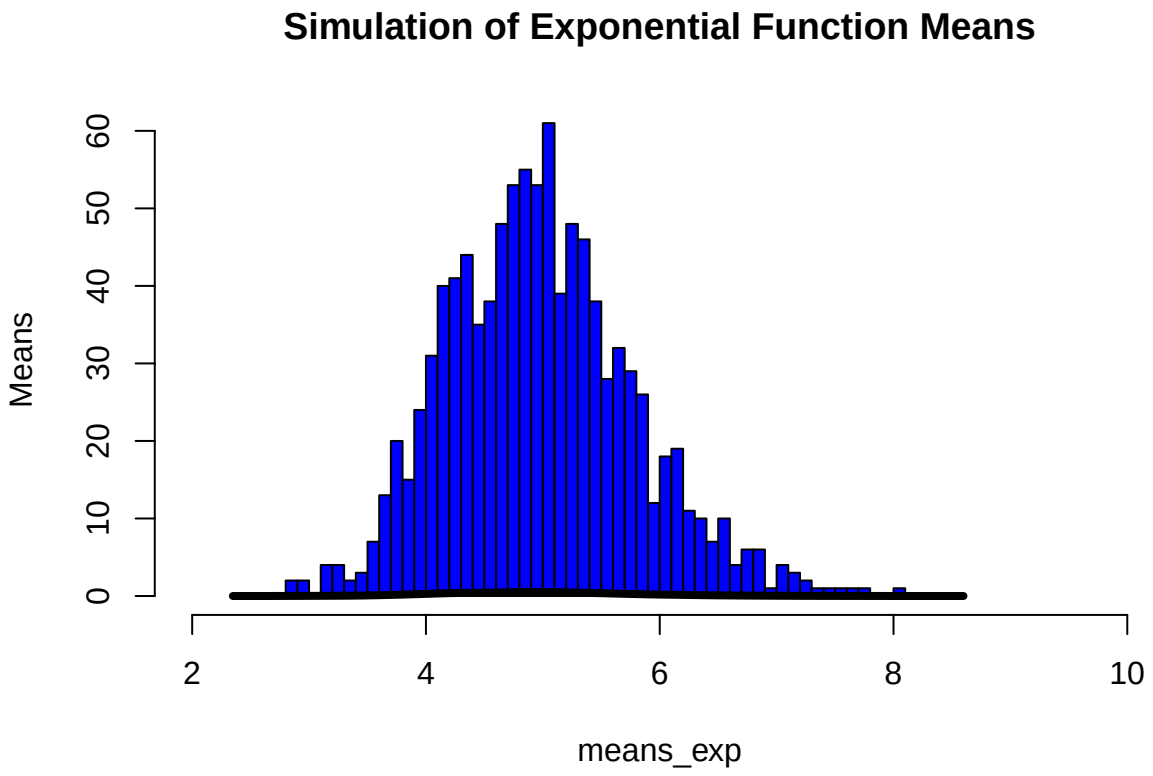
```
## [1] 0.6384844
```

Hence sample variance is close to theoretical variance

Question 3- Illustration if distribution is normal

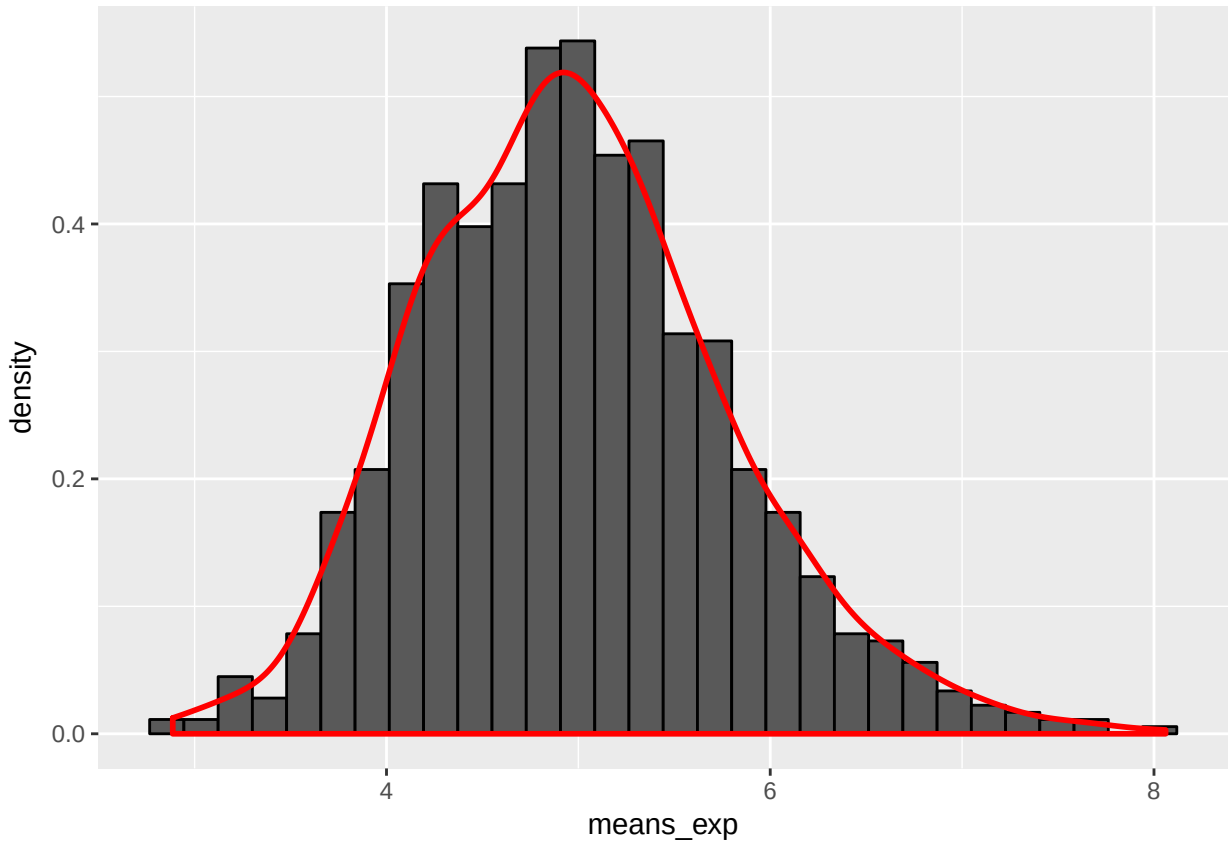
Due to central limit theorem , the means of sample should follow a normal distribution

```
hist(means_exp,breaks=60,xlim=c(2,10),col='blue',main='Simulation of Exponential Function Means',ylab='
lines(density(means_exp),lwd=4,col="black")
```




```
pdata<-data.frame(means_exp)
m<-ggplot(pdata,aes(x=means_exp))+geom_histogram(aes(y=..density..),color="black")+geom_density(color="
print (m)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



As per the above figure, the distribution of the means of sample seems to follow normal distribution .As we increase our number of samples, the distribution is more close to normal distribution.Dotted line is normal distribution curve and it is very close to sample distribution,which is black line in the above figure.