Statistical Inference — Project – Part 1

The exponential distribution can be simulated in R with rexp(n, λ) where λ is the rate parameter. The mean of exponential distribution is 1/ λ and the standard deviation is also also 1/λ. **Set λ = 0.2 for all of the simulations.** In this simulation, you will investigate the distribution of averages of 40 exponential(0.2)s. Note that you will need to do a thousand or so simulated averages of 40 exponentials.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s. You should

1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.

   The theoretical mean of this distribution is located at $\mu=1/\lambda=5$ based on CLT. Figure 1 illustrates that the mean of simulation converges to the theoretical mean, i.e. $\mu=1/\lambda=5$, as the number of sample increases.

   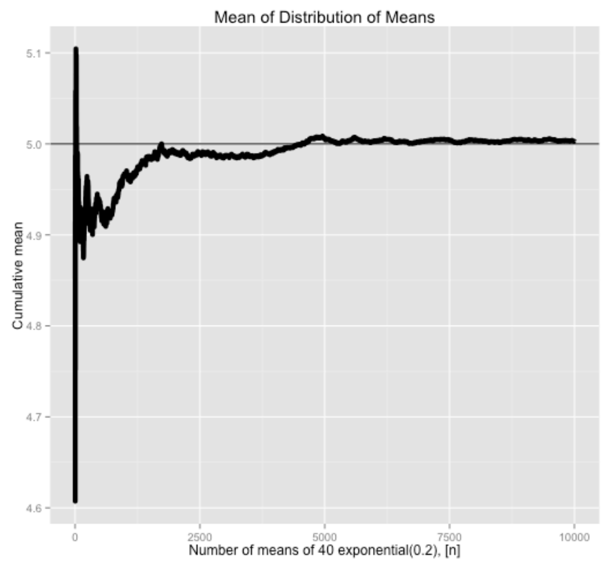

   Figure 1. Mean of averages of 40 exponential(λ=0.2) vs number of samples (n).

2. Show how variable it is and compare it to the theoretical variance of the distribution.

   The theoretical variance of this distribution is equal to σ^2/n of the exponential(0.2) with n equal to 40 based on CLT. Therefore, the theoretical variance equals $1/\lambda^2/n=0.625$. Figure 2 shows that how the variance of the simulation varies by the number of samples. As the number of samples increases, the variance of the simulation converges to the theoretical value, which is 0.625.

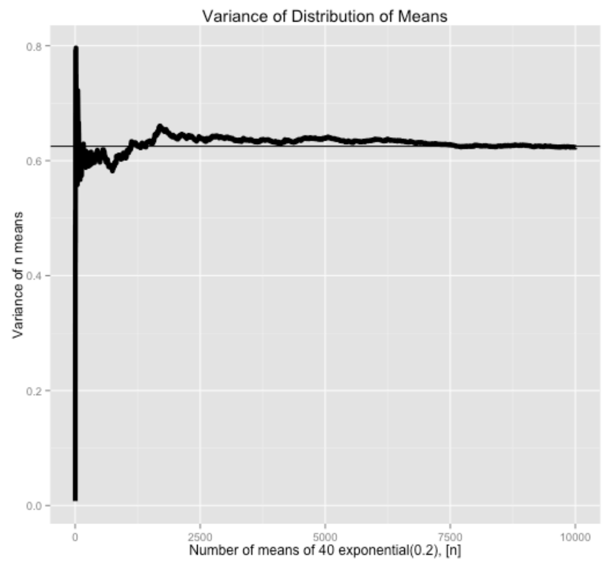

Figure 2. Variance of averages of 40 exponential(λ=0.2) vs number of samples (n).

3. Show that the distribution is approximately normal.

Theoretical approximation for this distribution is a normal distribution with mean equal to $1/\lambda=5$, i.e. the mean of exponential(0.2), and variance of $1/\lambda^2/n=0.625$, which is the variance of exponential(0.2) divided by n=40. Figure 3 displays the histogram and density of this simulation for 10000 samples and the theoretical estimate, which is $N(1/\lambda, 1/\lambda^2/n)$. It is seen that these two densities are very close.

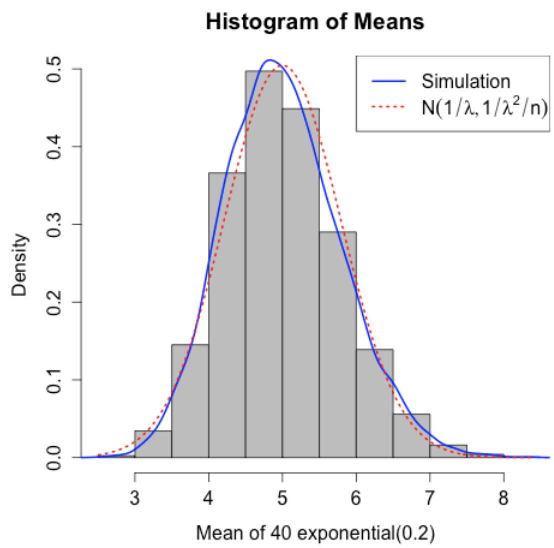

Figure 3. histogram and density of exponential(λ=0.2) for 10000 samples and the theoretical estimate, which is $N(1/\lambda, 1/\lambda^2/n)$.