Analysis of Central Limit Theorem on Exponential Distribution

Thawatchai Phakwithoonchai

12/21/2019

Overview / Synopsis

This assignment is to investigate the random exponential distribution in R and compare it with the Central Limit Theorem. The result indicates that the **distribution of averages of random exponential variables become closed to normal.**

Simulation Process

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 of the simulations. The distribution of averages of 40 exponentials will be investigated with a thousand simulations.

```
set.seed(1234)   #set.seed for reproducibility
lambda <- 0.2
n <- 40
simno <- 10000
```

1. Compared the sample mean with the theoretical mean of the distribution.

```
simmeans = NULL
for (i in 1:1000) simmeans = c(simmeans, mean(rexp(n, lambda)))
mean(simmeans)
```
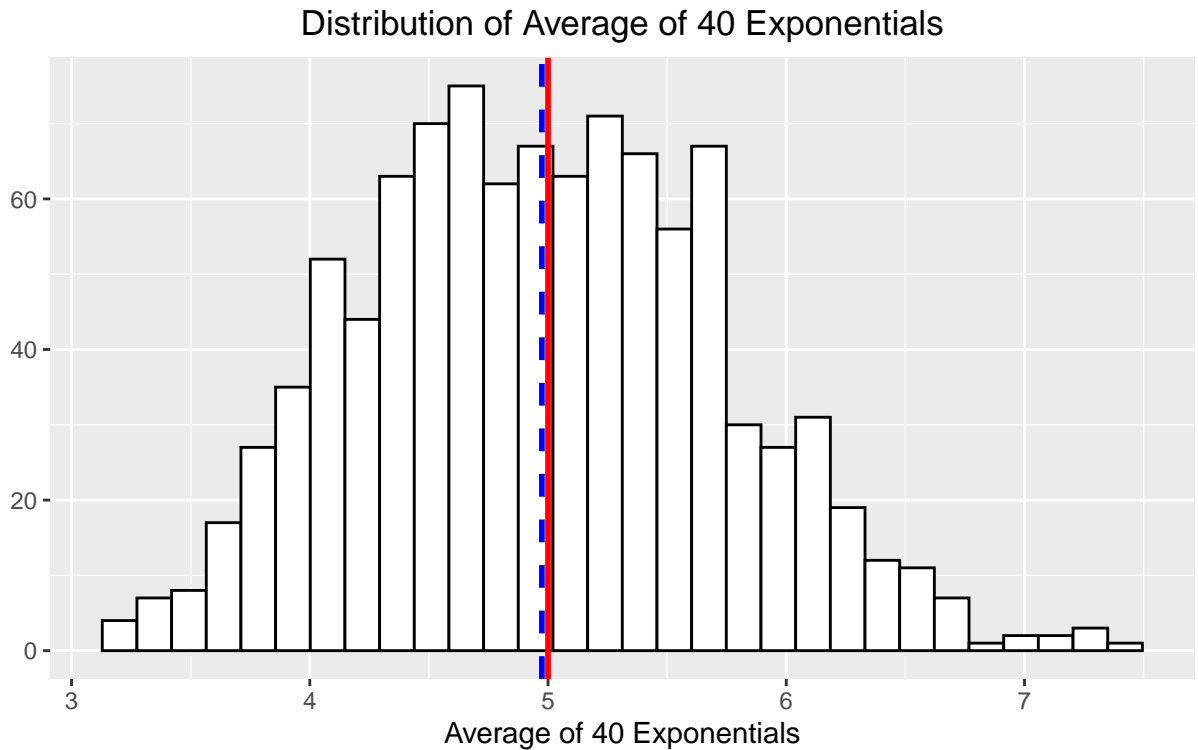
```
## [1] 4.974239
```

```
theomean <- 1/lambda
theomean
```

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

The following graph shows the distribution of average of 40 exponentials. The solid red line indicates the theoretical mean, while the dashed blue one indicates the sample mean.

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Distribution of Average of 40 Exponentials

2. Compared the sample variance with the theoretical variance of the distribution.

```
c(sd(simmeans), var(simmeans))
```

```
## [1] 0.7554171 0.5706551
```

```
c(1/lambda/sqrt(n), 1/lambda^2/n)
```

```
## [1] 0.7905694 0.6250000
```

The table highlighting basic features of the data is shown as following:

```
##                     Theoretical    Sample
## Mean                  5.0000000 4.9742388
## Standard Deviation    0.7905694 0.7554171
## Variance              0.6250000 0.5706551
```

3. Display that the distribution is approximately normal.

By focusing on the difference between the distribution of a large collection of random exponentials (positive skew shape) and the distribution of a large collection of averages of 40 exponentials. (approximately normal shape), the analysis result is apparently indicated that by applying the Central Limit Theorem, the distribution of averages of random exponential variables will be approximately normal.

```
rexpo <- rexp(simno, lambda)
par(mar = c(4, 4, 1, 1))
hist(rexpo, main = "Histogram of Random Exponentials", xlab = "Random Exponentials Value",
    breaks = 30)
```

Histogram of Random Exponentials
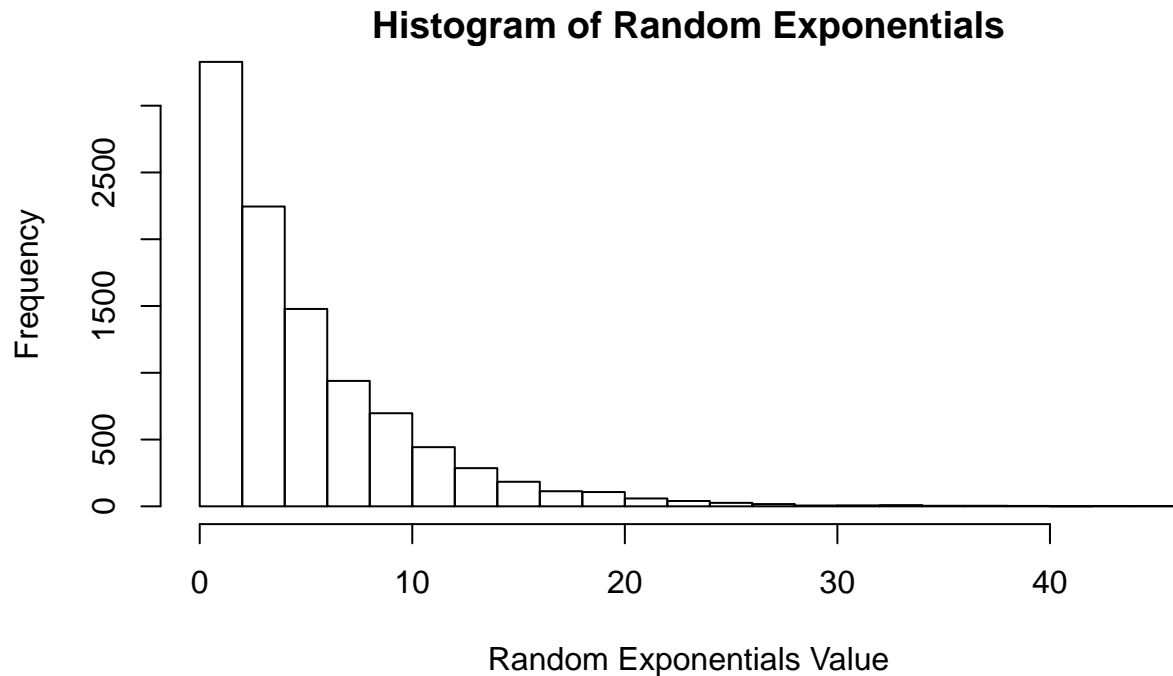

```
par(mar = c(4, 4, 1, 1))
hist(simmeans, main = "Histogram of Averages of 40 Random Exponentials",
    xlab = "Averages of 40 Random Exponentials", breaks = 20)
```

Histogram of Averages of 40 Random Exponentials