Simulation Project

Alper Halbutogullari, November 10, 2015

Overview

In this project we will investigate the exponential distribution in R and compare it with the CTL (Central Limit Theorem). The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` λ is the rate parameter. The mean of exponential distribution is $1/\lambda$ and the standard deviation is also $1/\lambda$. For this simulation, we set $\lambda = 0.2$ for all of the simulations. In this simulation, we will investigate the distribution of averages of 40 exponentials.

Our goals are to: 1. compare the sample mean with the theoretical mean of the distribution. 2. compare the sample variance with the theoretical variance of the distribution 3. show that the distribution is approximately normal.

Simulations

```r
# set the seed for reproducibility
set.seed(2015)

# rate parameter
lambda <- 0.2

# 1000 simulations with 40 samples each
sample_size <- 40
number_of_simulations <- 1000

sampled_exponential_means <- replicate(number_of_simulations, mean(rexp(sample_size, lambda)))
```

Analysis

Sample mean vs theoretical mean

```r
theoretical_mean <- 1/lambda
sample_mean   <- mean(sampled_exponential_means)
paste("theoretical_mean=",theoretical_mean,"   sample_mean=",sample_mean)
```

```
## [1] "theoretical_mean= 5    sample_mean= 5.0115634332513"
```

As expected they are in line. Let's see the theoretical_mean over the distribution.

```r
hist(sampled_exponential_means,
    freq=TRUE,
    breaks=50,
    xlab="Sample Means from 1000 Simulations")
abline(v=5, col="red", lwd=4)
```
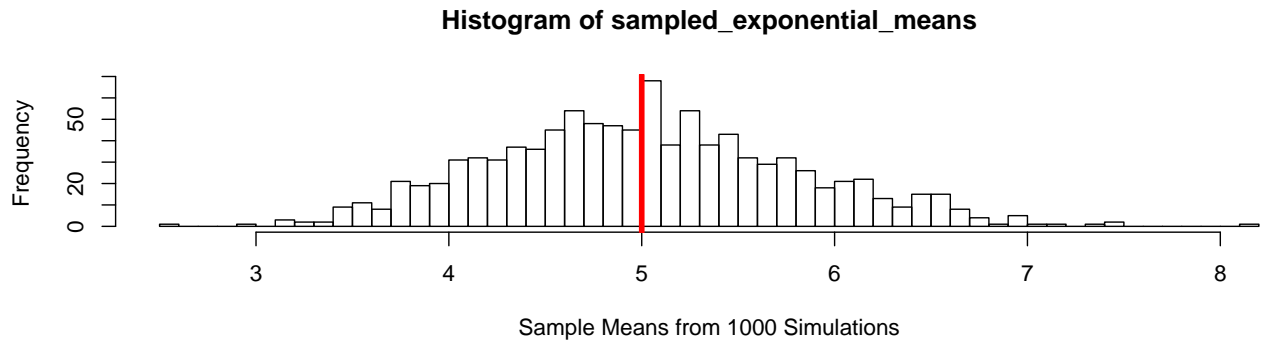
Histogram of sampled_exponential_means

Figure 1: Histogram of sampled exponent means

Red line shows the theoretical mean, which is in line with the expected value.

Sample variance vs theoretical variance

```
theoretical_variance <- (1/lambda)^2 / 40
sample_variance <- var(sampled_exponential_means)
paste("theoretical_variance=",theoretical_variance,"   sample_variance=",sample_variance)
```
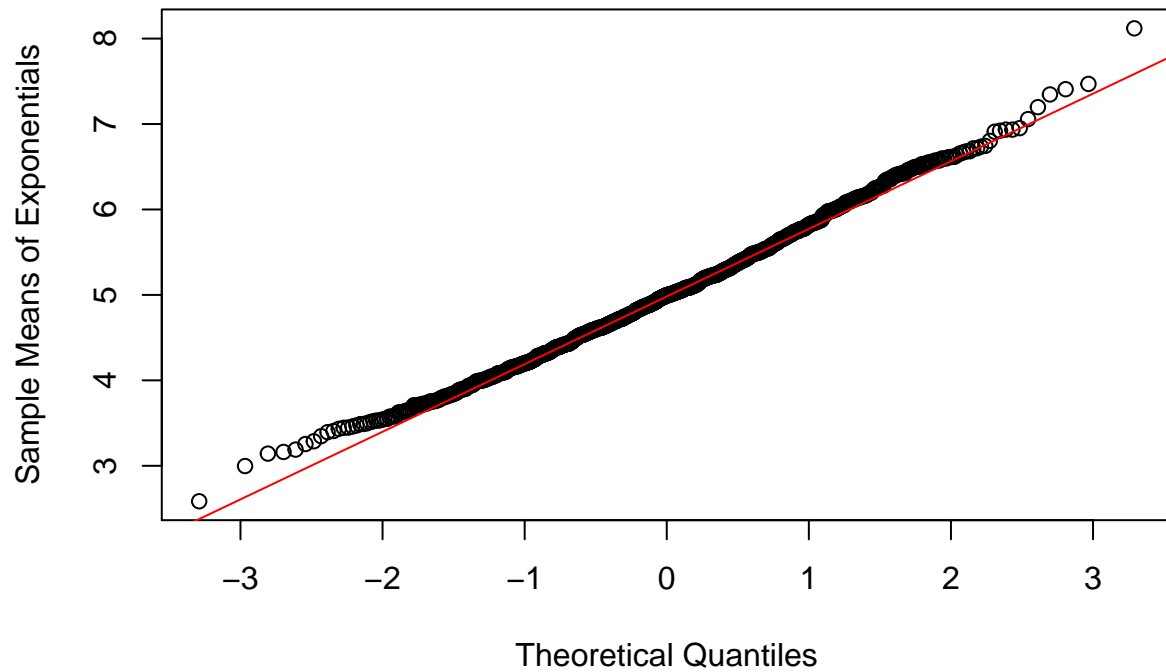
```
## [1] "theoretical_variance= 0.625    sample_variance= 0.626299229301796"
```

As expected the results are in the same ballpark.
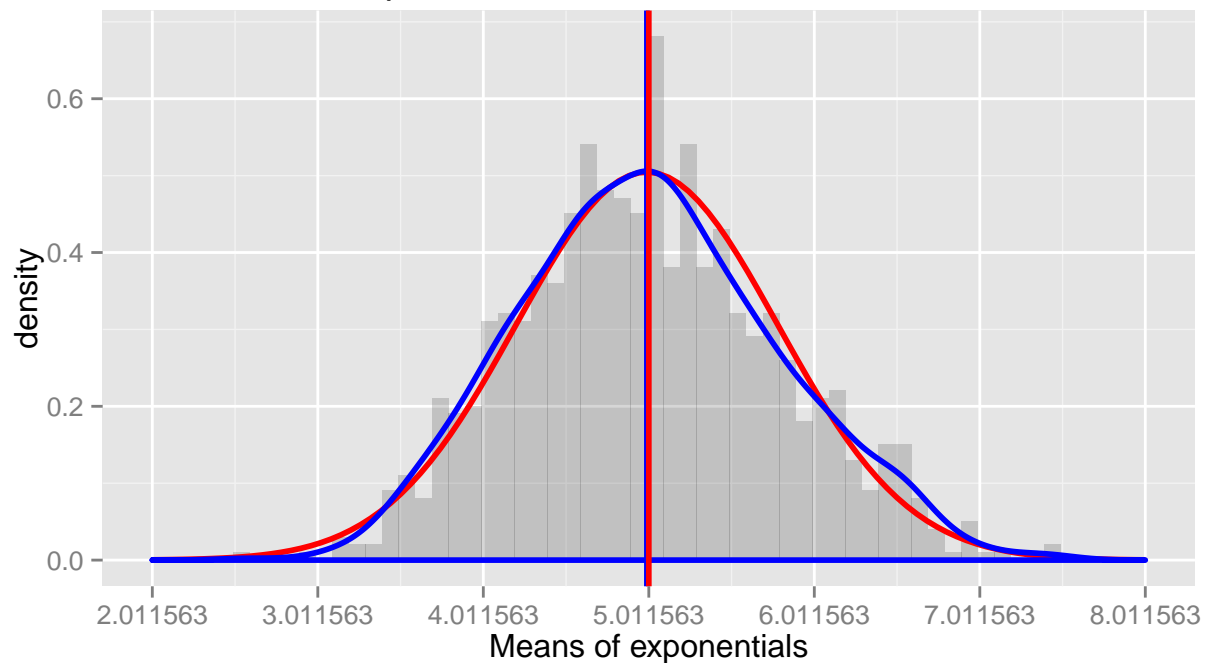
Sample distribution vs normal distribution

```
qqnorm(sampled_exponential_means, ylab = "Sample Means of Exponentials")
qqline(sampled_exponential_means, col = 2)
```

Normal Q-Q Plot

As can be seen above, the distribution is approximately normal.

Sample distribution vs normal distribution

Conclusion

The sampled_exponential_means distribution is very close to the theoretical normal distribution (with the theoretical mean and variance).