Statistical Inference Course Assignment

Carolinecoder

December 21, 2015

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.

1. Show the sample mean and compare it to the theoretical mean of the distribution.

```
# Create 1000  simulations of 40 exponentials with lamba = 0.2
set.seed(111)
n <- 40
lambda <- 0.2
sims <- seq(1:1000)
Averages <- sapply(sims, function(s) {
    n <- 40
    lambda <- 0.2
    exp_dist <- rexp(n, lambda)
    mean(exp_dist)
})
# Calculate the expected population mean, μ
1/lambda
```

```
## [1] 5
```

```
# Calculate the mean of the 1000 sample means
mean(Averages)
```

```
## [1] 5.02562
```

```
# The theoretical population mean of 5 is very close to the sample mean of 5.03.
```

2. Show how variable the sample is and compare it to the theoretical variance of the distribution.

```
# Calculate the expected variance of the population, using the Central Limit Theor
em
(1/lambda/sqrt(n))^2
```

```
## [1] 0.625
```

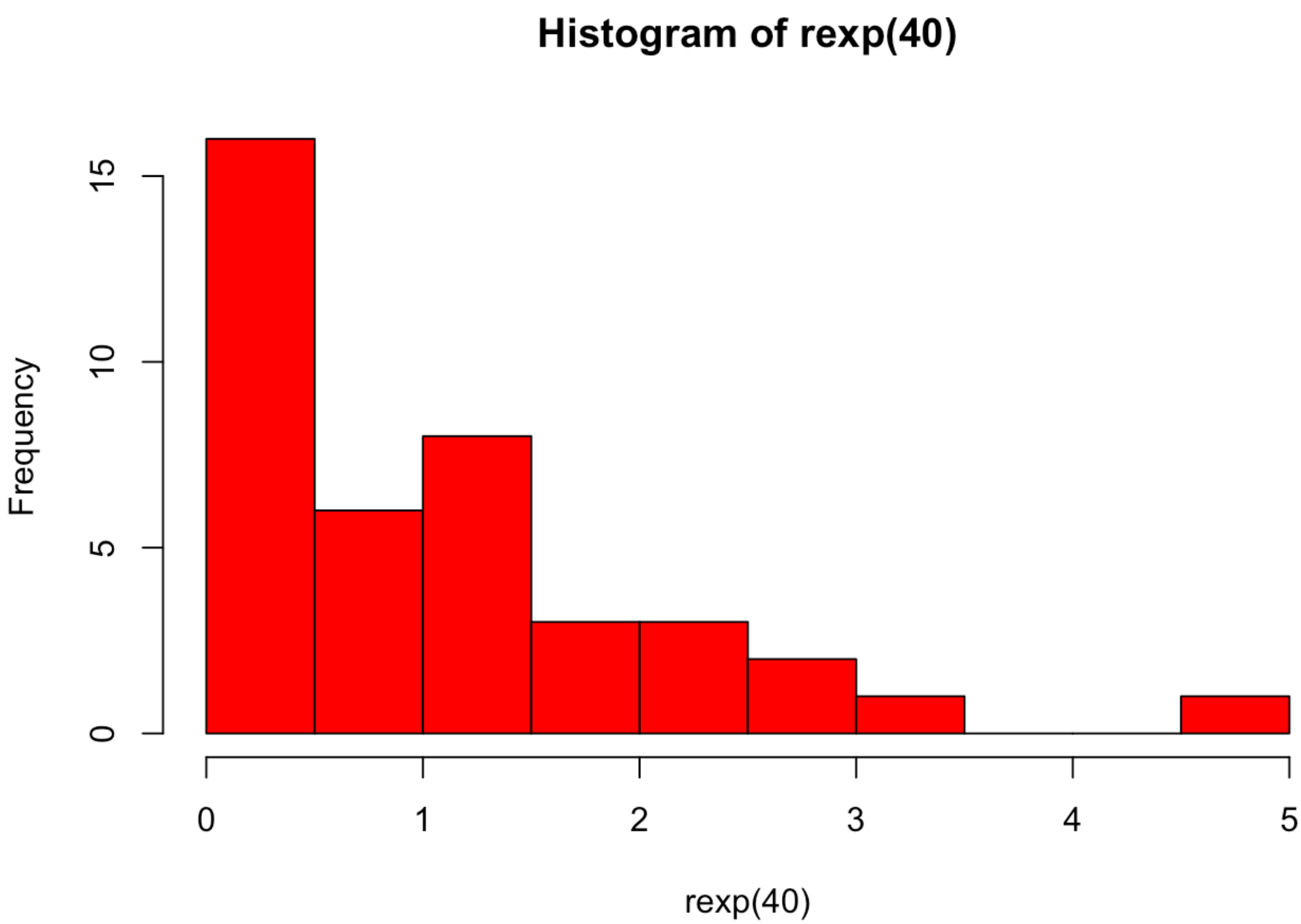
```
# Calculate the variance of the 1000 sample means
var(Averages)
```

```
## [1] 0.6069798
```

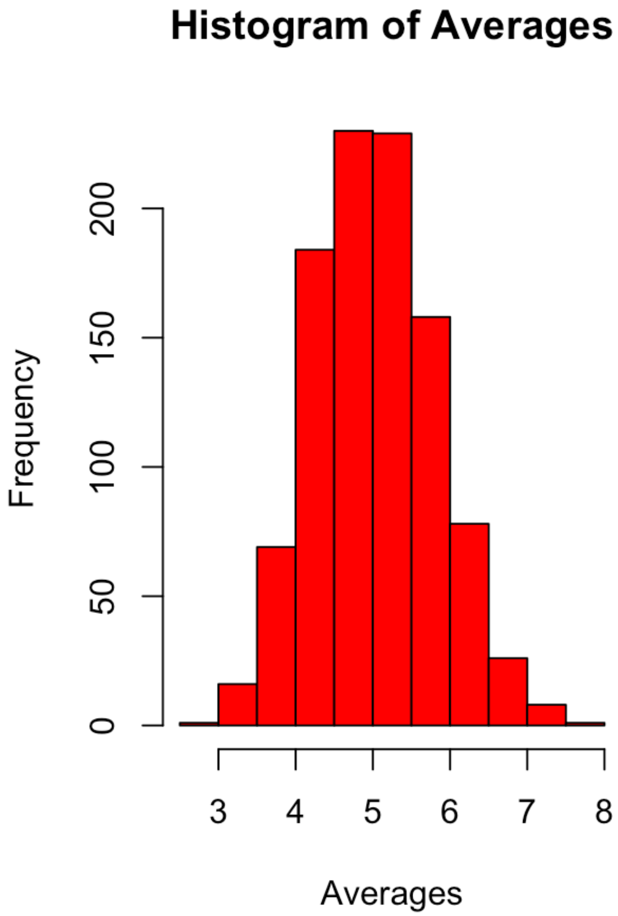
```
# Again, the theoretical population variance of 0.625 is quite close to the sample
variance of 0.607.
```

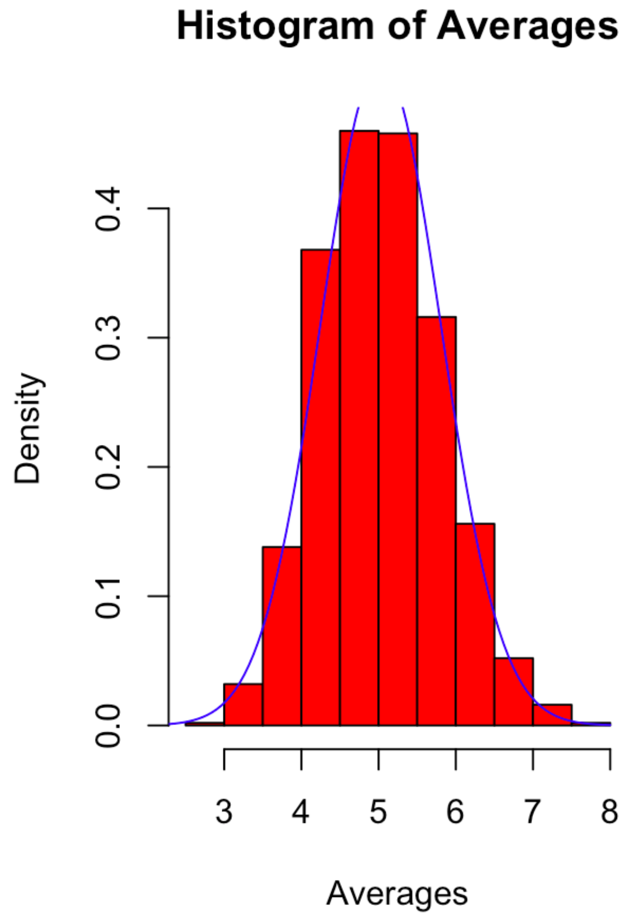
3. Show that the distribution is approximately normal.

```
# First, plot 40 random exponential variables
hist(rexp(40), col="red")
```




```
# The distribution of 40 exponentials is not normal
# Second, plot the averages of 1000 simulations of 40 exponentials
par(mfrow=c(1, 2))
hist(Averages, col="red")
hist(Averages, prob=TRUE, col="red")
curve(dnorm(x, mean=mean(Averages), sd=sd(Averages)), 2, 8, add=TRUE, col="blue")
```






```
# Here, we see that the sample averages are normally distributed
```