Exponential distribution

Marija Zoldin

04.07.2015

Overview

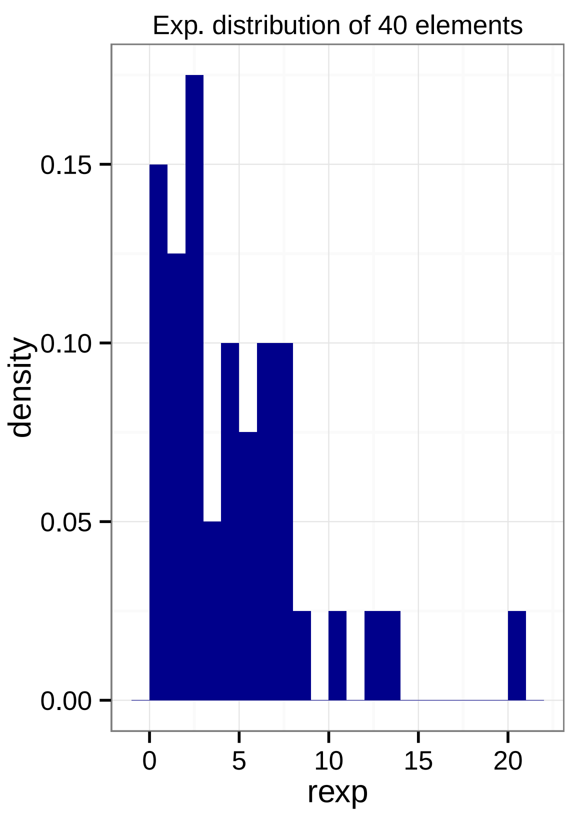
Exponential distribution is a probability distribution with density defined as :

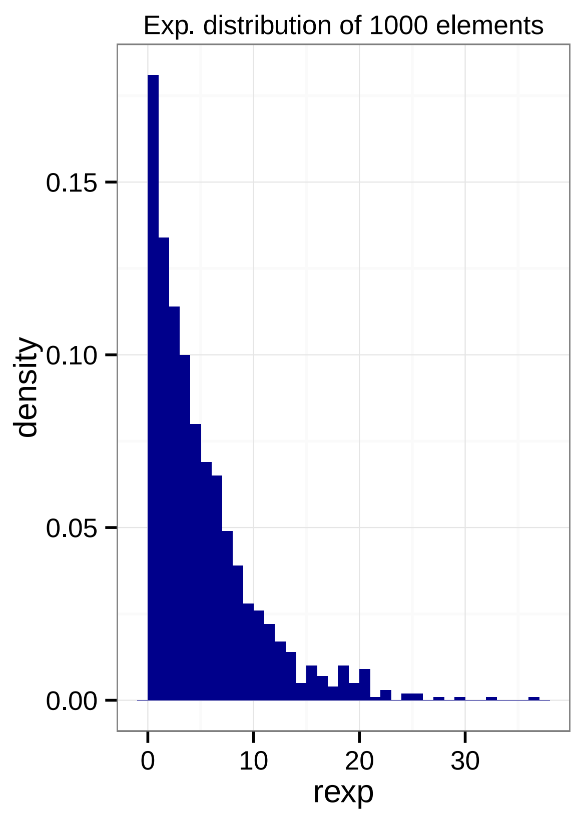
$$f(n) = \begin{cases} \lambda \times e^{-\lambda \times x} & \text{if } x >= 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Exponential distribution mean is defined as $\frac{1}{\lambda}$ and standard deviation (square root of the variance) is equal to the mean. As part of course project on Statistical inference Coursera class our task was to made a simulation of 1000 exponential distributions (each distribution contains 40 elements) and explore a relationship between sample means and population mean, sample variances and population variance. Our goal was to use during this assignment basic knowledge that we learned on this course – knowledge related to distribution of averages and central limit theorem.

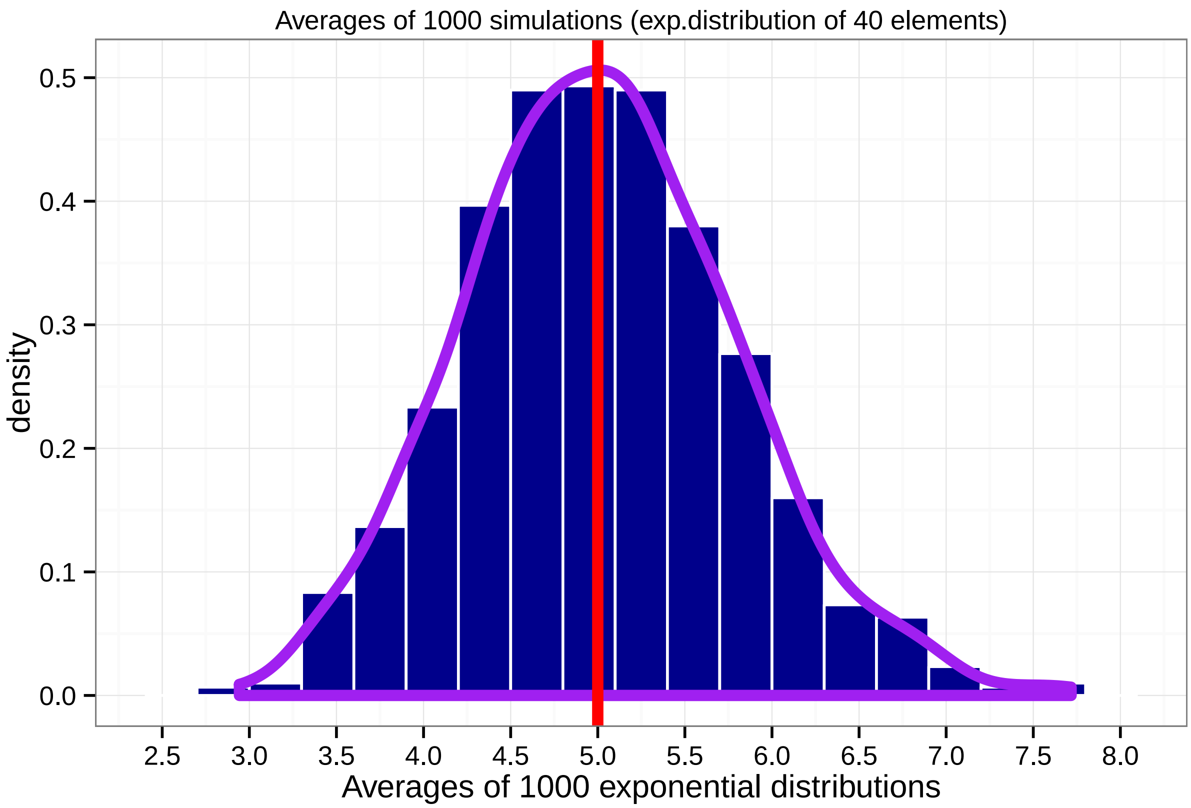
Exponential distribution simulation and central limit theorem

From the density distribution formula presented above it can be concluded that distribution is not normal. Two random exponential distribution of 40 and 400 elements were made via statistical language R and their distribution is presented on the graph below:



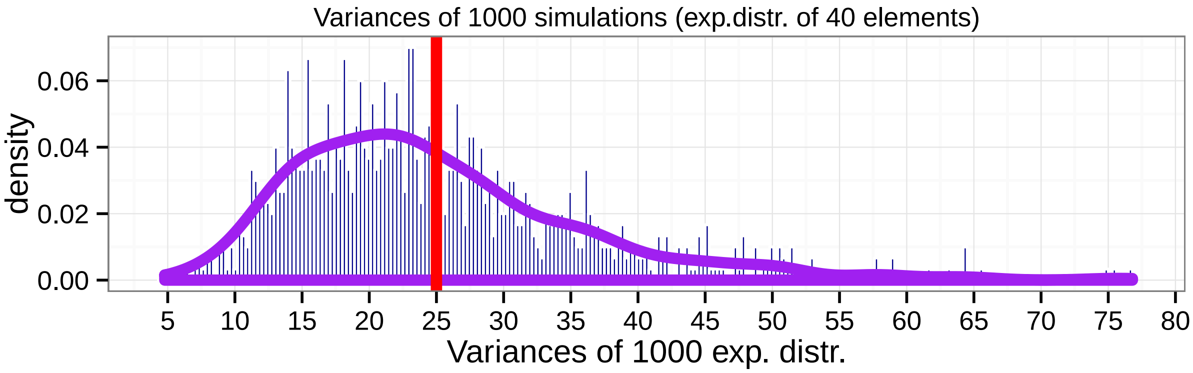


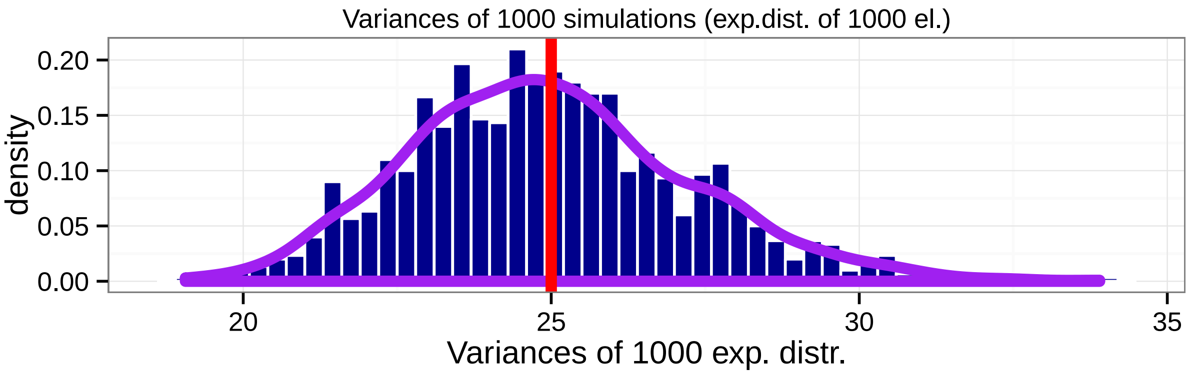
In the next step 1000 exponential distributions of 40 elements were simulated and for each of the 1000 distributions average values and variances were calculated. Because of the fact that we know that mean of averages has the distribution that is centered around population mean (which is $\frac{1}{\lambda} = 5$) we plotted distribution of the averages to convince ourselves in that fact:



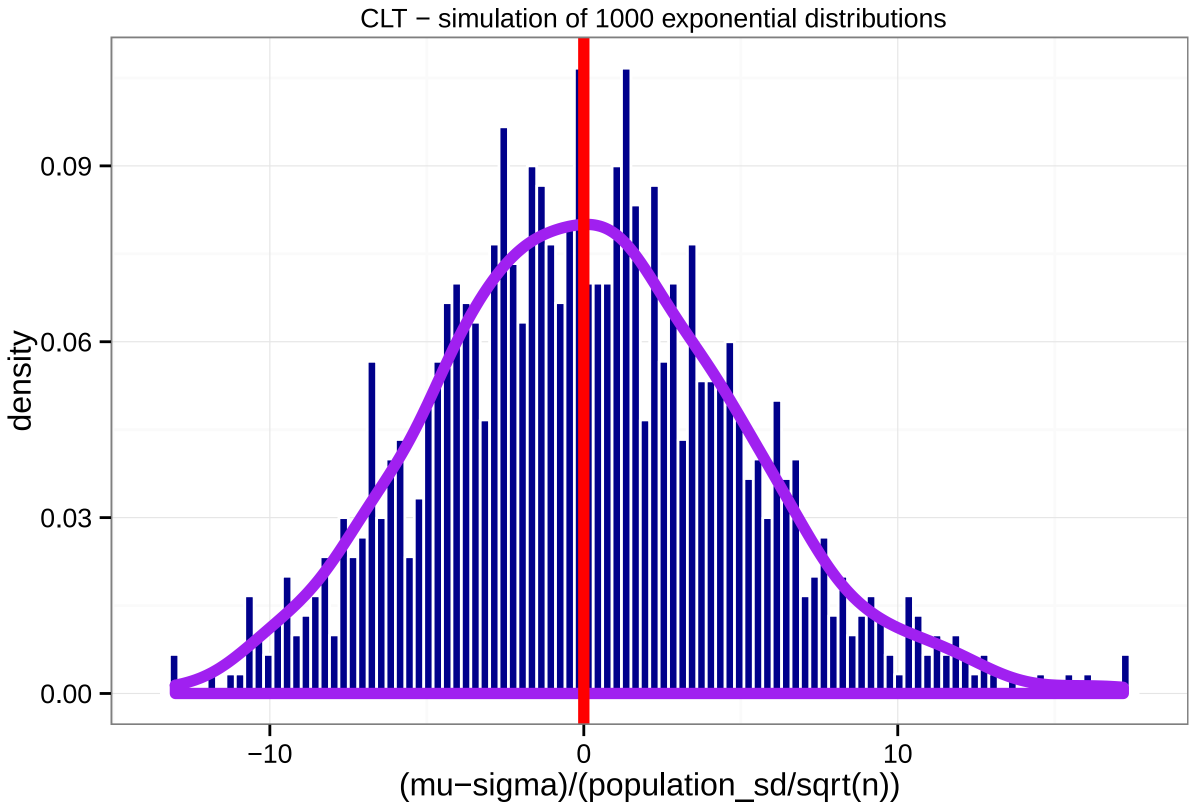
We can see from the plot above that the distribution of averages is normal, centered around 5 (which is population mean).

Also distribution of the variances was examined. It was expected that distribution of variances will be centered around population variance – for larger sample size estimation will be better. Because of that fact two simulation of 1000 distributions were made – one were each distribution contained 40 element and one for which each of the 1000 distributions contained 1000 elements. From the plot below it can be seen that with size of 1000 elements in each distribution estimation of the population variance was much better.





Central limit theorem was applied on the 1000 exp. distributions simulations with 40 elements. From each average value (of each distribution) population mean was subtracted and divided by standard error of the mean. By the central limit theorem distribution of calculated values needs to be standard normal distributed (with mean 0). From the plot below it can be noticed that distribution is std. normal with mean 0.



CLT − simulation of 1000 exponential distributions
density
0.00
0.03
0.06
0.09
−10
0
10
(mu−sigma)/(population_sd/sqrt(n))