

Statistical Inference Project

April L

April 7, 2018

Overview

In this project, you will investigate the exponential distribution in R and compare it with the Central Limited Theorem. The exponential distribution can be simulated 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 $1/\lambda$. Set `lambda = 0.2` for all the simulations.

Simulation:

We simulate 1000 'average of 40 random exponentials observations'

Sample Mean vs Theoretical Mean = $1/\lambda$

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

```
sample_mean <- mean(means)
print(paste('The sample mean = ', round(sample_mean,2), ' is comparable to the theoretical mean of the d

## [1] "The sample mean = 4.99 is comparable to the theoretical mean of the distribution 5"
```

Sample Variance vs. Theoretical Variance = $(1/\lambda)^2$

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

```
sample_var <- mean(variances)
print(paste('The sample mean = ', round(sample_var,2), ' is comparable to the theoretical mean of the di

## [1] "The sample mean = 25.11 is comparable to the theoretical mean of the distribution 25"
```

Distribution

3. Show the distribution is approximately normal.

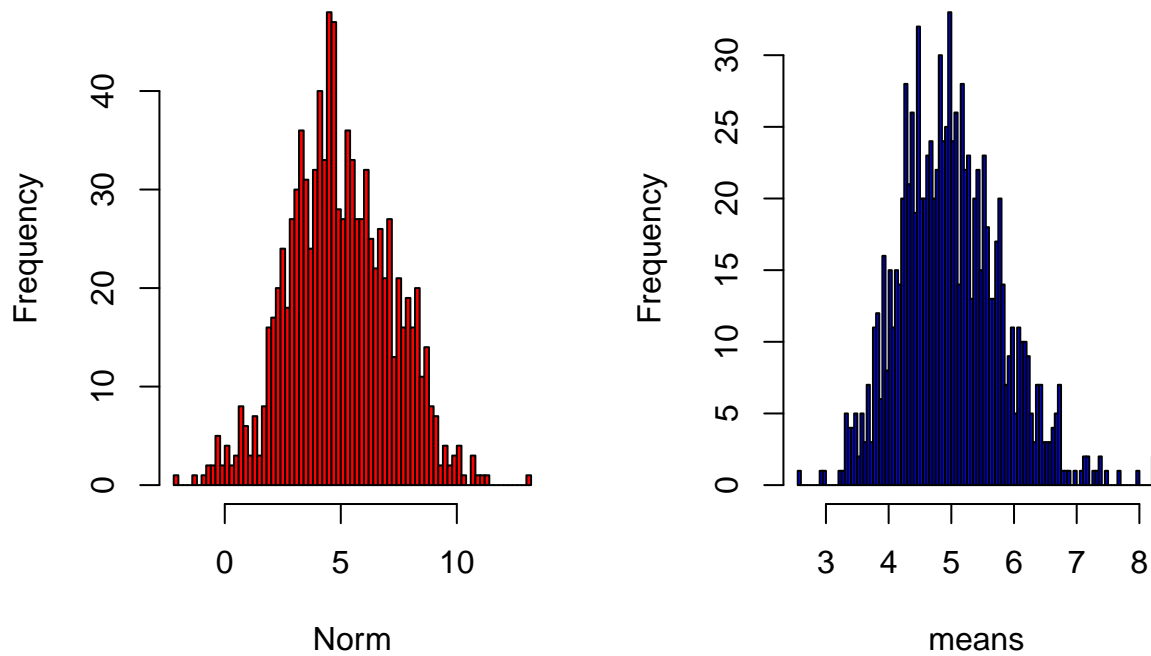
In point 3, focus on the difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials.

```
par(mfrow=c(1,2))

# a large collection of random exponentials of lambda = 0.2
Norm <- rnorm(1000, mean=1/lambda, sd = sqrt(1/lambda))
hist(Norm, breaks=100, main='Distribution of 1000 random exponentials', col='red')
```

```
# distribution of a large collection of averages of 40 exponentials
hist(means, breaks=100, main="Distribution of 1000 averages of 40 exponentials, lambda = 0.2", col='blue')
```

Distribution of 1000 random exponential of 1000 averages of 40 exponentia



Compare the 2 charts, it shows that the distribution of the means is approximately normal.