Statistical Inference Course Project

Carlos Barco

4/27/2017

Instructions

The project consists of two parts: ***1.- A simulation exercise. 2.- Basic inferential data analysis.***

You will create a report to answer the questions. Given the nature of the series, ideally you'll use knitr to create the reports and convert to a pdf. (I will post a very simple introduction to knitr). However, feel free to use whatever software that you would like to create your pdf. Each pdf report should be no more than 3 pages with 3 pages of supporting appendix material if needed (code, figures, etcetera).

Part 1: Simulation Exercise

Overview

It was analized the exponential distribution in R and compared with the Central Limit Theorem. The exponential distribution was simulated in R with rexp(n, lambda) being lambda the rate parameter. The mean of exponential distribution and the standard deviation, both were 1/lambda. Set lambda = 0.2 for all of the simulations. It was investigated the distribution of averages of 40 exponentials. It was carry out a thousand simulations.

Tasks

Show the sample mean and compare it to the theoretical mean of the distribution. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. Show that the distribution is approximately normal.

Simulations:

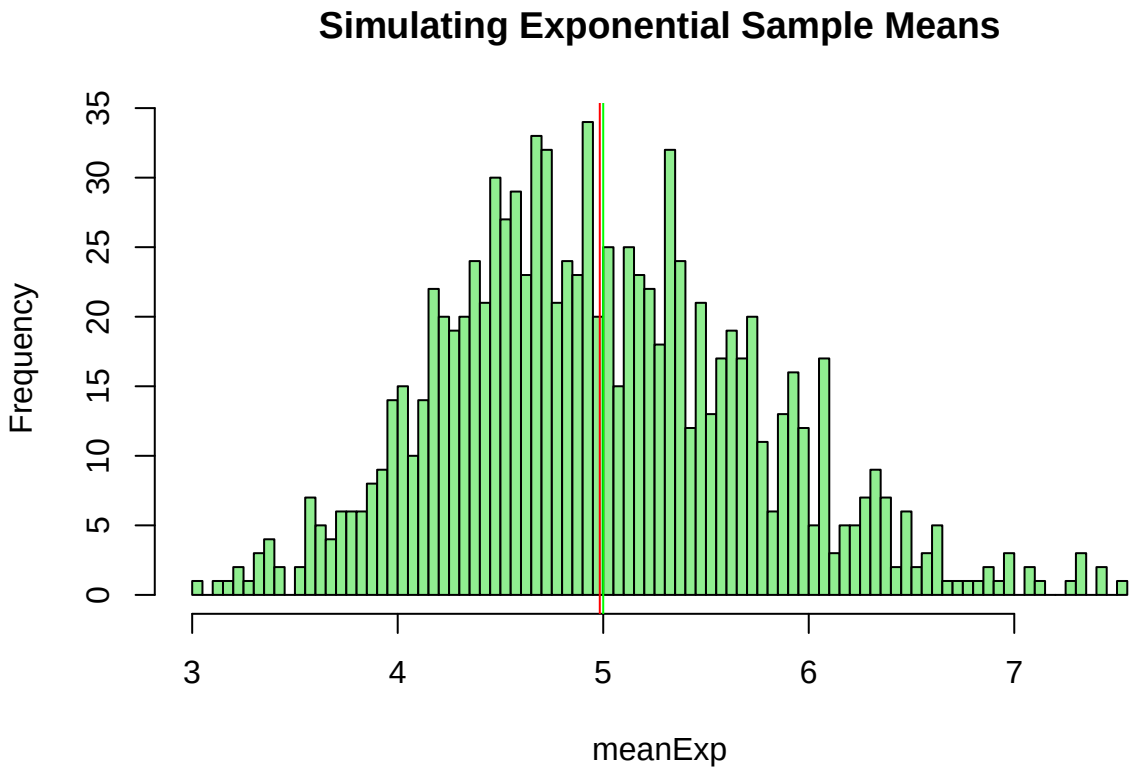
```
# Set Seed
set.seed(7)
# Rate Parameter
lambda <- 0.2
# Exponentials
n <- 40
# Simulations
nosim <- 1000
# Simulating Exponential distribution
sExp <- replicate(nosim, rexp(n, lambda))
# Mean of exponential distribution
meanExp <- colMeans(sExp)
# Sample Mean versus Theoretical Mean:
sampMean <- mean(meanExp)
sampMean
```

```
## [1] 4.983294
```

```r
theoMean <- 1/lambda
theoMean
```

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

```r
# Plots
hist(meanExp, main = "Simulating Exponential Sample Means", col = "lightgreen", breaks = 100)
abline(v = sampMean, col = "red")
abline(v = theoMean, col = "green")
```



For the Sample Mean, I got 4.983294. For the Theoretical Mean, I got 5.

Sample Variance versus Theoretical Variance:

Include figures (output from R) with titles. Highlight the variances you are comparing. Include text that explains your understanding of the differences of the variances.

```r
# Sample Variance
sampSd <- sd(meanExp)
sampVar <- sampSd^2
sampVar
```

```
## [1] 0.5792547
```

```r
# Theoretical Variance
theoSd <- (1/lambda)/sqrt(n)
theoVar <- theoSd^2
theoVar
```

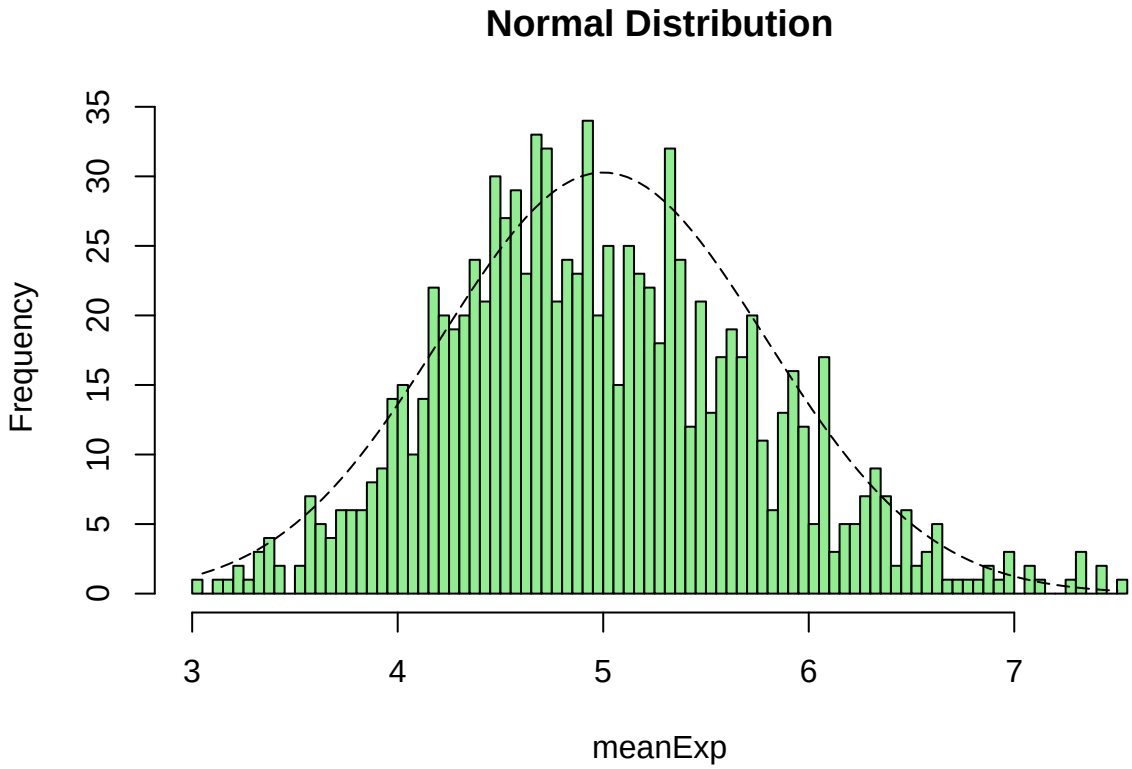
```
## [1] 0.625
```

For the Sample Variance, I got 0.5792547. For the Theoretical Variance, I got 0.625.

Distribution:

Via figures and text, explain how one can tell the distribution is approximately normal.

```
# Explanatory plot
hist(meanExp, main = "Normal Distribution", col = "lightgreen", breaks = 100)
xfit <- seq(min(meanExp), max(meanExp), length = 100)
yfit <- dnorm(xfit, mean = 1/lambda, sd = (1/lambda)/sqrt(n))
lines(xfit, yfit*60, lty = 5)
```




```
# Accordingly with the plot, our distribution is normal
```