Statical Inference Project Work - Question 1

Hardeep

25 April 2015

Overview

This is project work for statistical inference course, in this project we would simulate the exponential distribution 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 also `1/lambda`.

Simulations

We'll set `lambda = 0.2` for all of the simulations. In this simulation, we will investigate the distribution of averages of 40 `exponential(0.2)`s. Note that we will need to do a thousand or so simulated averages of 40 exponentials, and will generate those as follows:

```
set.seed(111527)
lambda <- 0.2
sample_size <- 40
num_simulations <- 1000

sim_data <- matrix(data=NA,nrow=0,ncol=sample_size)

for (i in 1:num_simulations) {
    sim_data <- rbind(sim_data, rexp(sample_size, lambda))
}
```

Now, we will illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 `exponential(0.2)`s.

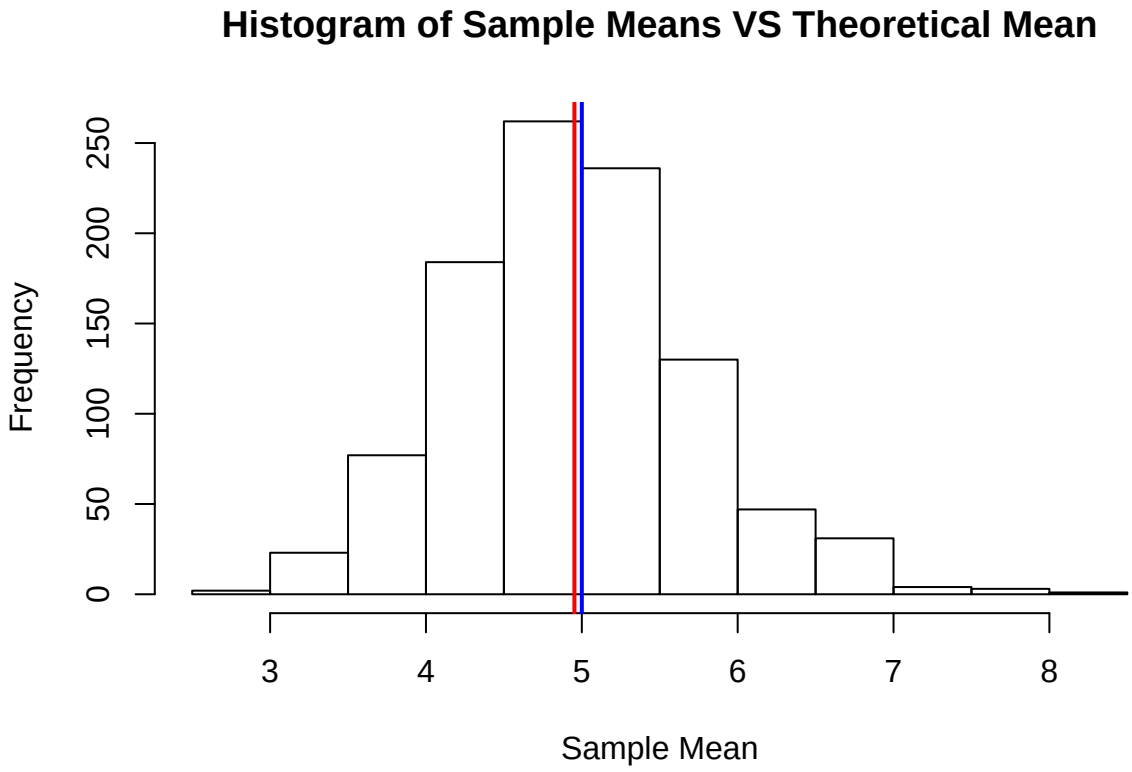
1. Sample Mean versus Theoretical Mean

As stated in the introductory paragraph to this section, the theoretical mean of the distribution is `1/lambda`, which, for a `lambda` of **0.2**, equates to **5**.

First, we'll calculate the mean of these sample means to see how closely it estimates the theoretical value:

```
sample_means <- c(1:num_simulations)
for (j in 1:num_simulations) {
    sample_means[j] <- mean(sim_data[j,])
}
mean_sample_means <- mean(sample_means)
```

So, as we can see, compared to the theoretical mean of **5**, the sample mean of **4.9525193** and we see the sample mean estimates theoretical mean nicely. We can also see this in the following figure where the red line denotes the sample mean and the blue line denotes theoretical mean



2. Sample Variance versus Theoretical Variance

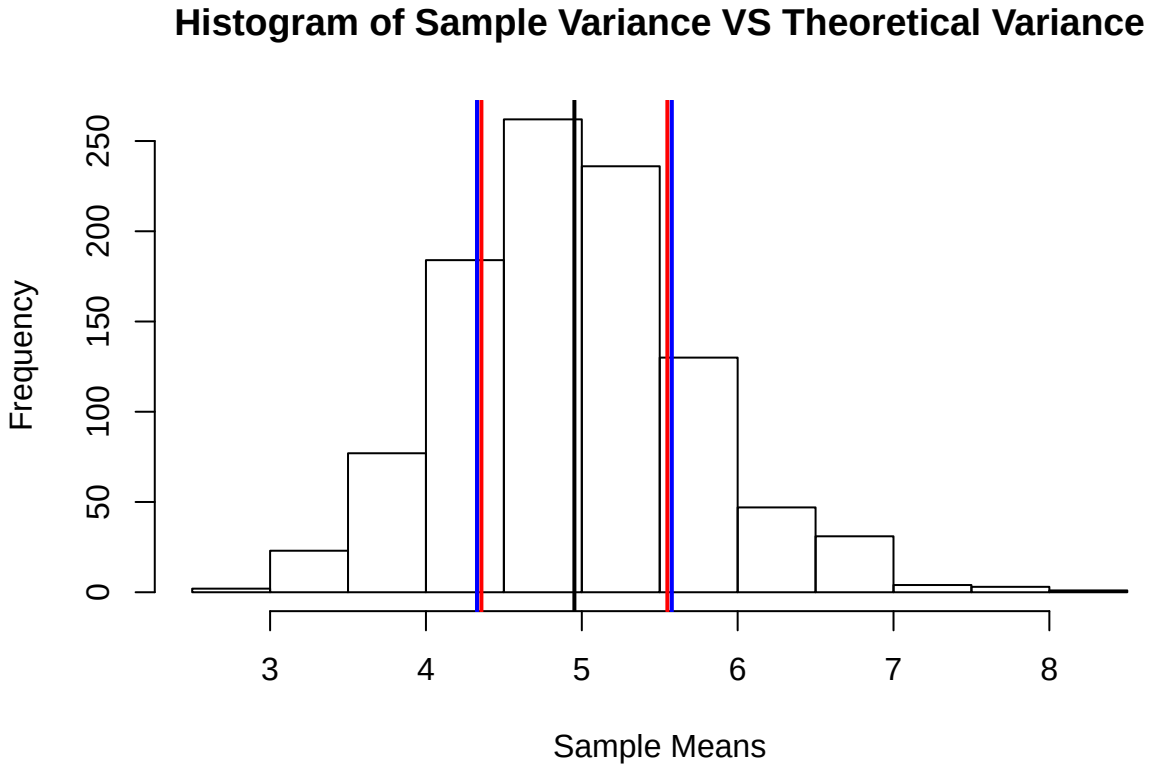
Lets first compute the theoretical variance first

```
pop_variance <- (1/lambda*(1/sqrt(sample_size)))^2
```

To get the observed variance, we'll calculate the variance of the sample means to see how closely it estimates the theoretical variance:

```
sample_variance <- var((sample_means))
```

So, as we can see, compared to the theoretical variance of **0.625**, the sample variance of **0.5966654** lines up quite closely. We can also see this in the following figure where the red line denotes the sample variance and the blue line denotes theoretical variance



3. Show that the distribution is approximately normal.

If we plot a curve with the mean and standard deviation of sample it comes out as normal curve. See figure below