Statistical Inference Course Project - Part 1

Mounika Yalamarthi

April 11, 2020

Overview

In Part 1 of this project, we will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution will be simulated in R with rexp(n, lambda), where lambda is the rate parameter. The mean of the exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Lambda = 0.2 for all simulations.

In this project, I will investigate the distribution of averages of 40 exponentials and do 1,000 simulations.

Simulations

First we will set up the simulation parameters.

```
set.seed(1111)

lambda = 0.2

n = 40
```

Next, we will run the simulations and calculate the means.

```
sims <- replicate(1000, rexp(n, lambda))

sims_means <- apply(sims, 2, mean)
```

Sample Mean vs. Theoretical Mean

The sample mean is the mean of the simulation means, and is calculated and graphically shown below:

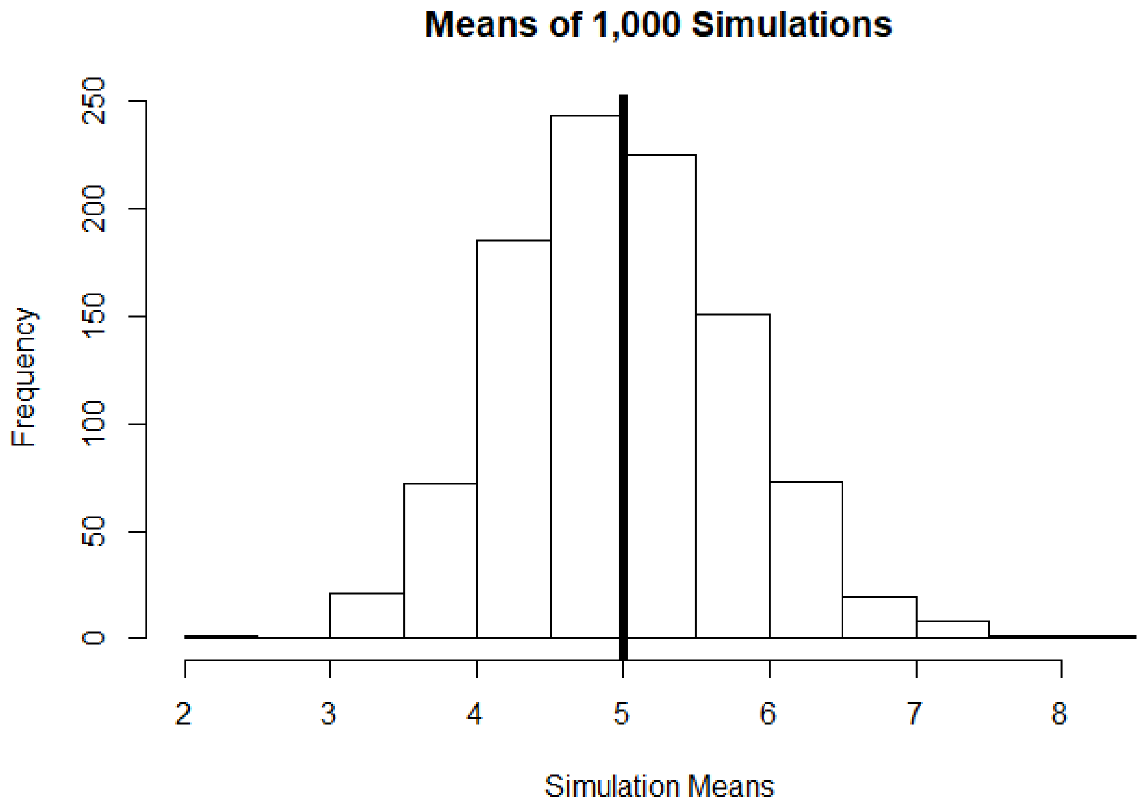
```
sample_mean <- mean(sims_means)

sample_mean
```

```
## [1] 4.994883
```

```
hist(sims_means, main = "Means of 1,000 Simulations", xlab = "Simulation Mean
s")

abline(v = sample_mean, lwd = 5)
```



The sample mean is shown as the thick line.

The theoretical mean is 1/lambda.

```
theo_mean <- 1/lambda
theo_mean
## [1] 5
```

So, the difference between the sample mean and theoretical mean is

```
abs(sample_mean - theo_mean)
## [1] 0.005117261
```

We can see that the sample mean is very close to the theoretical mean.

Sample Variance vs. Theoretical Variance

The sample variance is calculated below.

```
sample_var <- var(sims_means)
sample_var
## [1] 0.6105269
```

The theoretical variance is equal to [(1/lambda)/sqrt(n)]^2.

```
theo_var <- ((1/lambda)/sqrt(n))^2
theo_var
## [1] 0.625
```

So, the difference between the sample variance and theoretical variance is

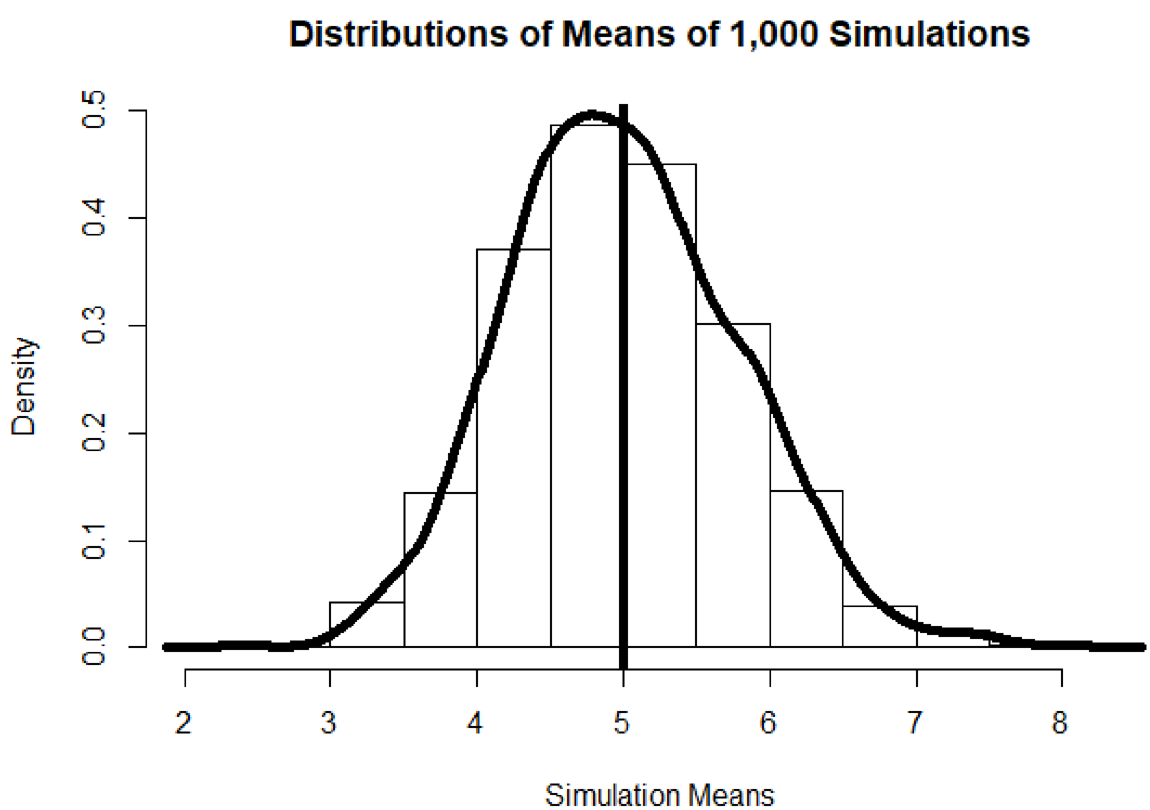
```
abs(sample_var - theo_var)
## [1] 0.0144731
```

We can see from the above that both the sample and theoretical variance are close to 0.6.

Distribution

In order to quickly tell if the distribution is normal, we can plot the means on a histogram and see if the overall shape resembles that of the normal distribution bell curve.

```
hist(sims_means, prob = TRUE, main = "Distributions of Means of 1,000 Simulat
ions", xlab = "Simulation Means")
abline(v = sample_mean, lwd = 5)
lines(density(sims_means), lwd = 5)
```



We can see from the shape of the red line that this distribution is very close to a normal distribution. If we were to run more simulations, the curve would get closer and closer to resembling the normal distribution bell curve.