Analysis of the Exponential distribution using simulation

Chris Shaw

3 April 2016

Question

Remove this section in final report

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.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. You should

Show the sample mean and compare it to the theoretical mean of the distribution. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. Show that the distribution is approximately normal. In point 3, focus on the difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials.

Overview

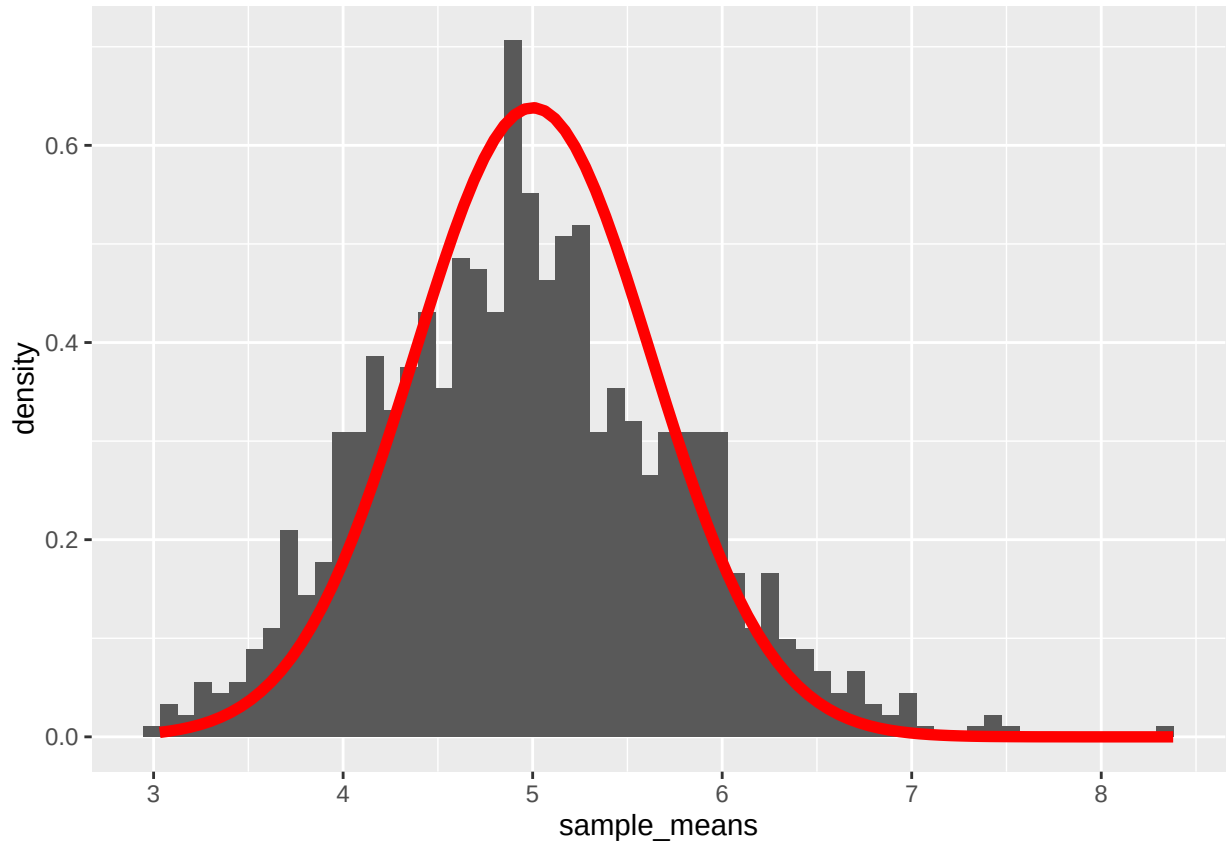
In a few (2-3) sentences explain what is going to be reported on.

Simulations

Include English explanations of the simulations you ran, with the accompanying R code. Your explanations should make clear what the R code accomplishes.

```
# set up parameters of the exponential distribution
lambda <- 0.2
sample_size <- 40
num_sims <- 1000

mean <- 1/lambda
std <- 1/lambda

# Ensure simulations are reproducible
set.seed(12345)

# create a matrix of simulation data (rows = simulations)

sim_data <- matrix(rexp(sample_size*num_sims, lambda), nrow = num_sims)
```

```r
sample_means <- as.data.frame(apply(sim_data, 1, mean))

ggplot(data=as.data.frame(sample_means),aes(x=sample_means)) +
        geom_histogram(bins=60, aes(y=..density..)) +
        stat_function(fun = dnorm,
                        args = list(mean=1/lambda, sd=1/lambda^2/sample_size),
                        color="red", lwd=2) +
        scale_x_continuous()
```



Sample Mean versus Theoretical Mean

Include figures with titles. In the figures, highlight the means you are comparing. Include text that explains the figures and what is shown on them, and provides appropriate numbers.

Sample Variance versus Theoretical Variance

Include figures (output from R) with titles. Highlight the variances you are comparing. Include text that explains your understanding of the differences of the variances.

Distribution

Via figures and text, explain how one can tell the distribution is approximately normal.