Stat Simulation

Yanyu

October 12, 2017

Simulation Assignment

Requirement

In this project you 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. Set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Summary

Compared with Central Limit Theorem, the distribution I got from exponential distribution is almost identical.

Research

Initialize simulation set

```
set.seed(102)
lambda = 0.2
set = 40
n = 1000
simulation_set <- matrix(rexp(n*set,lambda), set)
```

Calculte the mean and the variance / standard deviation of the rows

```
row_mean <- apply(simulation_set, 1, mean)

sim_mean <- mean(row_mean)
sim_sd <- sd(row_mean)
sim_var <- sim_sd^2

theo_mean = 1/lambda
theo_sd = (1/lambda)/sqrt(n)
theo_var =  (1/lambda)^2/n

print_result <- matrix(c(theo_mean, theo_sd, theo_var, sim_mean, sim_sd, sim_var), nrow = 3, nco
l = 2)
dimnames(print_result) =list(c("mean","standard deviation", "variance"),c("theoratical", "simula
tion"))
print(print_result)
```

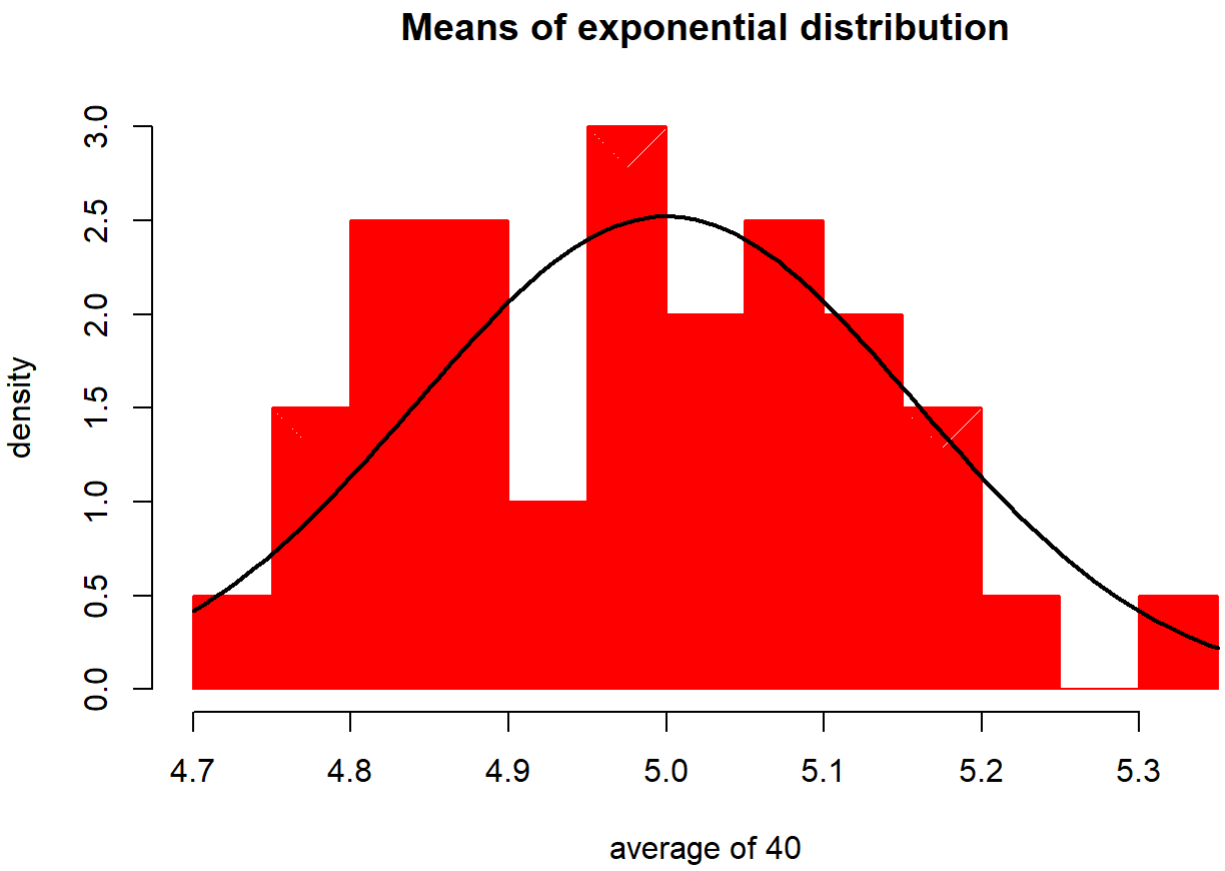
```
##                    theoratical simulation
## mean                 5.0000000 4.98155916
## standard deviation   0.1581139 0.14344280
## variance             0.0250000 0.02057584
```

The differences between the theoratical characteristics of the distribution and the simulation seems to be negligible. If mean is rounded on three figures, it is only .01 off. The standard deviation and the variance are even less off (.004 and .007).

Plotting

In this step the average of the means of each row is plotted.

```
hist(row_mean, density=100, breaks=20, prob=TRUE, col = "red", xlab="average of 40", ylab = "den
sity", main="Means of exponential distribution")
curve(dnorm(x, mean=theo_mean, sd=theo_sd), col="black", lwd=2, add=TRUE, yaxt="n")
```



In the bars the mean is given. In black the normal function based on the theoratical characteristics (mean and standard deviation) is printed. The distribution based on the plot looks normally distributed and this means that the Central Limit Theory is proven.