Statistical Inference: Project 1

Mrugank Akarte

27 February 2016

Overview

In the following report, exponential distribution was investigated using R. It was then compared with *The Central Limit theorem*. The data was simulated in R using rexp(n, lambda) where lambda is the rate parameter.

Given theoretical mean and theoretical standard deviation of the distibution as (1/lambda). Also it is mentioned that lambda =0.2 for all simulations.

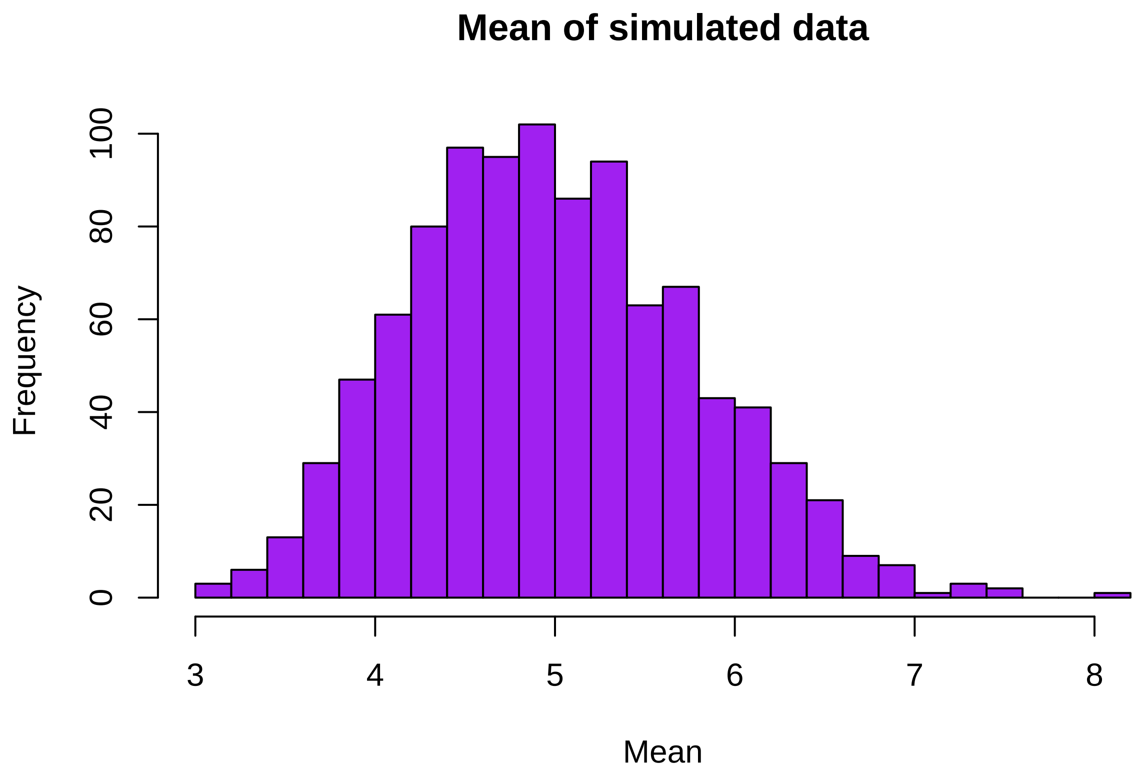
Simulations

Setting up parameters

```
lambda=0.2
sim<-1000
n<-40
set.seed(1)
```

Creating thousand simulations each with 40 random exponentials and calculating average for each row.

```
x<- matrix(rexp(sim*n,lambda),sim)
x_mean<-rowMeans(x)
hist(x_mean, col = "Purple",breaks = 30,xlab = "Mean",main = "Mean of simulated data")
```



Mean comparison

Calculating sample and theoretical mean of the data.

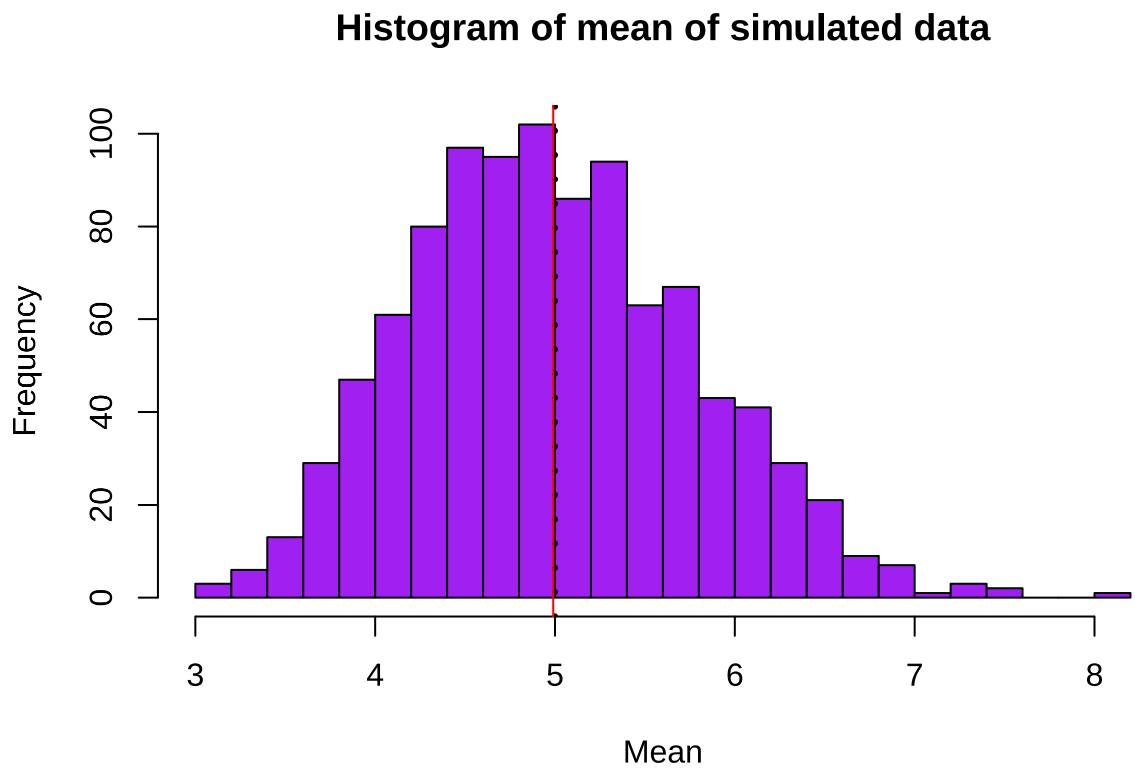
```
theoretical_mean<-(1/lambda)
theoretical_mean
```

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

```
sample_mean<-mean(x_mean)
sample_mean
```

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

```
hist(x_mean, col = "Purple",breaks = 30,xlab = "Mean",main = "Histogram of mean of simulated data")
abline(v = theoretical_mean, lty=3, lwd=3)
abline(v = sample_mean, lwd=1, col="Red")
```



The vertical red line signifies the sample mean of the data where as the black dotted line signifies the theoretical mean.

Thus from above graph and calculations we can see that the sample mean is very close to the theoretical mean.

Variance comparison

Calculating sample and theoretical variance of the given data.

```
std_deviation<-(1/lambda)
theoretical_var<-(std_deviation^2)/n
theoretical_var
```

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

```
sample_var<-var(x_mean)
sample_var
```

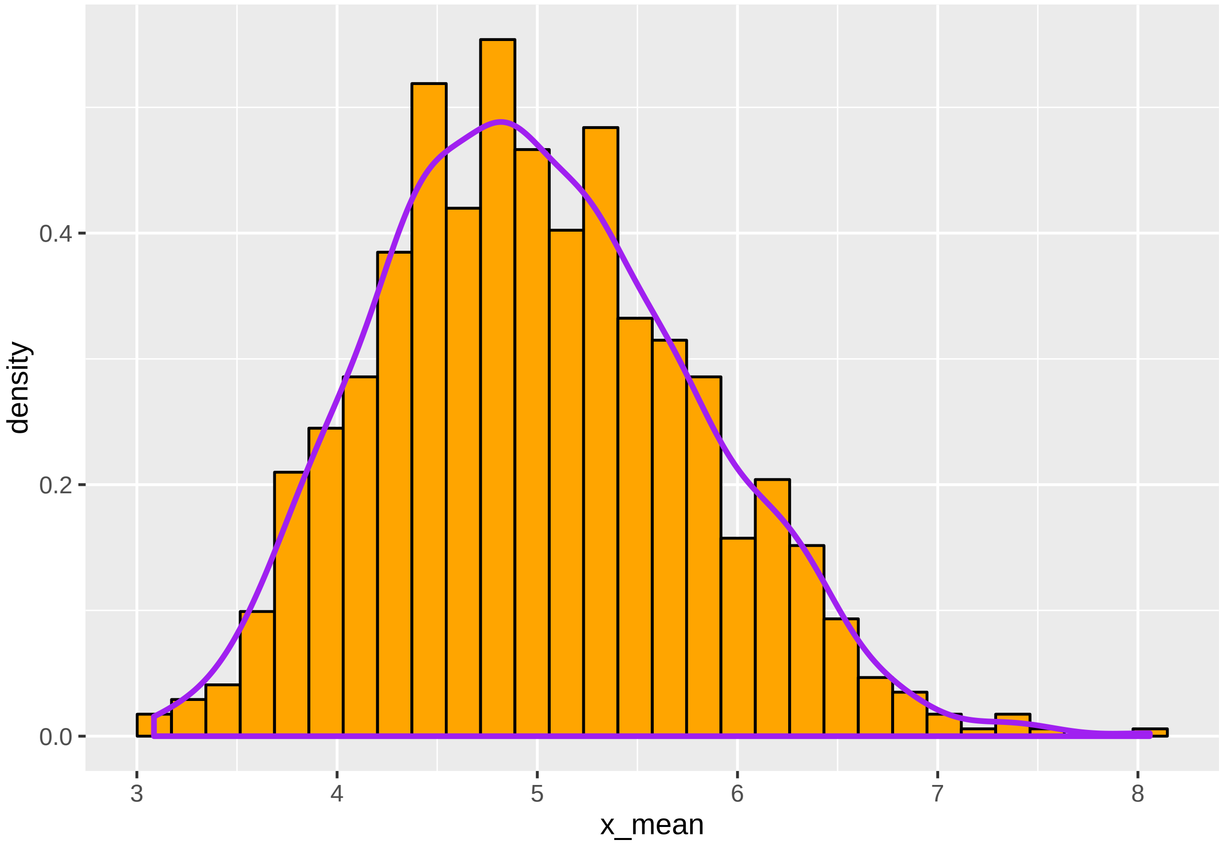
```
## [1] 0.6177072
```

From above we can see that the values of theoretical and sample variance are close to each other.

Approximating Normal distribution

Using graphical method

```
library(ggplot2)
graph1<-ggplot(data.frame(x_mean),aes(x=x_mean,y=..density..))+geom_histogram(colour='black',fill='orang
graph1
```



Comparing confidence intervals

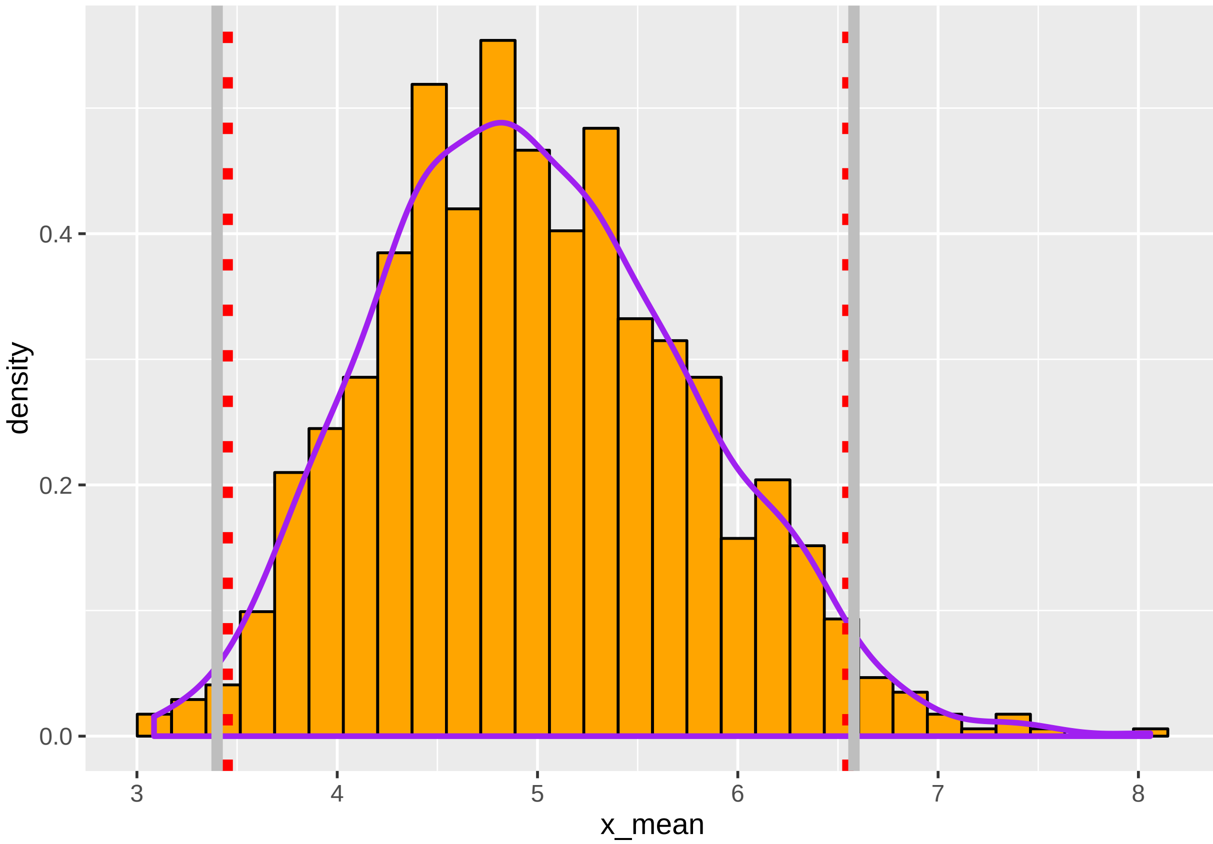
Since 95^{th} percentile lies 1.96 standard deviation from mean, theoretical confidence interval is calculated as follows.

```
theoretical_ci<-theoretical_mean+c(-1,1)*1.96*(std_deviation/sqrt(n))
theoretical_ci
```

```
## [1] 3.450484 6.549516
```

```
sample_ci<-sample_mean+c(-1,1)*qt(.975,(n-1))*sd(x_mean)
sample_ci
```

```
## [1] 3.400304 6.579746
```



We can clearly see that the confidence intervals are almost equal in both the cases.

Hence, from the graph and by comparing the mean, variance and the 95^{th} percentile confidence interval we can conclude that the exponential disrtibution is approximately normal disrtibution for given simulation.