Statistical Inference Course Project: Part 1

Author: Amin Mousavi

Overview

The goal of this project is to investigate and explore the exponential distribution in R and compare it with Central Limit Theorem. The exponential distribution can be simulated in R by rexp(n, λ), where λ is the rate parameter.

Simulation Code

We set λ=0.2 for all simulations, and will investigate the distribution of average of 40 exponentials for 1000 simulations. The code for generating the samples is as follows:

```
set.seed(11)
lambda <- 0.2
nosim <- 1000
n <- 40
meanSamples = matrix(rexp(nosim*n,lambda), nosim)
Means = apply(meanSamples, 1, mean)
```

Sample Mean

We set λ=0.2 for the exponential distribution. So, the theoretical mean is:

$$\mu = \frac{1}{\lambda} = \frac{1}{0.2} = 5$$

The average of 1000 sample means is:

```
> mean(Means)
[1] 4.987157
```

The theoretical mean and the sample mean for the given parameters are very close.

The Variability of Sample Mean

For λ=0.2 the theoretical variance is:

$$\sigma^2 = \left(\frac{1}{\lambda}\right)^2 = 25$$

The variance of the sample mean is:

$$\frac{\sigma^2}{\mathrm{n}} = \frac{25}{40} = 0.625$$

The variance of the sample mean from simulations is:

```
> var(Means)
[1] 0.6509313
```

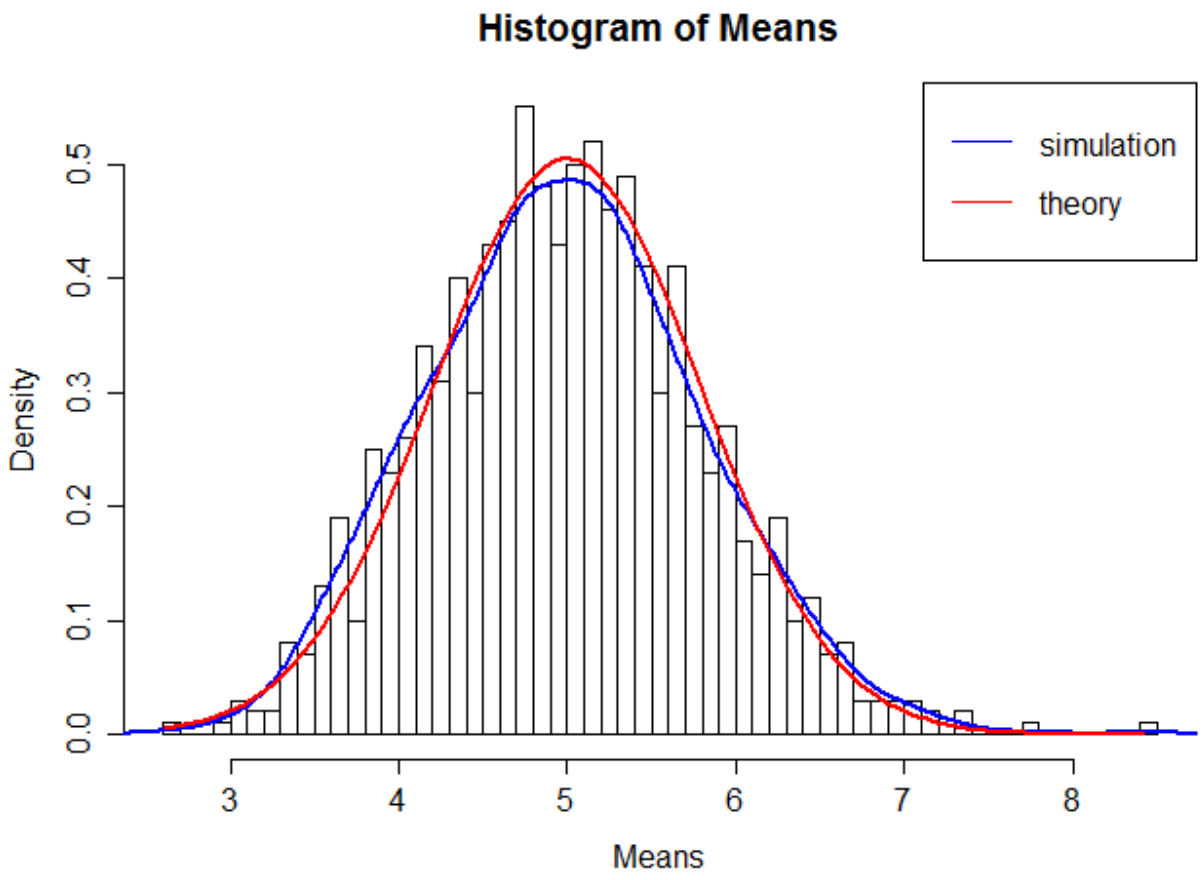
These values are also pretty close.

Distribution

For estimating the distribution of sample mean, first we plot the histogram of sample means. Then, we plot a theoretical normal distribution and compare the plots. The code for generating the plot is:

```
hist(Means, breaks=50, prob = TRUE)
lines(density(Means), col="blue", lwd=2)
x <- seq(min(Means),max(Means), length=100)
lines(x,dnorm(x,mean=5, sd=sqrt(0.625)), col="red",lwd=2)
legend("topright", c("simulation","theory"), lty=c(1,1), col=c("blue","red"))
```

and the result is:



The plot clearly shows that the distribution of the sample mean is very close to the normal distribution.