Statistical Inference Course Project

Joshua Chua

Monday, August 17, 2015

Overview

In this project we investigate the exponential distribution in R and compare it with the Central Limit Theorem. We investigate the distribution of averages of 40 exponentials and illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials.
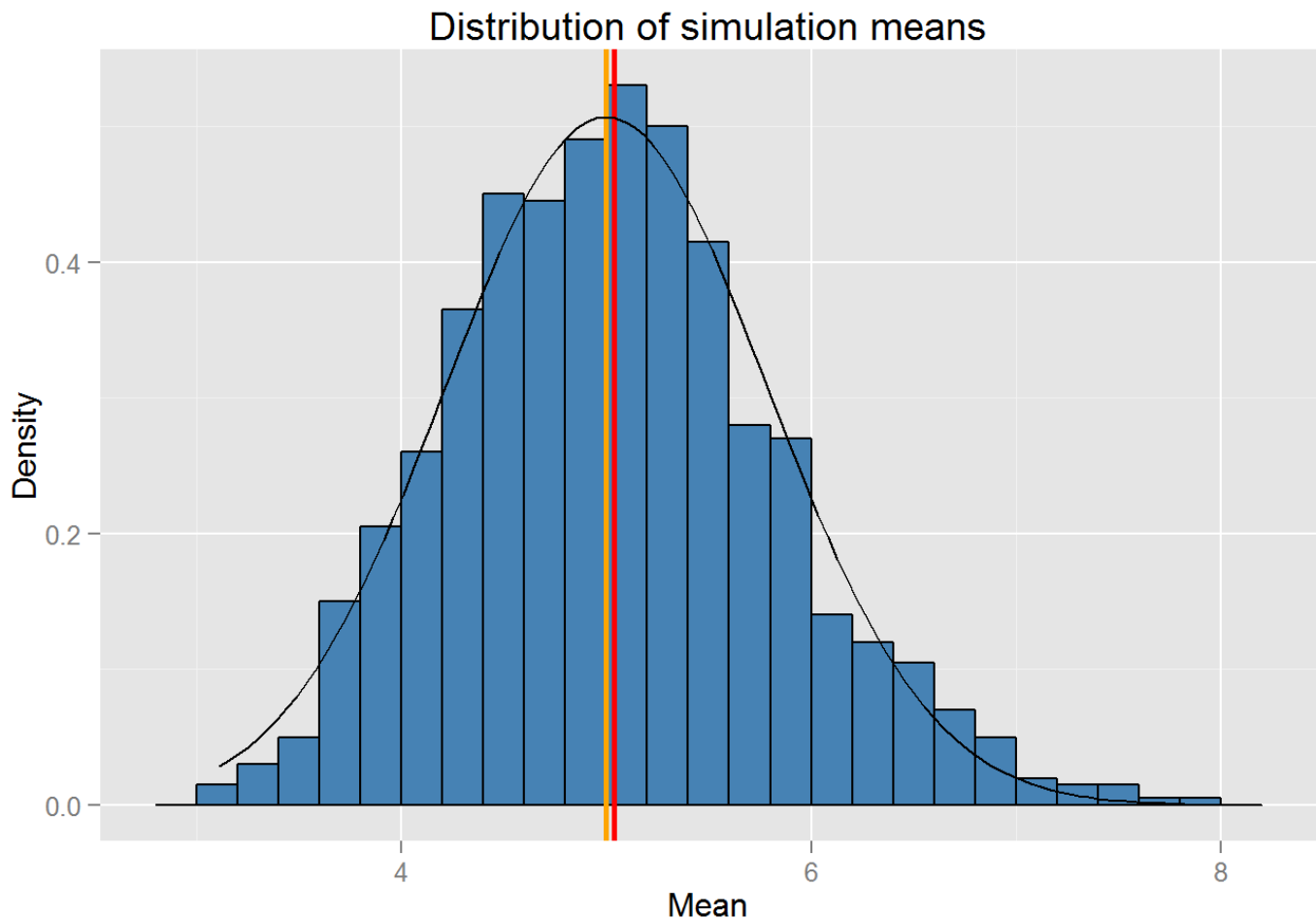
Simulations

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.

Results:

We plot the density function of our simulation means, and overlay it with a normal distribution we can see that the distribution of the simulation means is approximately normal, as the Central Limit Theorem exerts. Also included are the mean of our simulations (red line) and our expected mean (orange line),

and we can see they are pretty close.

Distribution of simulation means

Sample Mean versus Theoretical Mean

Question 1: Show the sample mean and compare it to the theoretical mean of the distribution.

The expected (theoretical) mean of an exponential distribution is 1/lambda, in our simulation case this would be 1/0.2 = **5** From our results we can see that the sample mean of the 1000 simulations is

```
## [1] 5.040009
```

which is pretty close to our expected mean.

Sample Variance versus Theoretical Variance

Question 2: Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

The theoretical variance of our simulation is (1/lambda)*(1/sqrt(n)))^2 = 0.625 Sample variance is

```
## [1] 0.6197137
```

We can see that the theoretical and sample variance are rather close.

Distribution

Question 3: Show that the distribution is approximately normal.

When compared with the actual distribution of the 1000 random exponentials, we can see that the distribution of a large collection of random exponentials is still exponential but the distribution of a large collection of averages of 40 exponentials is approximately normal.

Histogram of simulations