Statistical Inference Course Project

Overview

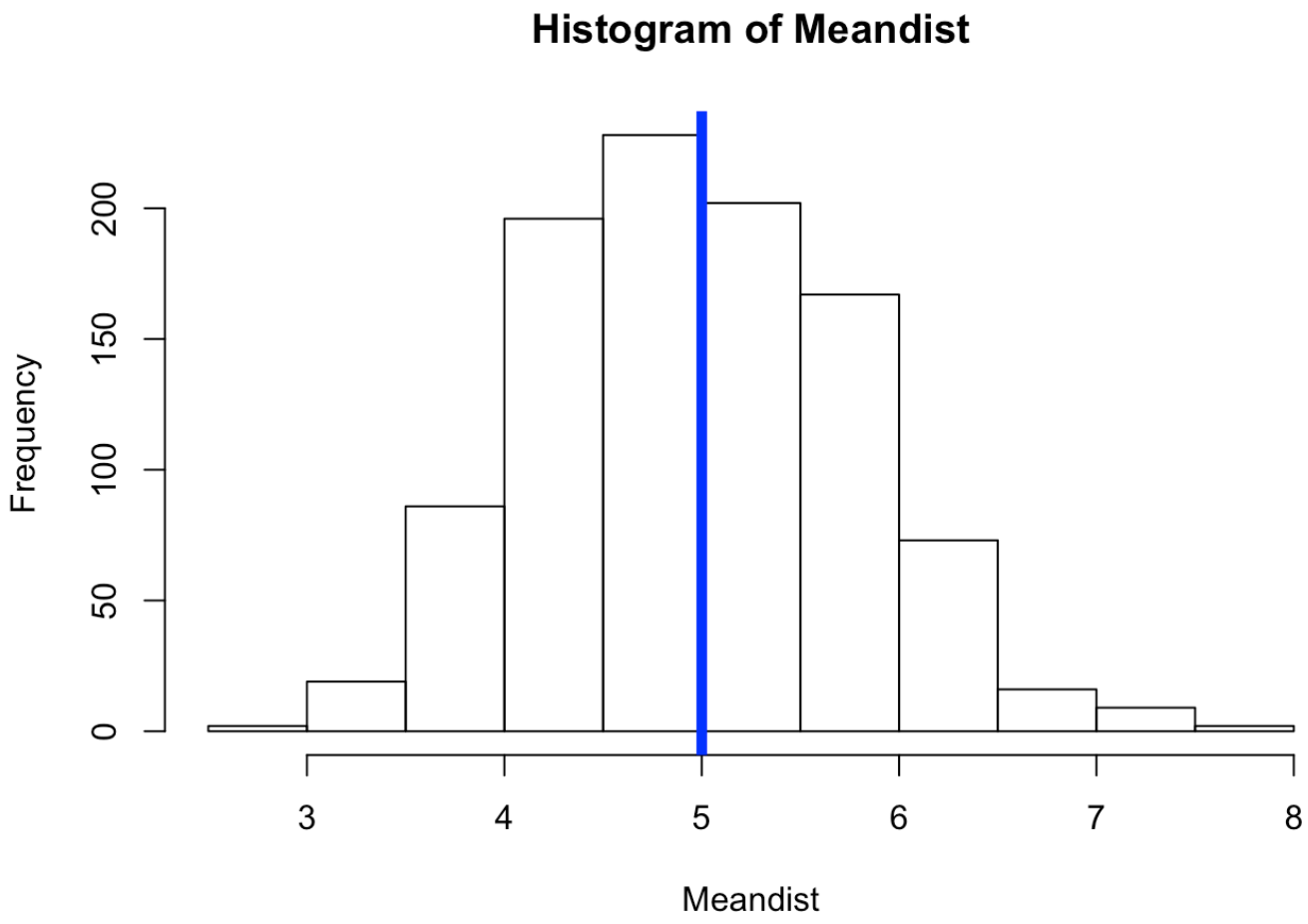
This project is spilt into 2 parts. They will be spilt into simulation and statistical analysis of tooth growth data. For the simulation part, we will investigate the exponential distribution in R and compare it with the Central Limit Theorem.

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

Part 1: Simulation

A. Mean

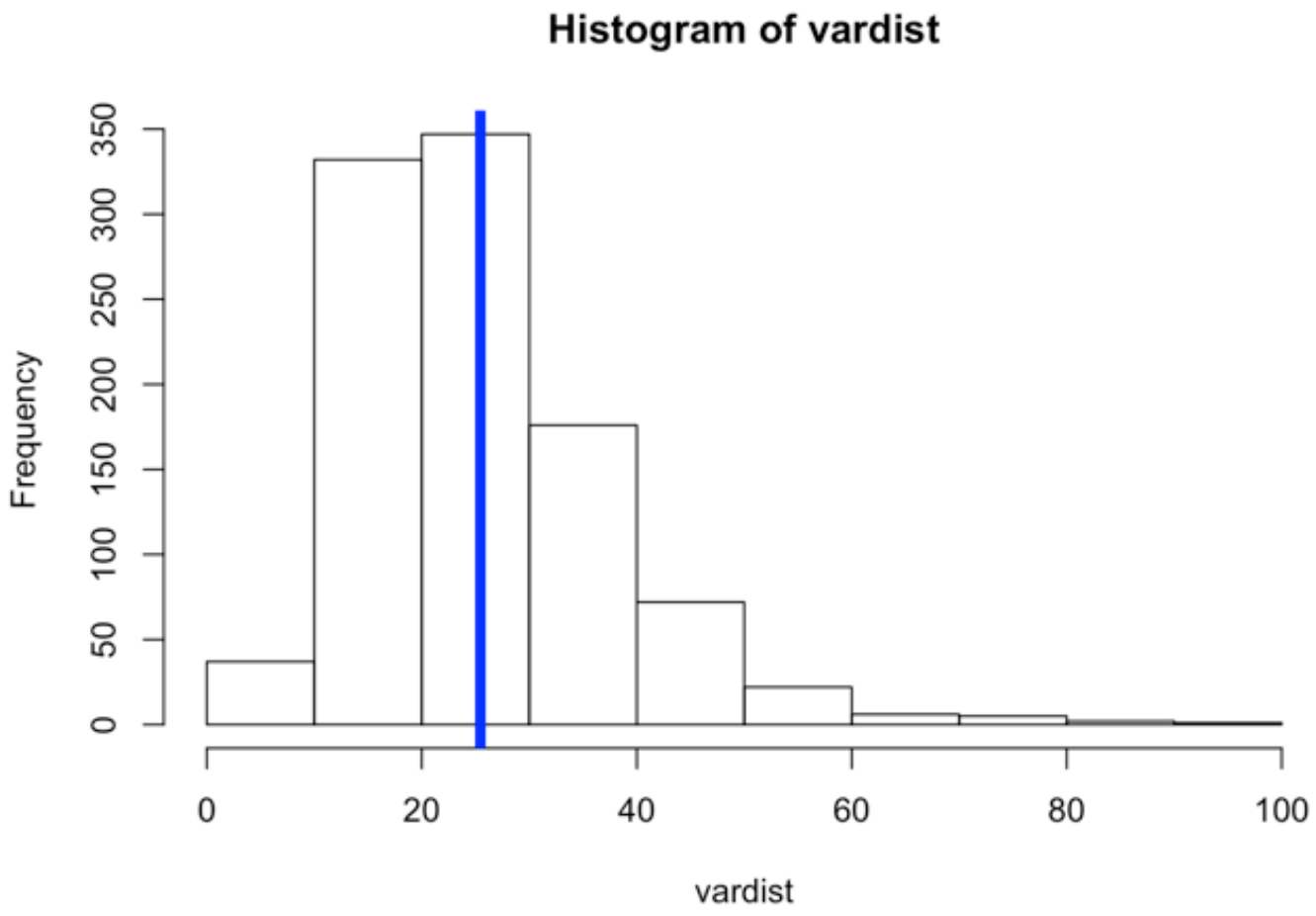
```
lambda = 0.2
mean_exp = 1/lambda
sd_exp = 1/lambda
sim_mat <- matrix(rexp(1000*40, lambda), nrow = 1000, ncol = 40)
Meandist <- apply(sim_mat, 1, mean)
hist(Meandist)
abline(v = 1/lambda, lty = 1, lwd = 5, col = "blue")
```



Simulated mean is the barchart - in this case is 4.9671627 Theoretical mean is the blue line 5

B. Variance

```
vardist <- apply(sim_mat, 1, var)
hist(vardist)
abline(v = mean(vardist), lty = 1, lwd = 5, col = "blue")
```




```
#theoretical variance
(1/lambda)^2
```

```
## [1] 25
```

```
#sample variance
mean(vardist)
```

```
## [1] 25.45396
```

Theoretical variance is the black line - in this case is 5, because the theoretical variance is 25 Sample variance is 25.4539556 Thus, they are quite close.

From the density and histogram, we can see that it is approx normal, since they concentrate near the mean, and is approximately symmetrical about the mean.