project1

Mohamed Elashri

March 22, 2020

Statistical Inference Course Project

In this project you 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. 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. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Simulation

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```
set.seed(11081979)

lamda <- 0.2
n <- 40 # exponetials number
NS <- 1000 # tests number

         exp_sim <- function(n, lamda)
           {
                     mean(rexp(n,lamda))
           }

           sim <- data.frame(ncol=2,nrow=1000)
           names(sim) <- c("Index","Mean")

           for (i in 1:NS)
           {
                     sim[i,1] <- i
                     sim[i,2] <- exp_sim(n,lamda)
           }
```

Sample Mean vs Theoretical Mean

Sample Mean

```
sample_mean <- mean(sim$Mean)

sample_mean
```

```
## [1] 5.027126
```

Theoretical Mean

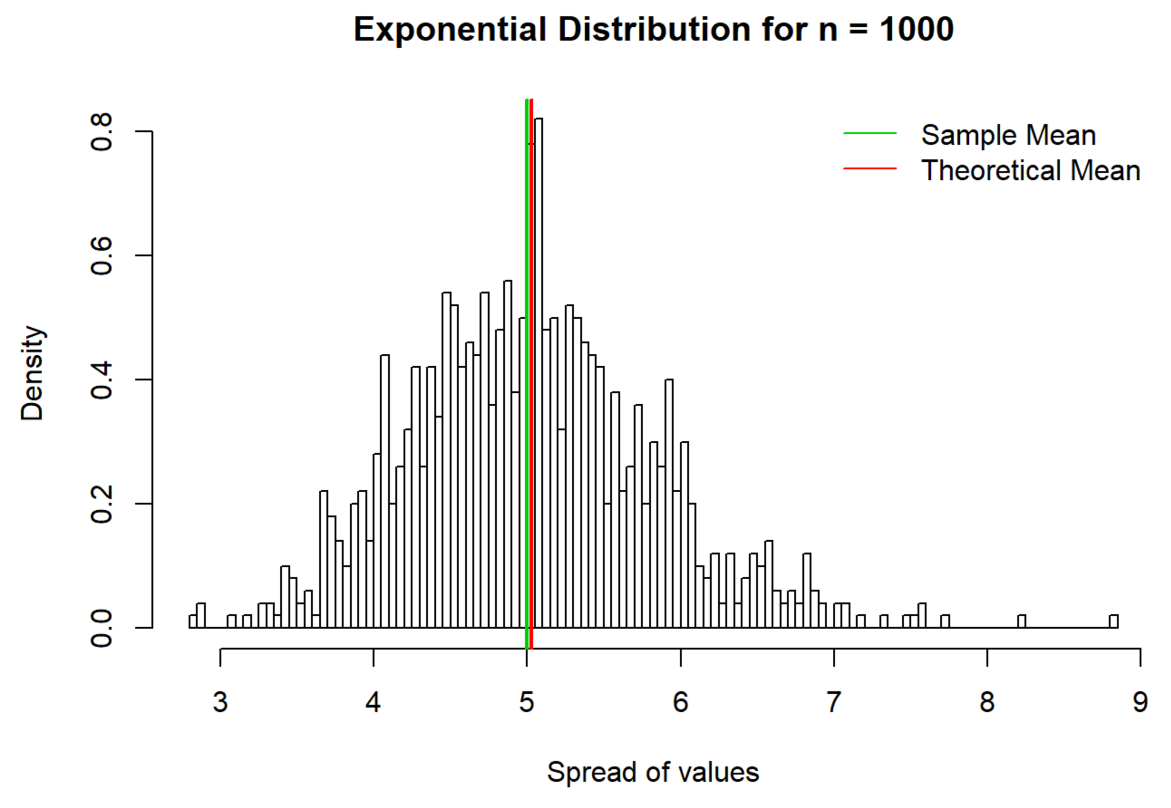
```
mean_theory <- 1/lamda
mean_theory
```

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

Histogram

```
hist(sim$Mean,
  breaks = 100,
   prob = TRUE,
    main="Exponential Distribution for n = 1000",
    xlab="Spread of values")
        abline(v = mean_theory,
            col= 3,
             lwd = 2)
                abline(v = sample_mean,
                   col = 2,
                     lwd = 2)

legend('topright', c("Sample Mean", "Theoretical Mean"),
   bty = "n",
    lty = c(1,1),
     col = c(col = 3, col = 2))
```



Sample Mean vs Theoretical Mean

The expected mean ?? of a exponential distribution of rate ?? is

$?? = \frac{1}{??}$

```
sample_var <- var(sim$Mean)
theor_var <- ((1/lamda)^2)/40
```

so the theoretical variance of the population is

```
theor_var
```

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

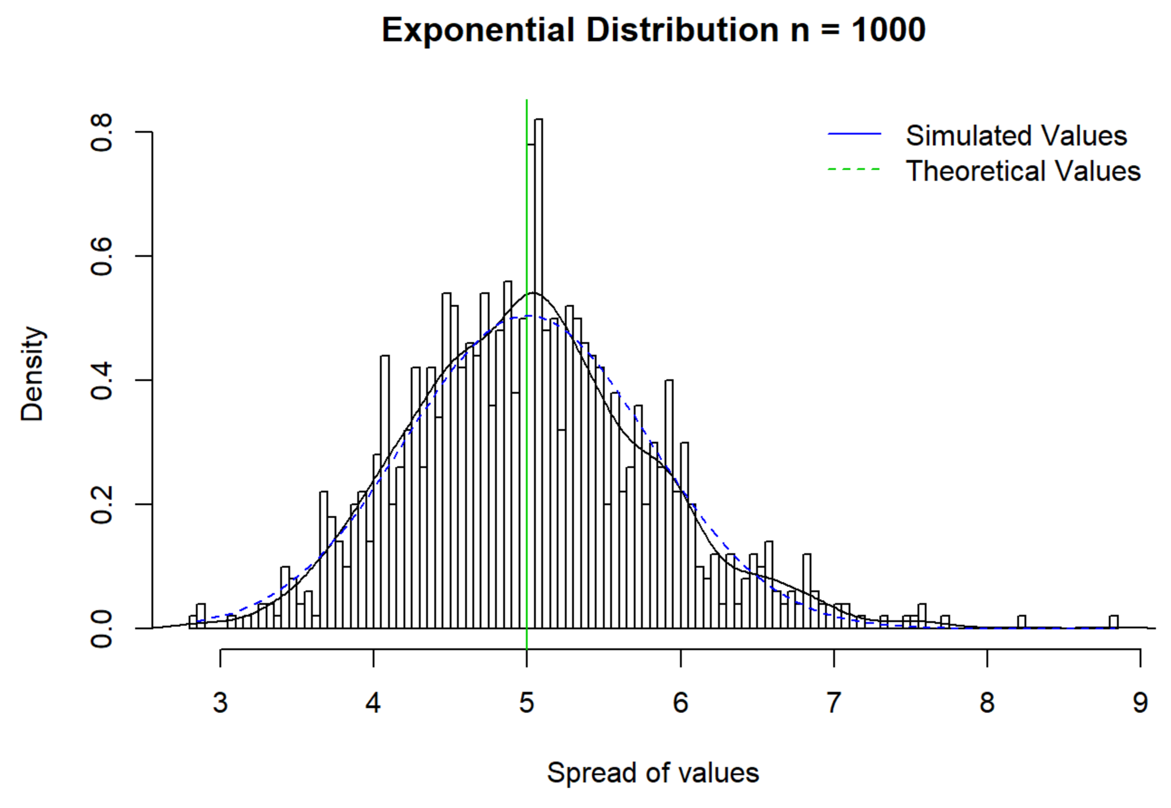
and sample variance is

```
sample_var
```

```
## [1] 0.6374592
```

and this is Histogram of values

```
         hist(sim$Mean,
                     breaks = 100,
                     prob = TRUE,
                     main = "Exponential Distribution n = 1000",
                     xlab = "Spread of values")
                     lines(density(sim$Mean))
                     abline(v = 1/lamda, col = 3)
                     xfit <- seq(min(sim$Mean), max(sim$Mean), length = 100)
                     yfit <- dnorm(xfit, mean = 1/lamda, sd = (1/lamda/sqrt(40)))
                     lines(xfit, yfit, pch = 22, col = 4, lty = 2)
                     legend('topright', c("Simulated Values", "Theoretical Values"),
                     bty = "n", lty = c(1,2), col = c(4, 3))
```



As we can see the standard deviations are very close Since variance is the square of the standard deviations.

Distribution

we can see that this distribution is nearly linear as follows

```
qqnorm(sim$Mean,
        main ="Normal Plot")
        qqline(sim$Mean,
        col = "3")
```