Statistical Inference - Simulation

Ajani Gyasi

25 October, 2015

Overview

In this project, we will examine whether an exponential distribution is comparable to the Central Limit Theorem. The exponential distribution will simulate 1000 trails with a sample size of 40, while a lambda rate is set to 0.2. As we investigate the distribution of averages, we will conclude whether its comparability satifies the Central Limit Theorem.

Simulations

We are doing a thousand simulations of the average (**means**) of 40 samples drawn from the exponential distribution.

```
sim <- 1000 #number of simulations
lambda <- 0.2 #lambda rate
n <- 40 #number of samples
m_exp <- 1/lambda #exponential mean
sd_exp <- 1/lambda #exponential standard deviation


set.seed(348) #seed for reporoducibility
dat_simulation <- matrix(rexp(n * sim, lambda), sim) #calculation
means <- rowMeans(dat_simulation) #simulation
```

Sample Mean vs. Theoretical Mean

The sample mean or the simulated mean is 5.035, while the theoretical mean is 5. Notice, that this gives us a close estimate of the population mean.

```
## [1] "Sample mean: 5.035"
```

```
## [1] "Theoretical mean: 5"
```

Sample Variance vs. Theoretical Variance

The sample variance or the simulated variance is 0.634, while the theoretical variance is 0.625. Again, these numbers are fairly close.
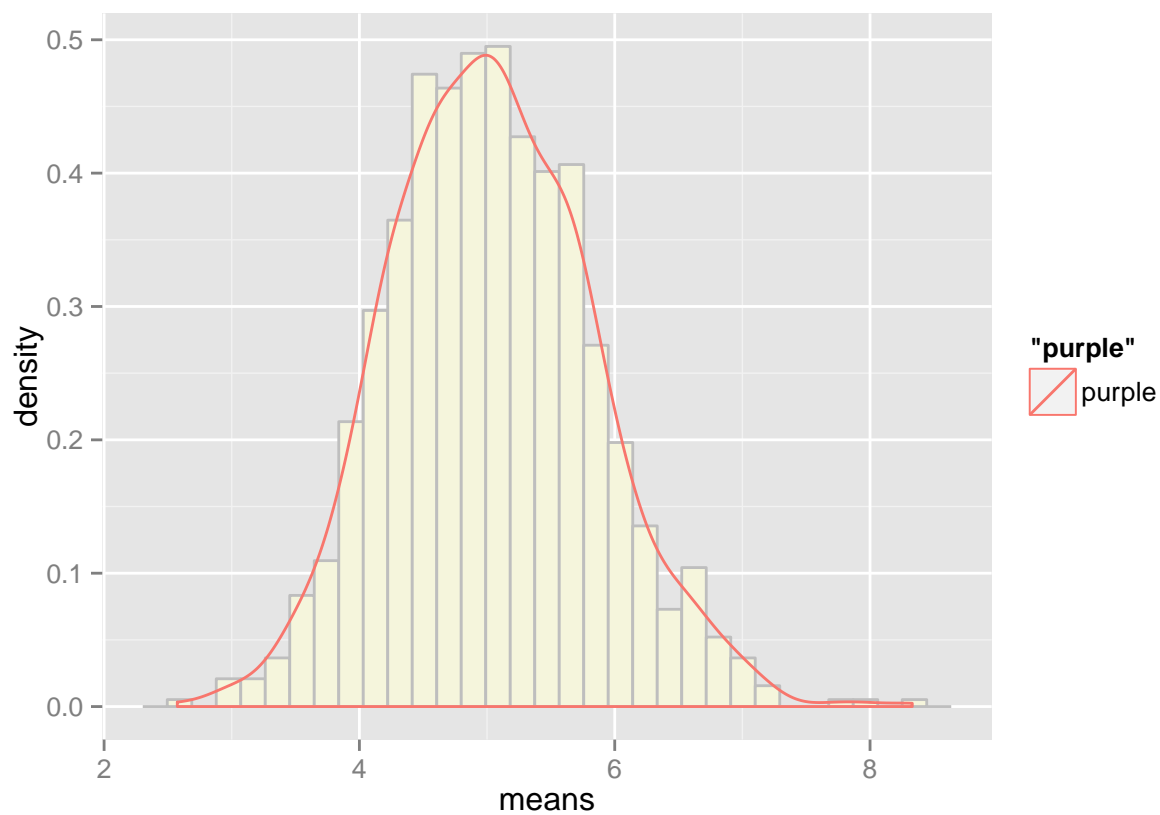
```
## [1] "Sample variance: 0.634"
```

```
## [1] "Theoretical variance: 0.625"
```

Distribution

The diagrams below showcase the exponential distribution being approximately normal. The Central Limit Theorem tells us that the averages of samples will follow a normal distribution. The two histograms prove the sample density follows a normal distribution. The qqplot double checks whether our simulation is in agreeance with the Central Limit Theorem.

Distribution of sample averages
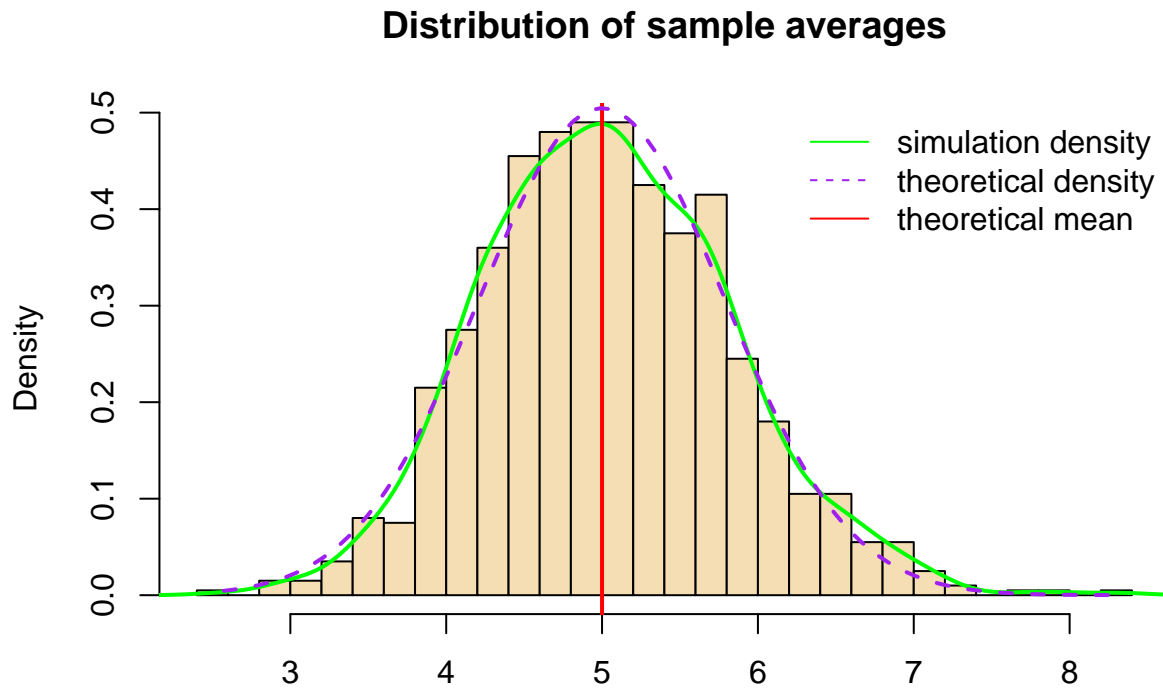

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```

Normal Q–Q Plot