Exponential Distribution and Central Limit Theorem

Exponential Distribution and Central Limit Theorem

CJ

Thursday, January 08, 2015

Overview

Simulation

Sample Mean versus Theoretical Mean

Sample Variance versus Theoretical Variance

Comparison with the Central Limit Theorem

Overview

In this project we investigate the exponential distribution and compare it with the Central Limit Theorem. We sample 40 exponentials and calculate their mean and standard deviation. After a thousand of simulations, we compare the average of such mean and standard deviation with its theoretical values, whose value should be both 1/lambda for the mean and the standard deviation. Finally, the distributions of the sample mean and standard deviation are normalized and compared with the standard normal.

Simulation

Firstly, calculate the theoretical values of exponential mean and standard deviation. When lambda= 0.2, both mean and standard deviation equal to 5.

Now, we generate 1000 averages of 40 random exponentials. In the chunk, a function called exps generate a vector of size n contains transformation of m exponentials. For example, here we use function mean, which returns a vector of means of m exponetials. By passing the function var to our exps function, we generate 1000 variances of 40 random exponentials.

Sample Mean versus Theoretical Mean

Plot the histogram of a thousand of simulated means and compare the average of these means with its theoretical value.

As we can see from Fig1, the center of sample means is very close to the theoretical mean with a standard error 0.02558013.

Sample Variance versus Theoretical Variance

Next, we investigate on the sample variance.

Plot the histogram of a thousand of simulated variances and compare the average of these variances with its theoretical value.

Comparison with the Central Limit Theorem

The law of large numbers (LLN) states that the average of the results obtained from a large number of trials should be close to the expected value. As shown in both Fig1 and Fig2, mean and variance are close to their theoretical values. Futher more, the Central Limit Theorem (CLT) states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases.

Nomarlization

Here we normalize both sample mean and variance and compare with the standard normal distribution.

Now plot both distributions of sample mean and variance and overlay with the standard normal distribution.

As we can see in Fig3, both (a) and (b) are close to standard normal distribution, which center at 0 and most of the data are within 3 times σ. But variance distribution is more skewing to the right. This might due to the limitation of the CLT that doesnâ€™t give the exact sample size.