

Statistical Inference Course Project Question 1

Sheetal

13 October 2018

Simulation Exercise

The purpose of this project is to investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with `rexp(n, lambda)` where λ is the rate parameter. compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with `rexp(n, lambda)` where λ is the rate parameter. The mean of exponential distribution is $1/\lambda$ and the standard deviation is also $1/\lambda$.

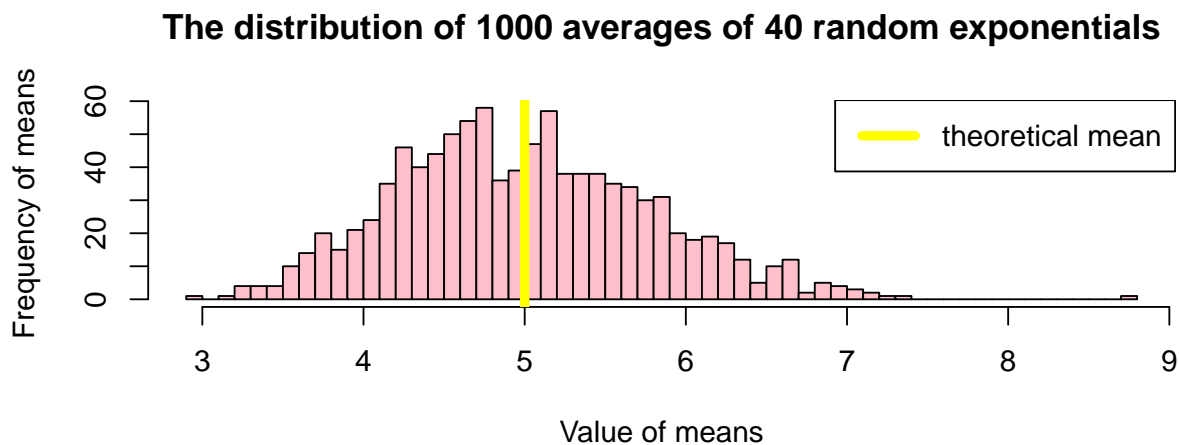
$\lambda = 0.2$ Simulations = 1000

Question 1

Show the sample mean and compare it to the theoretical mean of the distribution.

```
lambda <- 0.2
Data <- matrix(rexp(1000*40, lambda), nrow = 1000, ncol = 40)
distMean <- apply(Data, 1, mean)

hist(distMean, breaks = 50, main = "The distribution of 1000 averages of 40 random exponentials",
      xlab = "Value of means", ylab = "Frequency of means", col = "pink")
abline(v = 1/lambda, lty = 1, lwd = 5, col = "yellow")
legend("topright", lty = 1, lwd = 5, col = "yellow", legend = "theoretical mean")
```



In the above chart, I have simulated the exponential distribution 1000 times, and created a distribution of the means. The AB line is at the mean. The results show that the sample mean is close to the theoretical mean.

the below is the mean for all the simulations one at a time #

```
simMean <- rowMeans(Data)
```

the below is the mean for the data, in this case 4.9632

```
simMeanMean <- mean(simMean)
```

the theoretical mean is $1/\lambda$ i.e $1/0.2$, in this case 5

```
theatricalmean <- 1/0.2
```

both the sample mean and the theatrical mean are similar

Question 2

Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

```
simulationvar <- var(simMean)
```

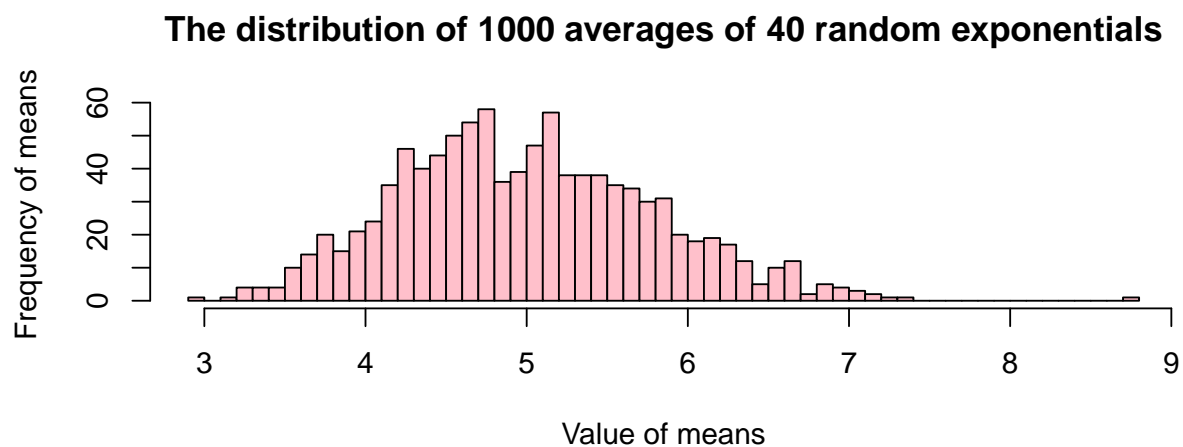
```
theatricalvar <- (1/lambda)^2/(40)
```

simulation variance is 0.61, theatrical variance is 0.62 and hence they are close

Question 3

Show that the distribution is approximately normal

```
hist(distMean, breaks = 50, main = "The distribution of 1000 averages of 40 random exponentials",  
     xlab = "Value of means", ylab = "Frequency of means", col = "pink")
```



The output graph looks close to a normal distribution