

Using exponential distribution in R to validate the Central Limit Theorem

Avranil Ghosh

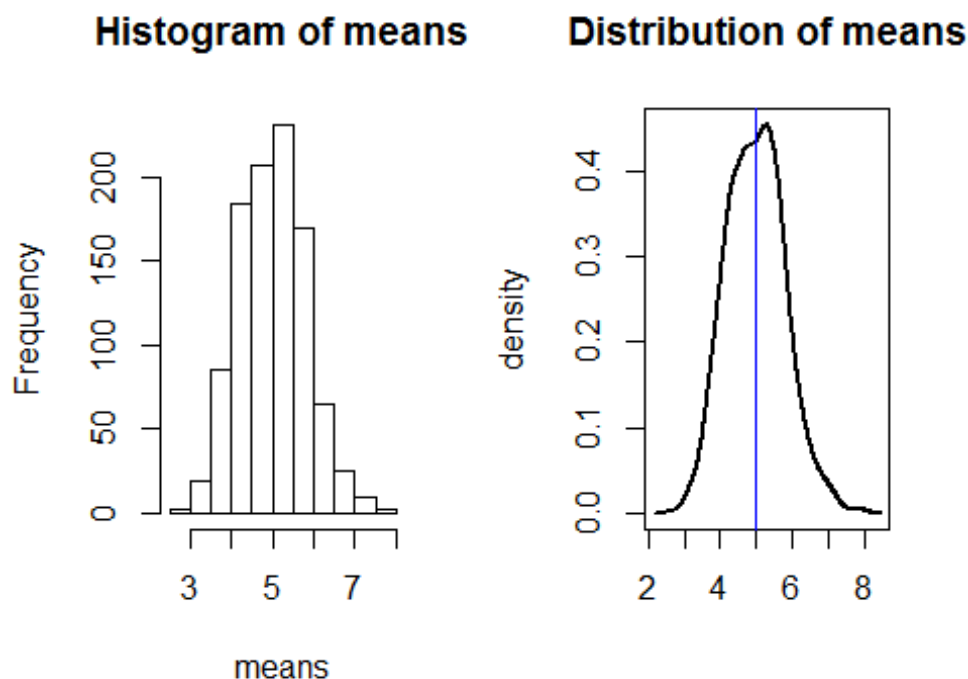
Sunday, January 25, 2015

Overview:

In this project we 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 $1/\lambda$. We set `lambda = 0.2` for all of the simulations.

Analysing sample mean and comparing it to the theoretical mean of the distribution.

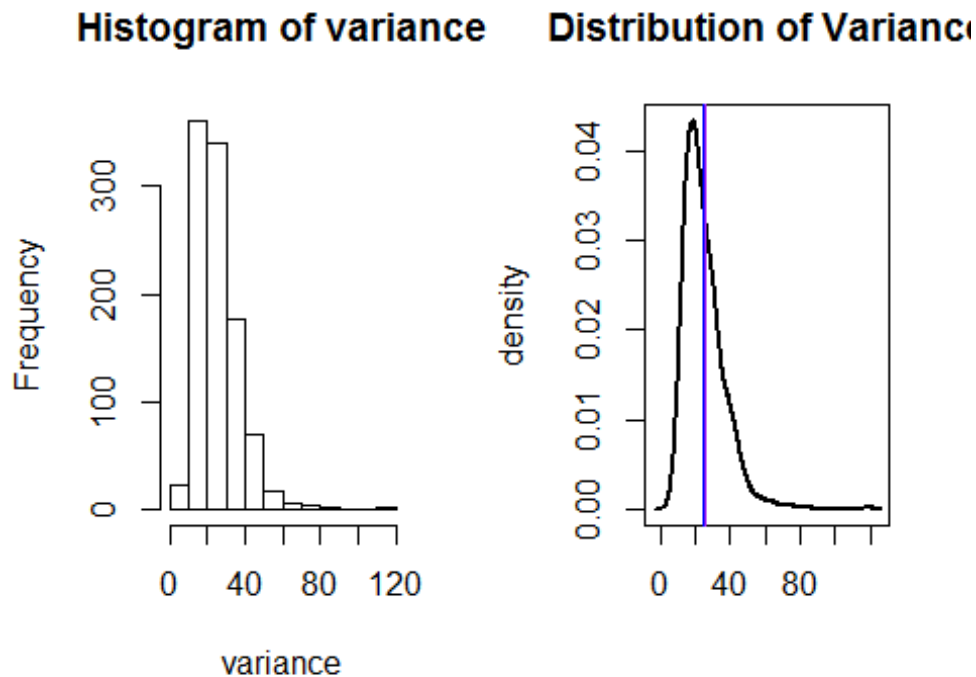
Here we plot the histogram and density distribution of averages of 40 random exponentials simulated over 1000 times.



The above figures illustrate that the mean of distribution of a large collection of averages of 40 exponentials that comes out to be 5.0020869 is very close to the actual population theoretical mean which is 5. We can conclude that the average limits to what it is estimating - the population mean.

Analysing sample variance and comparing it to the theoretical variance of the distribution.

Here we plot the histogram and density distribution of variances of 40 random exponentials simulated over 1000 times.

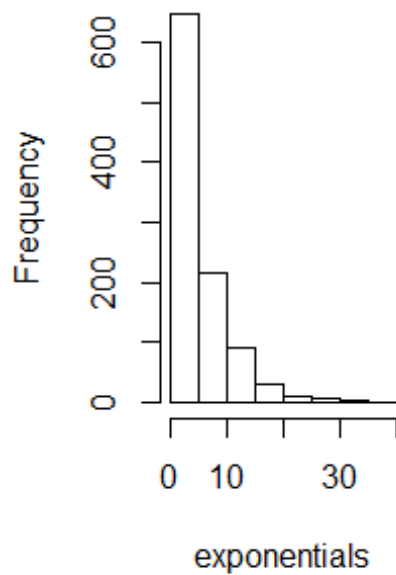


The above figures illustrate that the mean of distribution of a large collection of variances of 40 exponentials that comes out to be 25.2815371 is very close to the actual population theoretical variance which is 25.

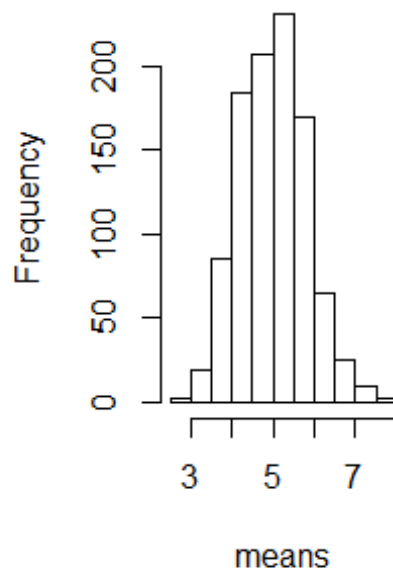
Analysing the nature of the distribution.

Here we plot the histogram of 1000 random exponential variables and compare it with the histogram of the averages of 40 random exponentials simulated over 1000 times.

Histogram of exponentials



Histogram of means



Here we see that although the original distribution of random exponentials is not normal, however the distribution of mean of averages assumes a normal bell shaped distribution. The same is true for the distribution of mean of variances.