Statistical Inference Course Project_Part 1

Shee

Overview

In this project we will 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 lambda is the rate parameter.

Simulation

```
#set seed to reproduce the data
set.seed(7714)
#set constants
lambda = 0.2
n = 40
nsims = 1:1000

Mean_ExpDist <- data.frame(x = sapply(nsims, function(x) {mean(rexp(n, lambda))}))
```

Sample Mean versus Theoretical Mean

The theoretical mean value of an exponential distribution of lambda is:

```
Emean <- 1/lambda
Emean
```

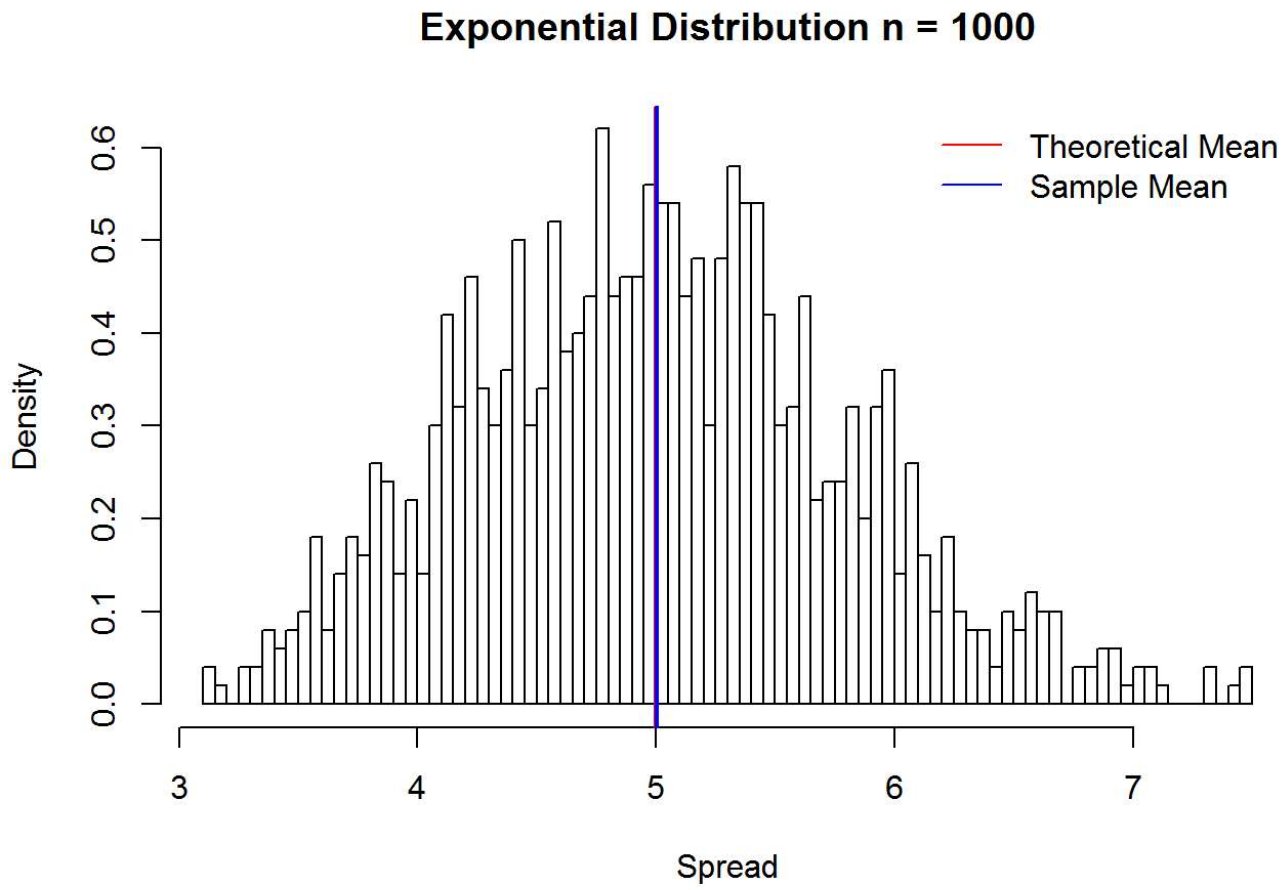
```
## [1] 5
```

The average of sample mean value of 1000 simulations of 40 random samples of exponential distributions is:

```
Smean <- mean(Mean_ExpDist$x)
Smean
```

```
## [1] 5.005964
```

Histogram plot of the exponential distribution n = 1000:



Thus this showed that the theoretical mean value of the exponential distribution and the avaerage of sample mean of 1000 simulations are very close.

Sample Variance versus Theoretical Variance

The theoretical variance of a exponential distribution of lambda is:

```
EVar <- (1/lambda)^2/n
EVar
```

```
## [1] 0.625
```

The variance of sample mean of 1000 simulations of 40 random samples of exponential distributions is:

```
SVar <- var(Mean_ExpDist$x)
SVar
```

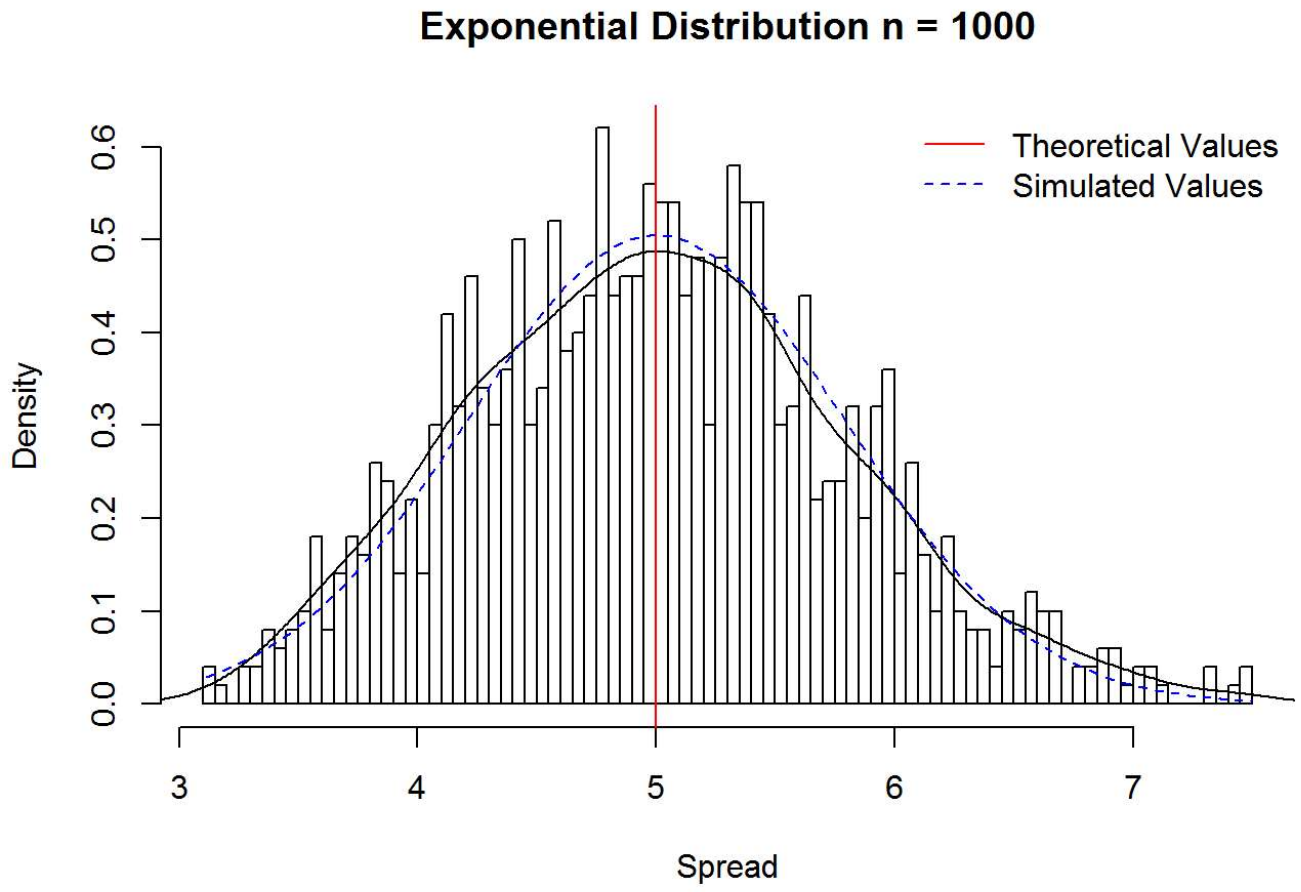
```
## [1] 0.6284007
```

Again, the sample variance and the theoretical variance are pretty close.

Distribution

Comparing the population means & standard deviation with a normal distribution of the expected values.

```
hist(Mean_ExpDist$x,
     breaks = 100,
     prob = TRUE,
     main = "Exponential Distribution n = 1000",
     xlab = "Spread")
     lines(density(Mean_ExpDist$x))
     abline(v = 1/lambda, col = "red")
     xfit <- seq(min(Mean_ExpDist$x), max(Mean_ExpDist$x), length = 100)
     yfit <- dnorm(xfit, mean = 1/lambda, sd = ((1/lambda)/sqrt(40)))
     lines(xfit, yfit, pch = 22, col = "blue", lty = 2)
     legend('topright', c("Theoretical Values", "Simulated Values"),
            bty = "n", lty = c(1,2), col = c("red", "blue"))
```



From the graph, the distribution of simulated means of random sampled exponantial distributions, is quite near to the normal distribution with the expected values based on the given lambda.

The figure has proved that the distribution is approximately normal.