Statistical Inference - Peer Assessment

Cleosson Souza

27 de setembro de 2015

Investigating the Exponential Distribution using R and compare it with the Central Limit Theorem

Introduction

The goal of this analysis is to explore the exponential distribution using R. To begin, we are going to do one thousand simulated averages of 40 exponentials. We will look at the mean distribution of these exponentials, illustrate their variability, and explore their similarity to the normal distribution.

For reference, the exponential distribution can be simulated 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. We will use lambda = 0.2.

```r
library(ggplot2)

set.seed(5)
lambda <- .2 # rate
samples <- 40
simulations <- 1000
s_means <- NULL

##Get one thousand mean observations from 40 expotentials
for (i in 1:simulations) { s_means <- c(s_means, mean(rexp(samples, lambda)))}

# sample mean
s_mean <- round(mean(s_means), 2)

# theoretical mean
t_mean <- round(1 / lambda, 2)

# sample sd
s_sd <- round(sd(s_means), 2)

# sample variation
s_var <- round(var(s_means), 2)

# theoretical sd
t_sd <- round((1 / lambda) * (1 / sqrt(samples)), 2)

# theoretical var
t_var <- round(t_sd ^ 2, 2)
```

|	Mean	Standard Deviation	Variance
Simulation	5.04	0.78	0.6
Theoretical	5	0.79	0.62

Conclusion: The mean of the sample distribution is approximately equal to the theoretical mean. The variance of the sample distribution is approximately equal to the theoretical variance. We can see from the results that the sample average mean, standard deviation and variance are all very close to the expected values for the exponential distribution.

Mean of 40 exponentials and it's relation to the normal distribution

The plot below shows the normal distribution. The dashed line is the mean.

Normal Distribution
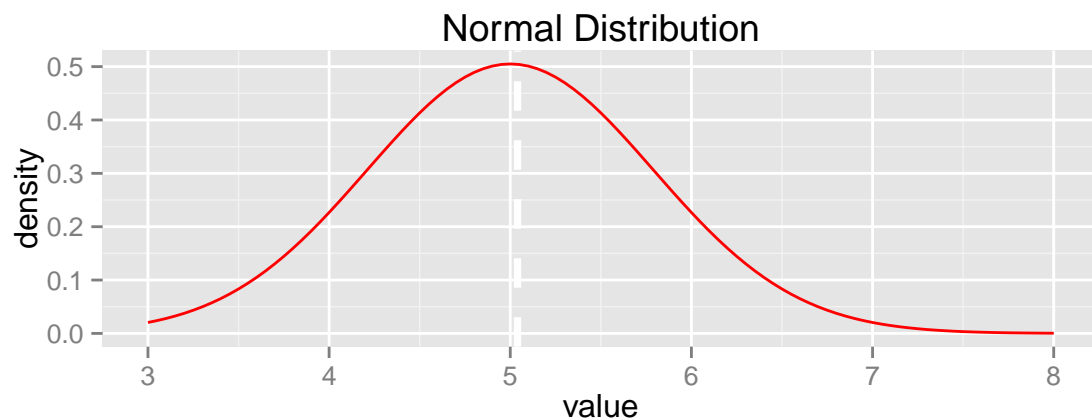
The plot below shows the sample mean distribution. The dashed line is the mean.

Sample Mean Distribution
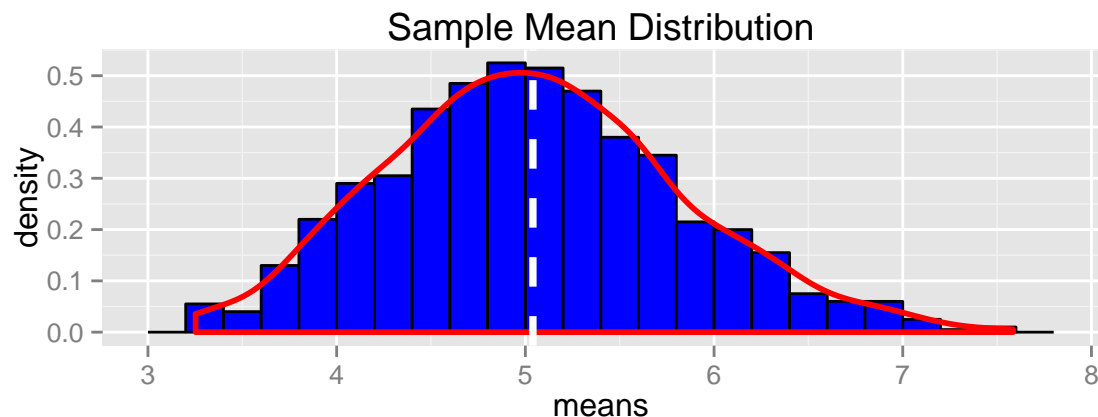
The plots demostrate that the sample mean distribution is approximately normal. Use the red line.

The plot below shows the exponential distribution of one thousand. The dashed line is the mean.

Exponential Distribution
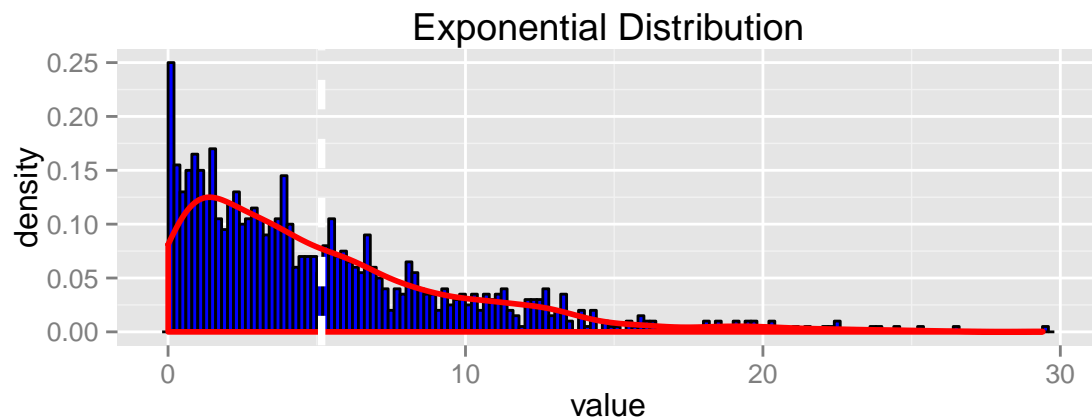
The plots demostrate that the exponential distribution does not look like a normal distribution