Assignment 1 exercise 1

Mike Vosskuhler

August 30, 2018

Simulation excercise

Exercise summary:

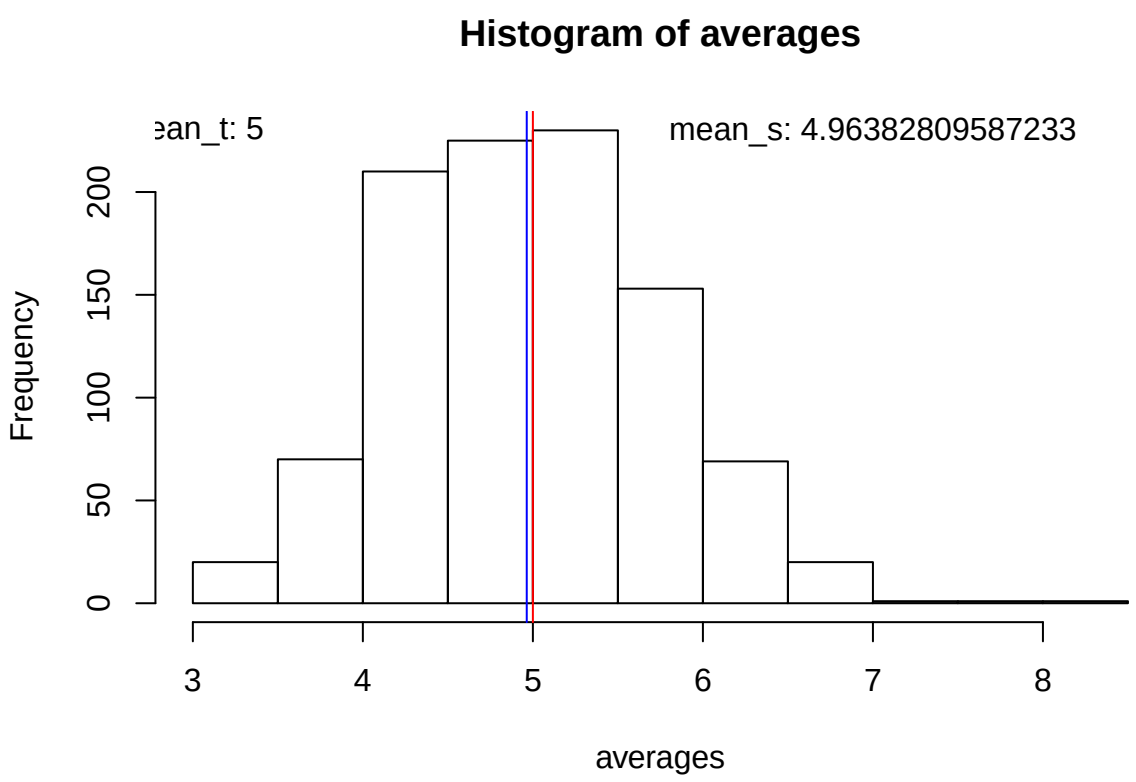
For this exercise we will generate some data using the rexp() function. This function generates random samples from an exponential distribution with a mean of 1/lambda and a standard deviation of 1/lambda. For this experiment we will use a lambda of 0.2. Using the rexp() function we will generate a sample containing 40 randomly generated values. Furthermore, we will use the set.seed() function to allow for reproducibility of the results. The data-dataset will contain 1000 of these samples

```
set.seed(20)
data <- rexp(40000, 0.2) %>% matrix(1000, 40)
```

Sample Mean versus Theoretical Mean:

We will first compare the sample mean to the theoretical mean, recall that the theoretical mean for the exponential distribution is 1/lambda. So, our theoretical mean = 1/0.2 = 5. The plot will show the mean of the averages of samples:

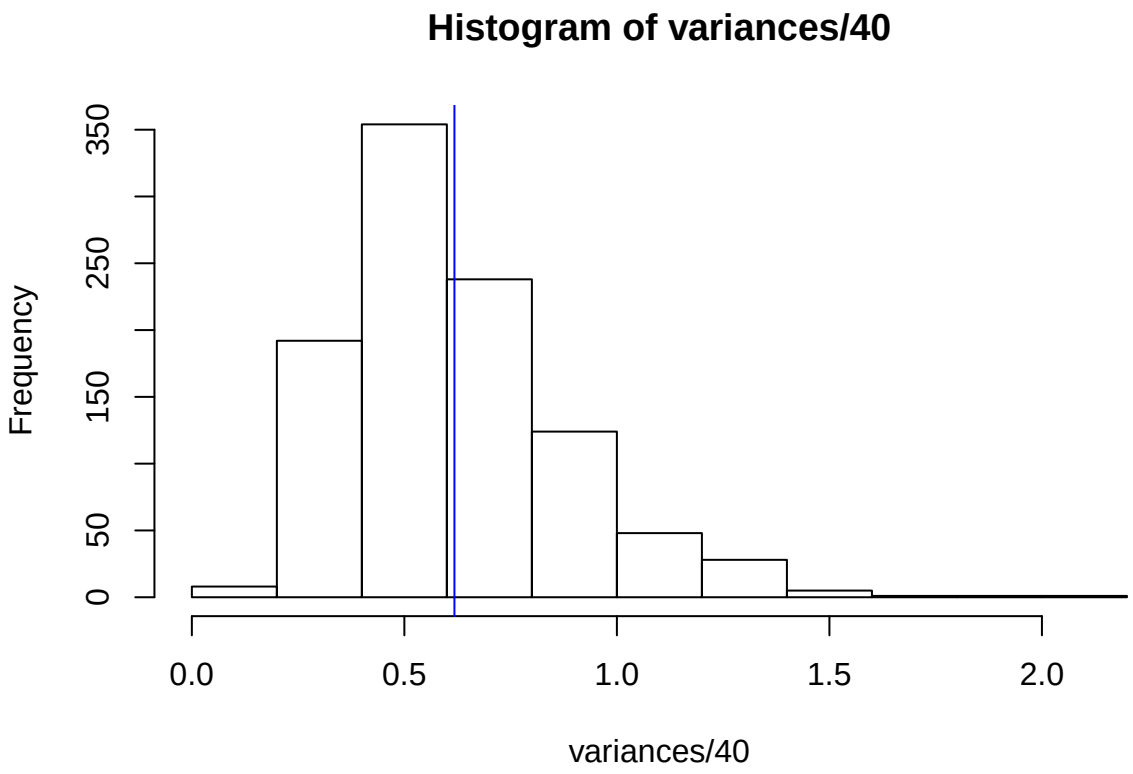
```
averages <- apply(data, 1, mean)
hist(averages)
abline(v = 5, col = "red")
abline(v = mean(averages), col = "blue")
text(3,230, labels = "mean_t: 5")
text(7, 230, labels = paste("mean_s:", mean(averages)))
```



Sample variance versus Theoretical variance:

Next we will compare the theorectical variance with the sample variance. Recall that theoretical standard deviation = 1/lambda. So, theoretical variance = 1/lambda^2 For 40 measurements the variance will be divided by n = 40. Thus, theoretical variance = (1/0.2)^2/40 = (5)^2 = 25/40=0.625.

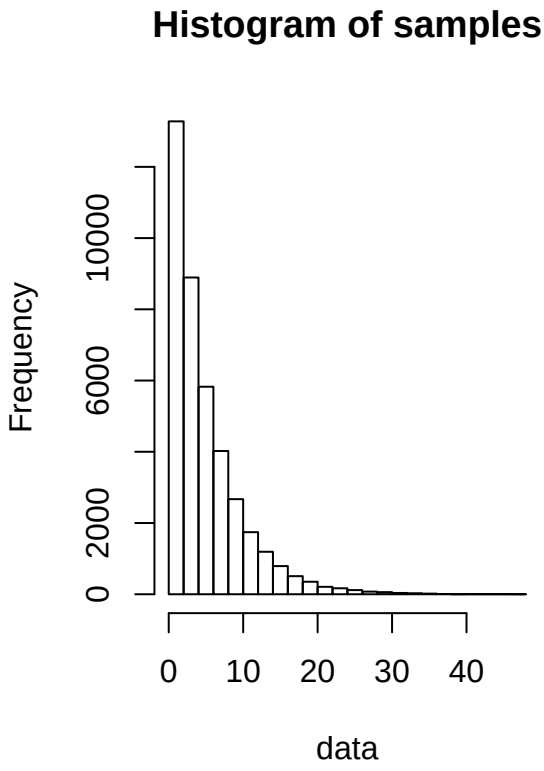
```
#calculating the variance of the averages:
variances <- apply(data, 1, var)
hist(variances/40)
abline(v = mean(variances/40), col = "blue")
abline(v = 6.25, col = "red")
```

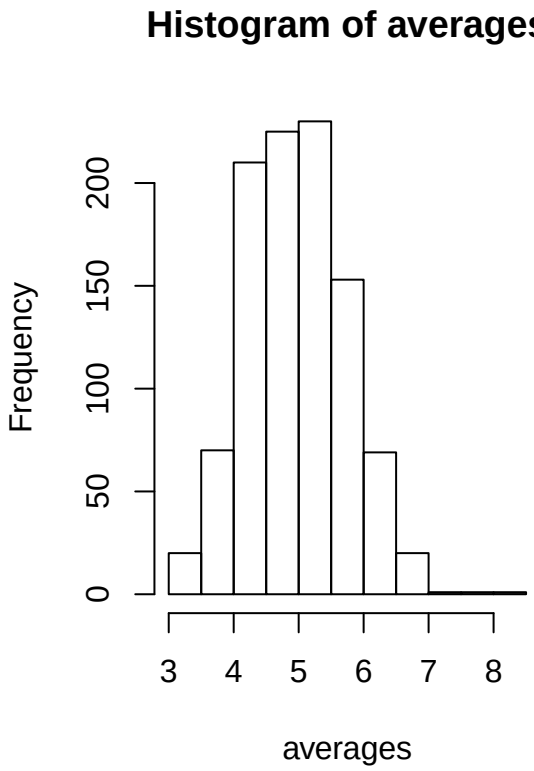


Distribution:

Lastly I will show the difference between the distribution of 40000 samples of the exponential distribution and the averages of 1000 collections of 40 samples. What this will demonstrate is that the averages are approaximately normally distibuted. This effect can be explained by the LLN.

```
par(mfrow = c(1,2))
hist(data, main = "Histogram of samples")
hist(averages)
```


Histogram of samples
Frequency
0
2000
6000
10000
0
10
20
30
40
data


Histogram of averages
Frequency
0
50
100
150
200
3
4
5
6
7
8
averages