Simulation of Exponential Distribution

zhenyue zhu

April 24, 2016

Perform simulation

In this project, I 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. I this simulation, I set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials (n=40). The number of simulation to do is nosim=10,000.

```
set.seed(1234)
lam<-0.2
nosim<-10000
n<-40
```

Then I create exponential distributed n*nosim random numbers. Store them in a matrix form, with nuumber of rows is number of simulations (nosim=10,000) and number of columns is the sample size (n=40). Then I calculate the mean for each simulation. Store them in rmean variable. Then I find the mean and standard deviation for these 10,000 simulation results.

```
sample<-matrix(rexp(n*nosim, lam), nosim)
rmean<-apply(sample,1,mean)
sim_mean<-mean(rmean)
sim_sd<-sd(rmean)
mean(rmean)
```

```
## [1] 5.005646
```

```
sd(rmean)
```

```
## [1] 0.7941202
```

The simulated values for the mean and standard deviation is 5.0056459 and 0.7941202.

Theoretical prediction.

Now calculate the theoretical value the mean and standard deviation.

```
1/lam
```

```
## [1] 5
```

```
1/lam/sqrt(n)
```

```
## [1] 0.7905694
```

```
t_mean<-1/lam; t_sd<-1/lam/sqrt(n)
```

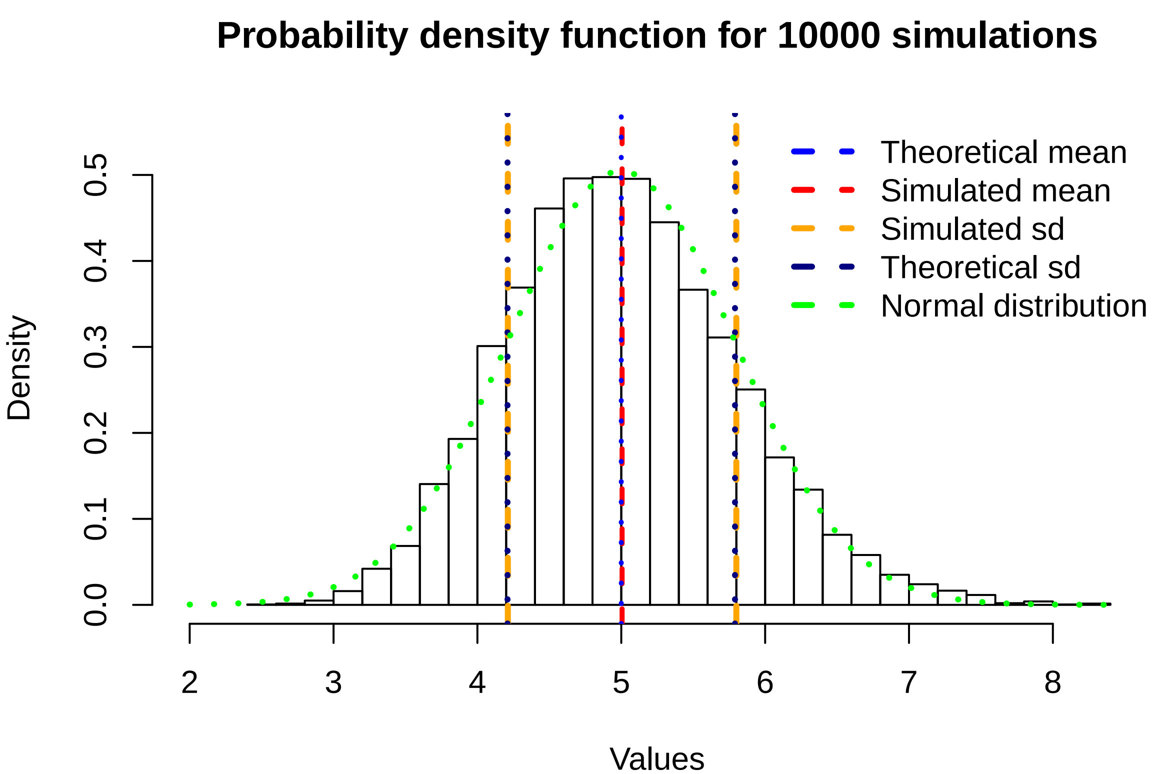
The theoretical value the mean and standard deviation are 5 and 0.7905694. As we can see that the simulation results match the theoretical value very perfectly. Thus confirms the central limit theorem.

For the variance, which is the square of the standard deviation, which is also similar for both simulation and theoretical value, since the standard deviation is almost equal. I will plot the histogram figures and only labeled the mean and standard deviation (instead of variance) for comparison, since mean and standard deviation have the same units.

Histogram of the distribution

You can also make plots for the density plot use histogram function. In this figure I plot the density for the simulation.

```
myhist <- hist(rmean, freq = FALSE, xlim = c(2, 8.5), ylim = c(0, .55), breaks=30,
main=paste("Probability density function for",nosim,"simulations"), xlab = "Values")
abline(v = sim_mean, col = "red", lwd = 2.5, lty = 2)
abline(v = 5, col = "blue", lwd = 2.5, lty = 9)
abline(v = sim_mean+sim_sd, lwd=3, lty = 2, col="orange")
abline(v = sim_mean-sim_sd, lwd=3, lty = 2, col="orange")
abline(v = t_mean+t_sd, lwd=3, lty = 9, col="navy")
abline(v = t_mean-t_sd, lwd=3, lty = 9, col="navy")
x <- seq(min(rmean), max(rmean), length = 100)
curve(dnorm(x, mean = t_mean, sd =t_sd), col = "green", lwd = 3, lty=3, add = TRUE)
legend('topright', c("Theoretical mean", "Simulated mean", "Simulated sd",
"Theoretical sd", "Normal distribution"), lty=2, col=c('blue','red',"orange",
"navy", "green"), bty='n',lwd=3)
```



The mean is labeled on the figure. The blue vertical line is the expected mean or theoretical of the mean value.

The red vertical line is the simulated mean. They are almost identical. I further label the one standard deviation for both the simulation and theoreitcal values in the histogram plot. Just like the mean, they almost overlap. Further confirm the validity of CLT theory.

I also plot the normal distribution density on the same figure, where I use the mean and the standard deviation from the theoretical value. As we can see the normal distribution density on this figure captured the simulated experiments. Thus the simulated distribution is approximately normal.