Statistical Inference - Course Assignment Part1

Praveen S

16 December 2017

Overview

This is the course project for the Coursera Statistical Inference class. It analyzes the exponential distribution in R and aplies the Central Limit Theorem to a thousand simulations.

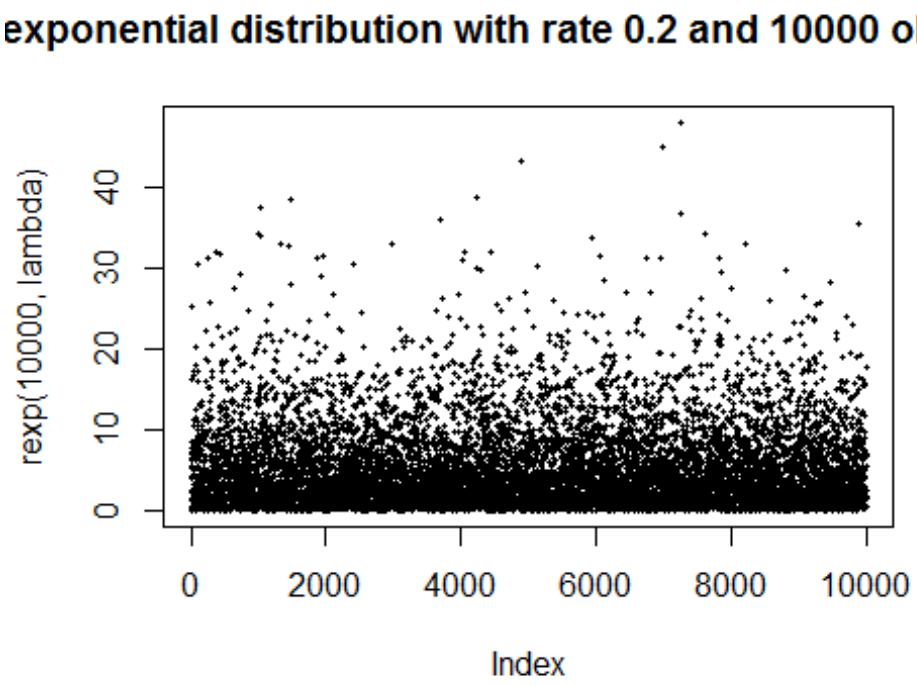
This work investigates the exponential distribution in R and compares it with the Central Limit Theorem.The project explains with simulation and explaratory texts through

1.Simulation Exercise 2.Basic Inferential Data Analysis

Simulation Exercise

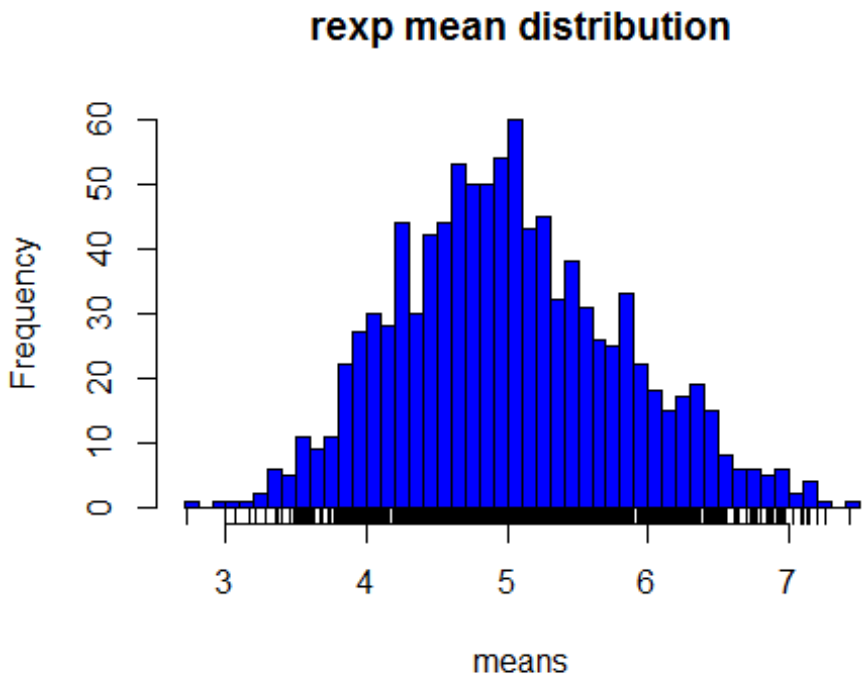
```
set.seed(5000)
lambda <- 0.2 # The rate parameter lambda has given
n <- 40 # Number of exponentials
sim <- 1000 # A thousand simulations

# The exponential distribution
plot(rexp(10000,lambda), pch=20, cex=0.6, main = "The exponential
distribution with rate 0.2 and 10000 observations")
```



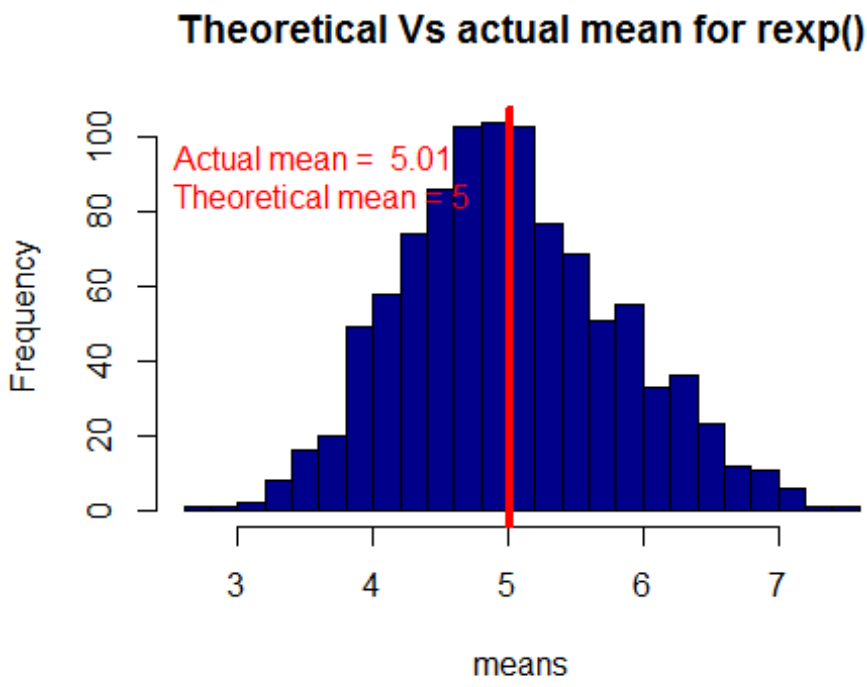
Generate th collection of means for 1000 simulations of the experimental distribution

```
means <- NULL
for(i in 1 : sim)
means <- c(means, mean(rexp(n,lambda)))
hist(means, col="blue", main = "rexp mean distribution", breaks =40)
rug(means)
```



Sample Mean Vs Theoretical Mean The mean of the exponential distribution is in theory 1/lambda. Since lambda is 0.2, the theoretical mean should render 5. Let's check if the numbers match

```
hist(means, col="darkblue", main = "Theoretical Vs actual mean for rexp()",
breaks=20)
abline(v=mean(means), lwd="4", col="red")
text(3.6, 90, paste("Actual mean = ", round(mean(means),2), "\n Theoretical
mean = 5"), col="red")
```



Sample Variance Vs Theoretical Variance The standard deviation of the exponential distribution is (1/lambda)/sqrt(n). Let's see if these match # theoretical standard deviation Vs practical standard deviation

```
print(paste("Theoretical standard deviation: ", round((1/lambda)/sqrt(n) ,4),
", Practical standard deviation", round(sd(means) ,4)))
```

```
## [1] "Theoretical standard deviation:  0.7906 , Practical standard
deviation 0.7998"
```

```
# The variance should be
print(paste("Theoretical variance: ", (1/lambda)^2/n, ",Practical variance",
round(var(means) ,4)))
```

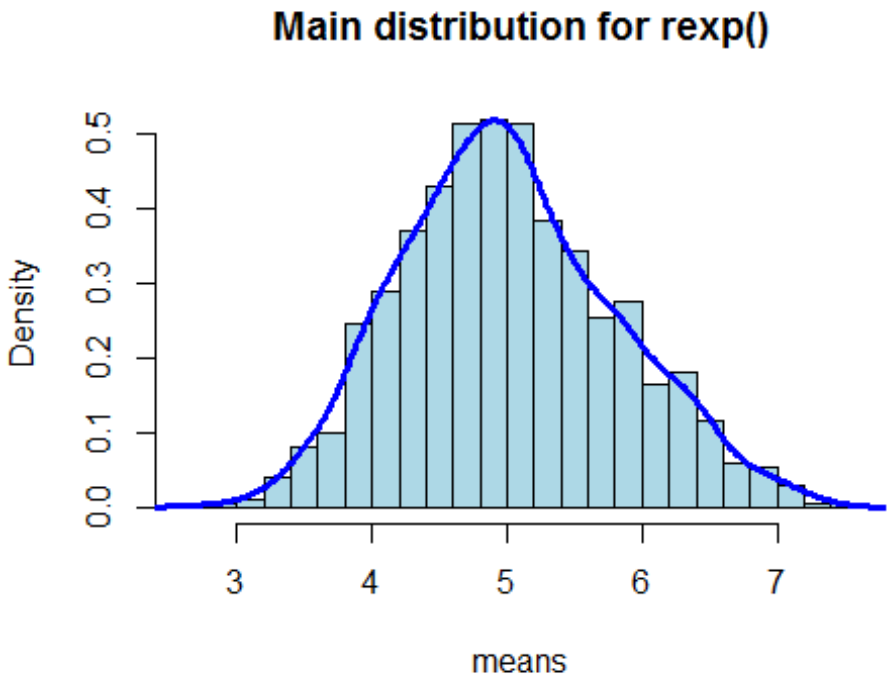
```
## [1] "Theoretical variance:  0.625 ,Practical variance 0.6396"
```

So the numbers match

Is the distribution of means normal ?

Finally, let's check if the exponential distribution is approximately normal. Due to the Central limit Theorem, the average of samples should follow a normal distribution.

```
hist(means, prob=TRUE, col = "lightblue", main = "Main distribution for
rexp()", breaks = 20)
lines(density(means),lwd=3, col="blue")
```



As shown in the above graph, the calculated distribution of means of random sampled exponential distributions overlaps with the normal distribution, due to the Central Limit Theorem. The more samples we would get, the closer will the density distribution be to the normal distribution bell curve.