Statistical Inference Assignment_Part1

Edie Hawthorne

7/17/2018

Overview

This is an analysis to investigate the exponential districution in R and compare with with the Central Limit Theorem.

The first part is an analysis for a simulation of an exponential distribution.

The second part is a basic inferential data analysis about tooth growth.

Simulation

The first simulation (simulate_pop) below is simulating 1000 random samples as the population. The second simulation (mns) below is taking 40 random samples from the population and then take the mean of the 40 exponential random samples. The third simulation (vars) below is taking 40 random samples from the population and then take the variance of the 40 exponential random samples.

```
lambda = 0.2

simulate_pop <- rexp(1000, lambda)

mns = NULL
for (i in 1:1000) mns = c(mns, mean(rexp(40, lambda)))

vars = NULL
for (i in 1:1000) vars = c(vars, var(rexp(40, lambda)))
```

Compare Sample mean versus Theoretical Mean

By taking the average of the 40 simulated exponential random samples, we get 5.04, which compared to the theoretical mean of 1/lambda which is 5. We can see the sample mean equals the theoritical mean.

```
mean_sample <- mean(mns)

mean_theore <- 1/lambda
```

Sample Mean steps: 5.04 Theoretical Mean steps: 5.00

Compare Sample Variance versus Theoretical Variance

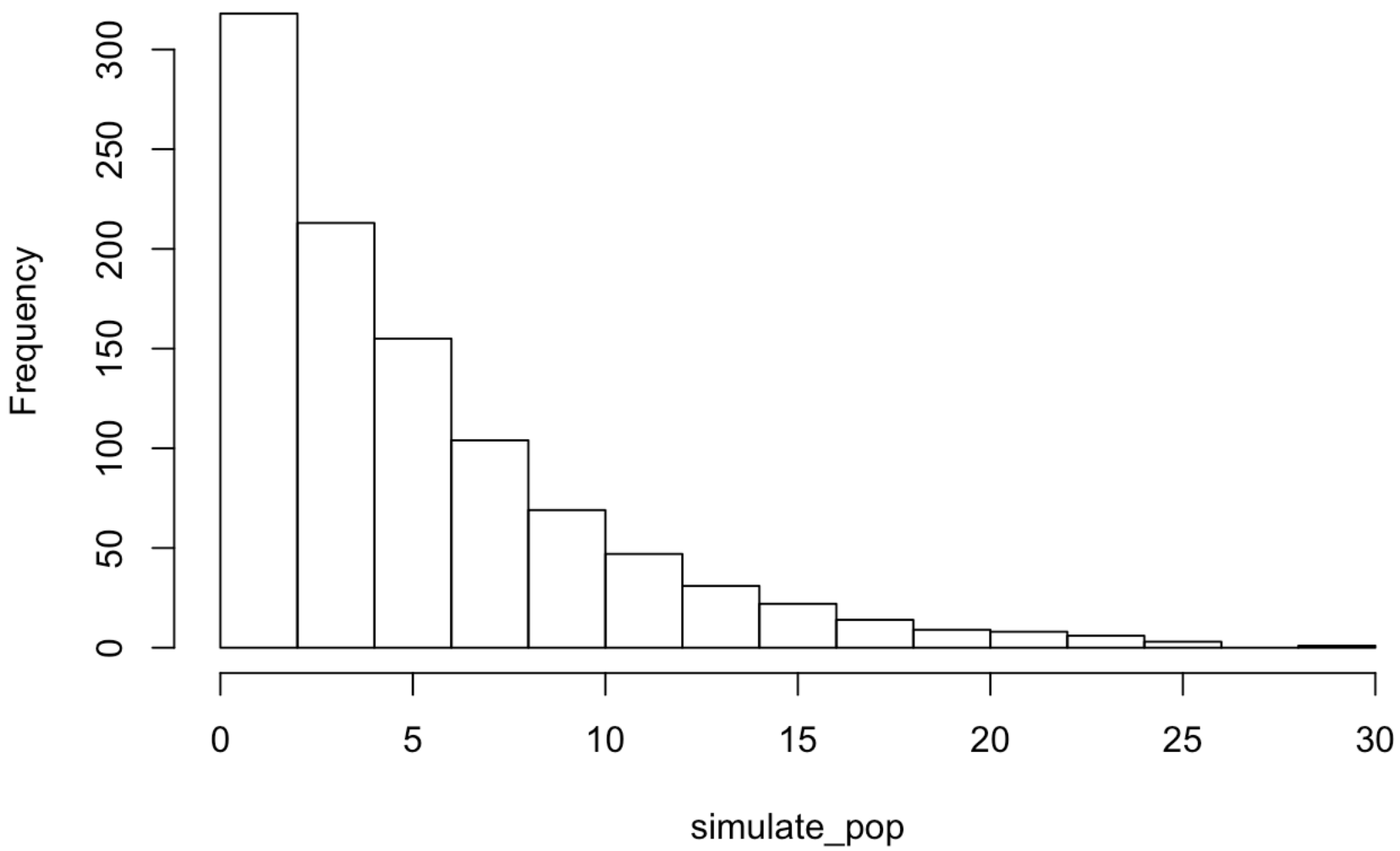
By taking the average of the variance of the 40 simulated exponential random samples, we get 25.16, which compared to the theoretical variance of 1/lambda^2 which is 25. We can see the sample variance is very close to the theoretical variance.

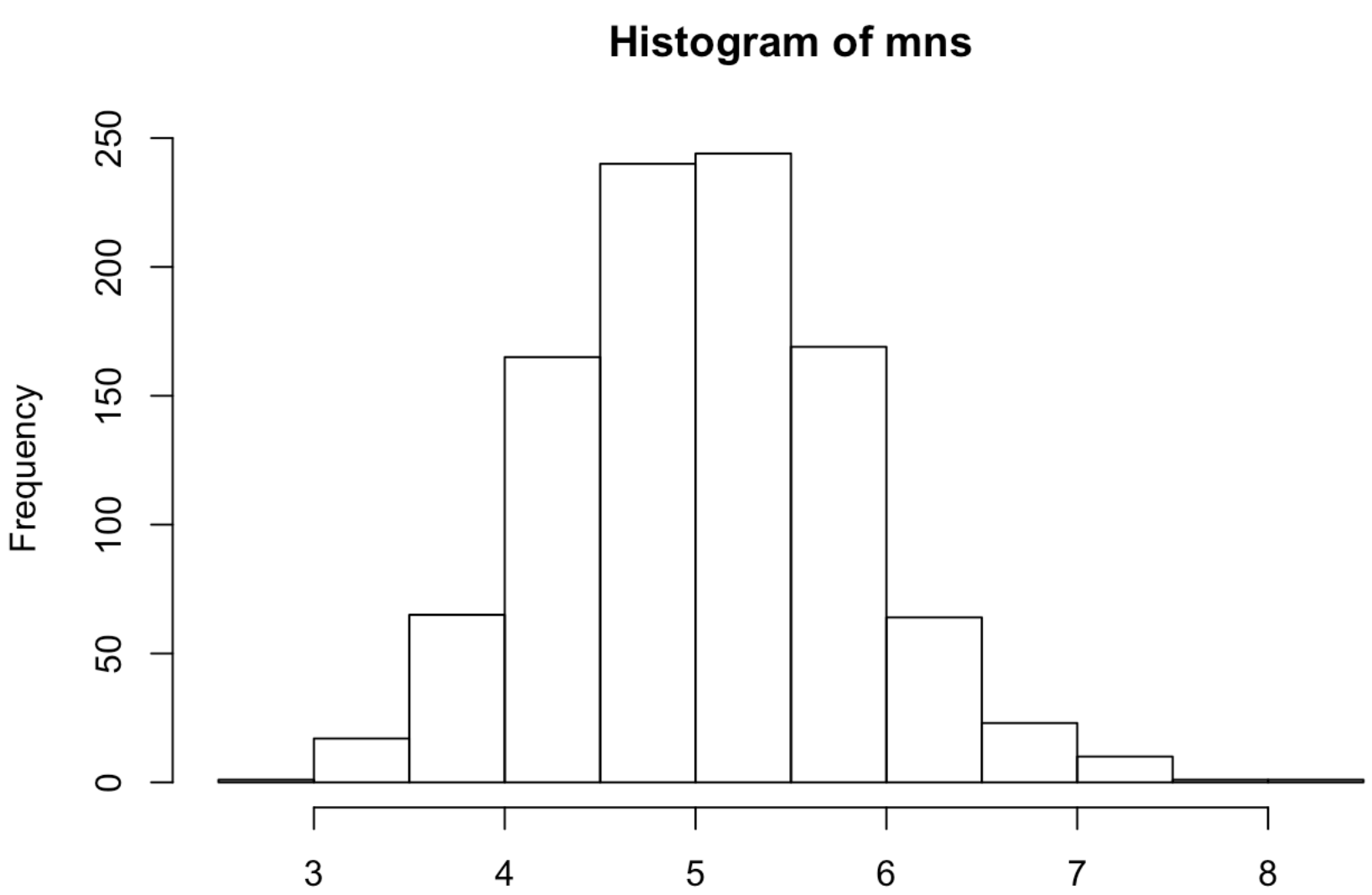
```
var_sample <- mean(vars)

var_theore <- (1/lambda^2)
```

Sample variance steps: 24.79 Theoritical variance steps: 25.00

Distribution

As you can see, after taking the avrage of mean of the 40 exponential random samples, the distribution (dist_samplemean) become more like standard normal distribution. Because the average of the sample mean are centralized around the population mean.


Histogram of simulate_pop
Frequency
0
50
100
150
200
250
300
0
5
10
15
20
25
30
simulate_pop


Histogram of mns
Frequency
0
50
100
150
200
250
3
4
5
6
7
8

mns