EXPONENTIAL DISTRIBUTION AND ITS COMPARISON WITH CENTRAL LIMIT THEOREM

MALLESH SUBRAMANI M

Overview

This paper concerns with the study of exponential distribution , which is a continous probabality density function.We are mainly concerned with the fitness of the theoretical and experimental values of the exponential distribution.Also the verification of central limit theorem for large number of samples has been performed.

Simulation

The exponential distribution is the probability distribution that describes the time between events in a Poisson point process.

```
exp40_ <- rexp(n=40,rate=0.2) #generating 40 random exp distribution values
exp40_
```

```
##  [1] 10.74959247  1.08245088  0.04577272  0.35084532  0.92900772
##  [6]  2.29551852  0.20170825  9.10159877  0.49508716  0.10950733
## [11]  5.48095667  0.97908622  3.31175156  1.22680343 12.47464481
## [16]  9.66682390  3.74600416 14.40635305  0.20127160  9.30507028
## [21]  2.46945177  5.01222198  5.64056610 16.67933320 10.36193997
## [26]  1.27575449  0.20566630 19.16185290  1.55336745  0.47965631
## [31]  1.25635129  6.81266698  4.76730306  9.04417434  3.63142432
## [36]  1.94200421 12.23418889  8.41252987  1.29688092  9.41761535
```

```
The above data set represents the random generated values from the rexp function, rexp is a p
redefined function in R, which is used to generate random required values of exponential dist
ribution.It takes in two values, that is "n" which is the required number of samples and "rat
e" is the lambda value. The above expression "exp40_"contains 40 random exponential distribut
ion values.
```

```
mean40 <- mean(exp40_)      #finding the mean of the obtained 40 values
mean40
```

```
## [1] 5.19537
```

```
lambda <- rate <- 0.2       #setting the lambda value
mean_th <- 1/rate           #calculating the theoritical mean values
mean_th
```

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

```
mean_th-mean40                #comparing the obtained mean values
```

```
## [1] -0.1953701
```

```
mean_th is the theoritical mean and mean40 is the experimental mean of the random samples obt
ained.

Therefore it can be seen that theoritical mean is approximately equal to the experimental mea
n for less number of samples.
```

```
## [1] 5.133028
```

```
var40 <- std40*std40        #std40 is the experimentally obtained standard deviation
var40
```

```
## [1] 26.34798
```

```
var_th <- std_th*std_th    #std_th is the theoritically computed standard deviation
var_th
```

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

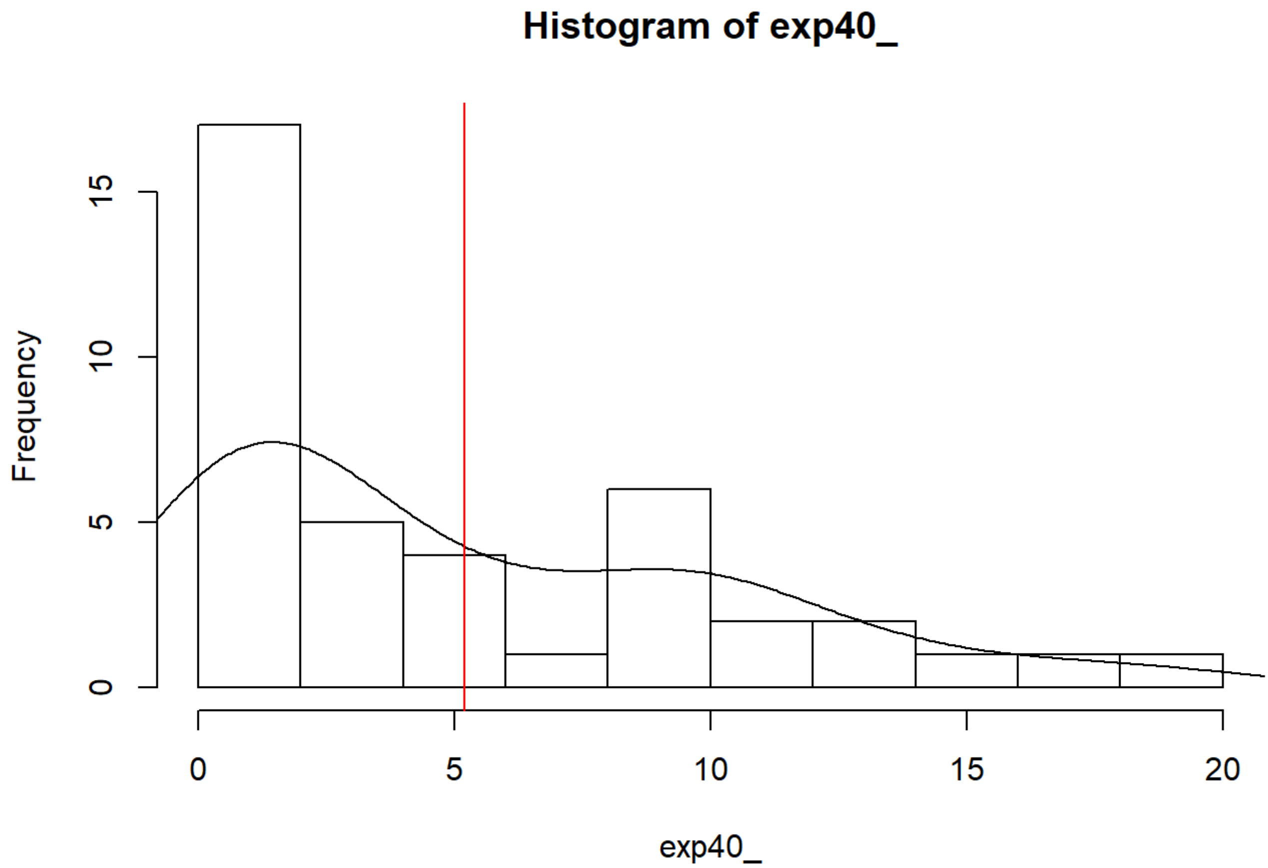
```
var_th-var40               #comparing the obtined variance
```

```
## [1] -1.347979
```

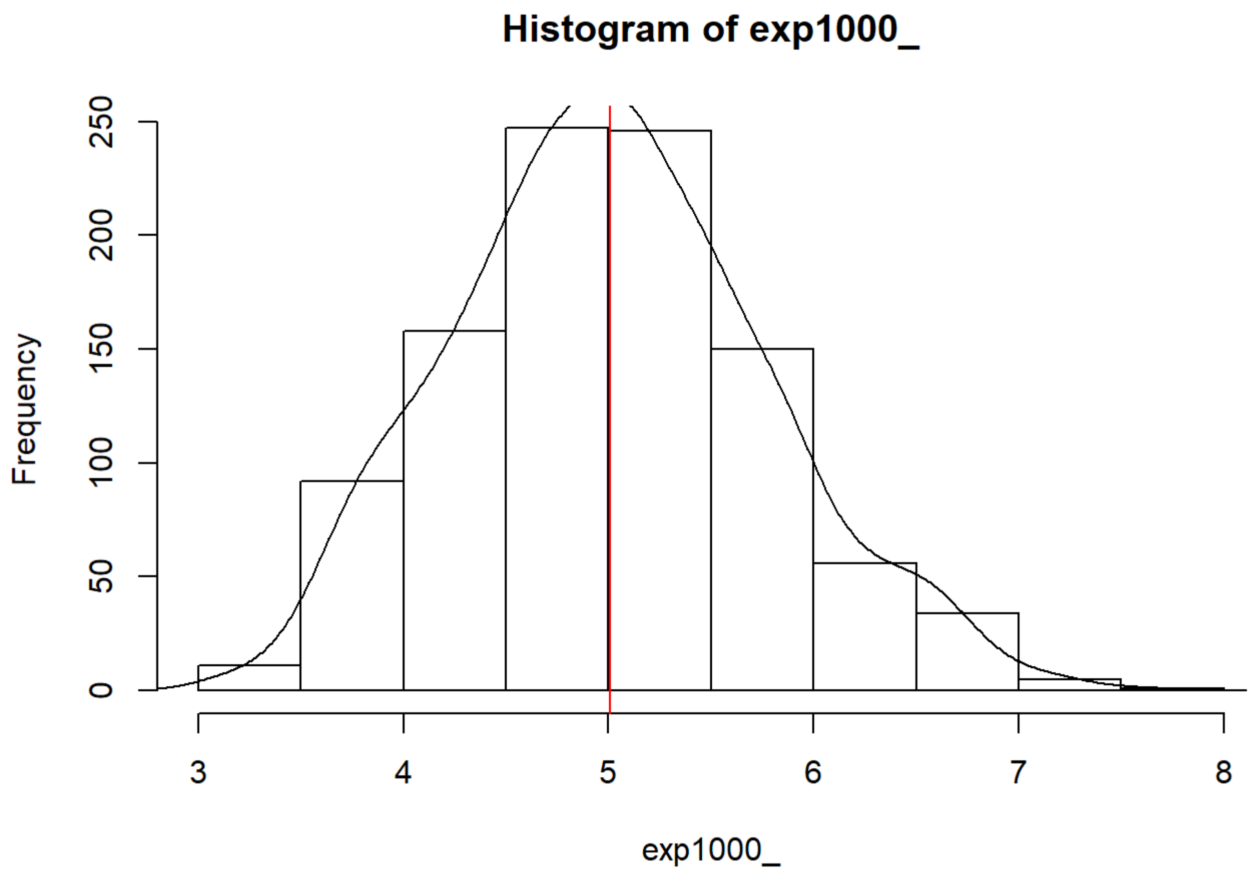
```
var40 and var_th are the experimentally calculated and theoritical variances

Therefore from the above calculations it can be seen that there is a difference in the varian
ce of the samples,as the difference in  variance value increases, the data is spreadout on a 
larger scale and hence many values are away from mean value, which signifies that the confidi
ence interval is narrow for one particular p value.Therefore the probabality of obtaining exp
ected value is less.

Central Limit Theorem: The central limit theorem (CLT) establishes that, in some situations, 
when independent random variables are added, their properly normalized sum tends toward a nor
mal distribution for large number of samples even if the original variables themselves are no
t normally distributed.
```




```
It can be seen from the above histogram that distribution is not gaussian, because the maximu
m point is not at the mean value depicted by red line moreover it is not bell shaped.
```



```
It can be seen from the above histogram that distribution is gaussian, because the maximum po
int is at the mean va-lue depicted by red line, and also it is bell shaped.
```

```
mean [1] 5.007041
```

```
The mean is almost equal to the calculated mean.Hence the obtained data set can be approximat
ed to that of a normal distribution,as the number of samples increases,that is tends to infin
ity the obtained mean will exactly be equal to theoritical mean.
```

Result

From the above computation and experimentation we can conclude that the theoritical and experimental values are approximately equal.Central Limit Theorem is verified for larger data samples,hence large data sets can be approximated as a normal distribution.