Statistical Inference Project - Simulation Exercise

Sherman Wood

July 25, 2016

Introduction

This is my submission for the Coursera Statistical Inference Course Project - Simulation Exercise.

Simulation Exercise

Overview

We will investigate the exponential distribution and compare it with the Central Limit Theorem through the use of R.

The Central Limit Theorem states that as n gets larger, the distribution of the difference between the sample average Sn and its limit mu, when multiplied by the factor sqrt(n) (that is sqrt(n)(Sn - mu)), approximates the normal distribution with mean 0 and variance sigma^2. For large enough n, the distribution of Sn is close to the normal distribution with mean mu and variance sigma^2/n.

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The theoretical mean of an exponential distribution is 1/lambda and the theoretical standard deviation is also 1/lambda.

Simulation

For this investigation, we will evaluate the distribution of means of 1,000 samples of 40 random exponentials with lambda = 0.2.

For this exponential distribution, the Central Limit Theorem says that:

mu = 1/lambda = theoretical sample mean (Sn) = 5

Theoretical Sample variance = (1/lambda)^2 / n = 0.625

First let's get our sample exponential distribution.

```
for (i in 1 : simulations) means = c(means, mean(rexp(n,lambda)))
```

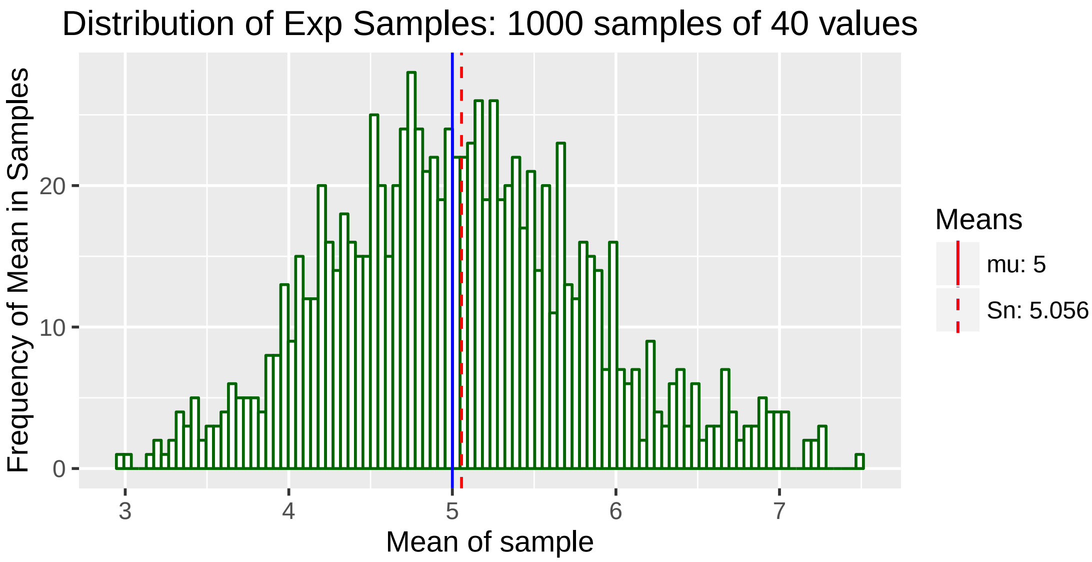
The statitics of the the sample are:

Sample mean (Sn) = 5.056

Sample variance = 0.6544

Sample standard deviation = 0.8089

Visually the sample is:



Comparison of Theoretical and Sample Means and Variance

Compare the sample distribution statistics to the values expected from the Central Limit Theorem :

Theoretical mean mu - Sn = -0.056

Theoretical - Sample variance = -0.0294

These differences are close to 0. The Central Limit Theorem assertions are thus supported.

Comparison of the sample to the normal distribution

A density plot allows us to compare the sample distribution to the normal distribution. The closeness of these curves highlights the Central Limit Theorem assertion that for large enough n, the distribution of Sn is close to the normal distribution with mean mu and variance sigma^2/n.