Coursera Statistical Inference - Simulation

pigmy

22 January, 2015

Contents

GitHub: https://github.com/perplexedpigmy/statinf-010

RPub: http://rpubs.com/perplexedpigmy/54934

Overview

We will perform an exponential distribution simulation and compare it to the CLT(Centeral Limit Theorem). The excercise we are undertaking is already restricted to specific values that we are documented in the code itself.

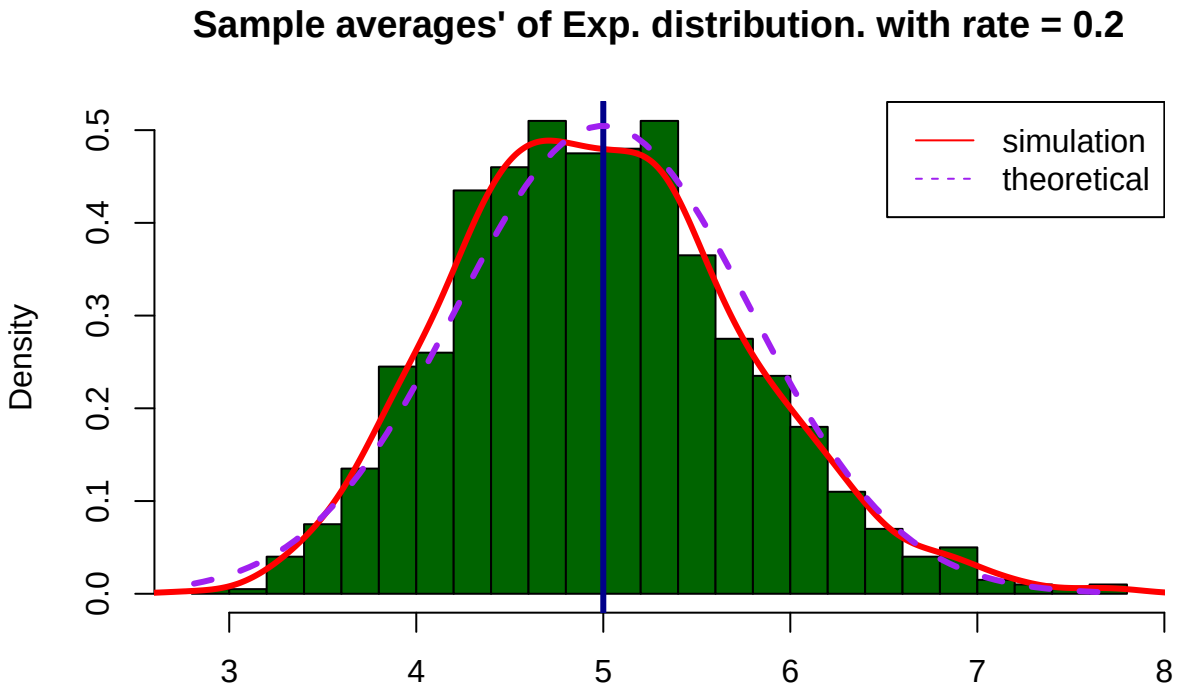
Simulations

As instructed we are going to perform 1000 simulations of 40 samples each, where `lambda = 0.2`

```
set.seed(1234)      # Just a seed
lambda <- 0.2       # Given by the exercise
sample.size <- 40   # Given by the exercise
sims <- 1000        # Suggested by the exercise
d.mean <- 1/lambda  # Distribution mean. Inherent property
d.sd   <- 1/lambda  # Disribution standard deviation. Inheret property

exp.sim <- matrix(rexp(sims*sample.size, rate=lambda), sims, sample.size)
means <- rowMeans(exp.sim)
```

Analysis



Sample mean vs. theoretical mean

```
sample.mean <- mean(means)
theoretical.mean <- 1/lambda
```

The deviation of the sample mean `4.9742388` from the theoretical mean `5` is approximately `0.5`%

2. Sample variance vs. Theoretical variance

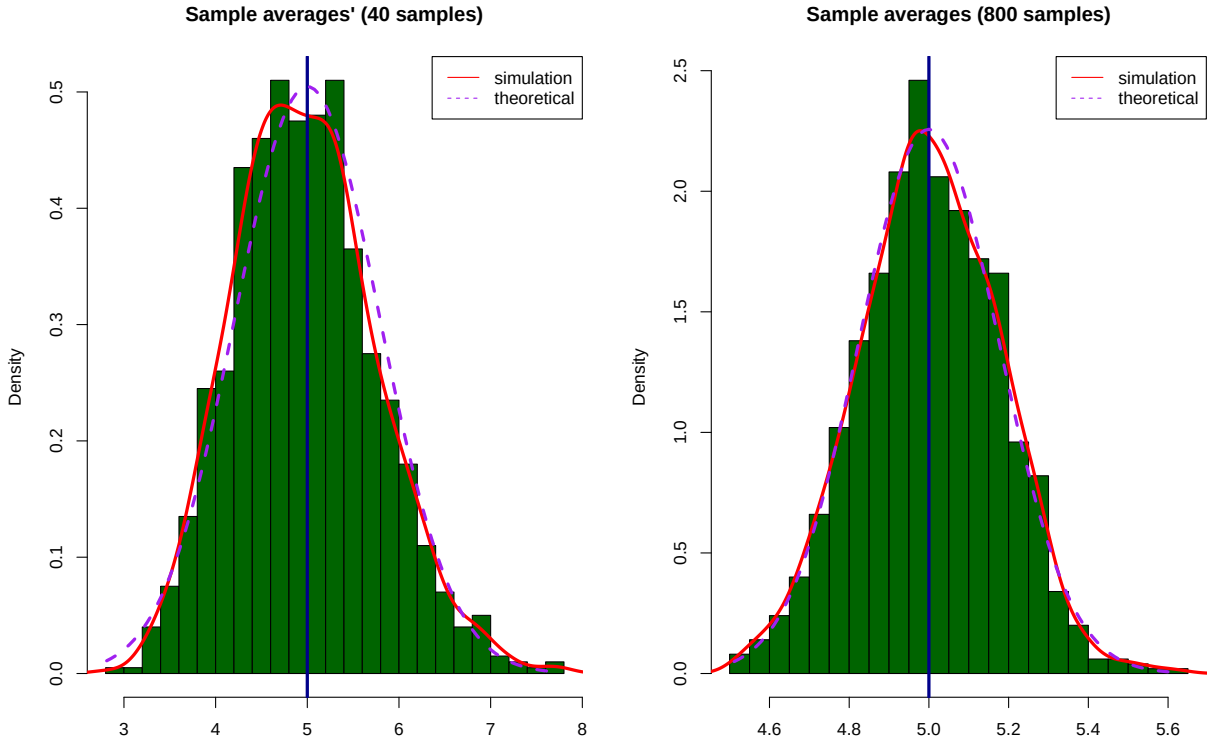
```
sample.sd <- sd(means)
theoretical.sd <- (d.sd)/sqrt(sample.size)

sample.var <- var(means)
theoretical.var <- 1 / (lambda ** 2 * 40)
```

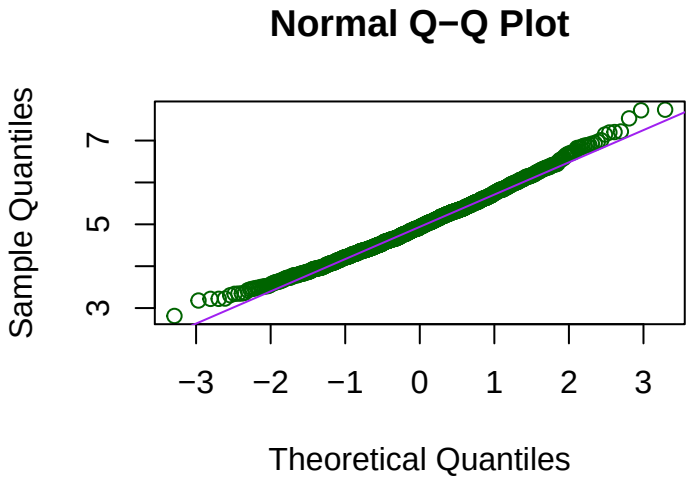
The sample variance `0.5949702` is different from the theoretical variance `0.625` by `4.8`%

and the sample standard deviation `0.7713431` diffrantiates from the theoretical standard deviation `0.7905694` by `2.4`%

3. Is the distribution normal?



Notice how the the averages of 800 samples from the exponential distrbution converges to the mean vis-a-vis the the averages of 40 samples, it is also visibly closer to the normal distribution.



Q-Q Plot

A Q-Q plot indicates as well that the 40 exponentials are quite close to the normal distribution.