Statistical Inference Project Part 1

Jack Chesson

Overview

Through the following processes, we will investigate the exponential distribution in R and compare the exponential distribution with lambda parameter 0.2 to the Normal Distribution using the Central Limit Theorem. This will be done by simulating the means of 40 exponentials and investigating its distribution.

Libraries

```
library(data.table)
library(ggplot2)
library(knitr)
```

Set Constants

```
set.seed(31)

lambda <- 0.2

n <- 40

sims <- 1000

exponential_sims <- replicate(sims, rexp(n, lambda))

exponential_means <- apply(exponential_sims, 2, mean)
```

Question 1

Compare the simulated mean to the theoretical mean of the distribution

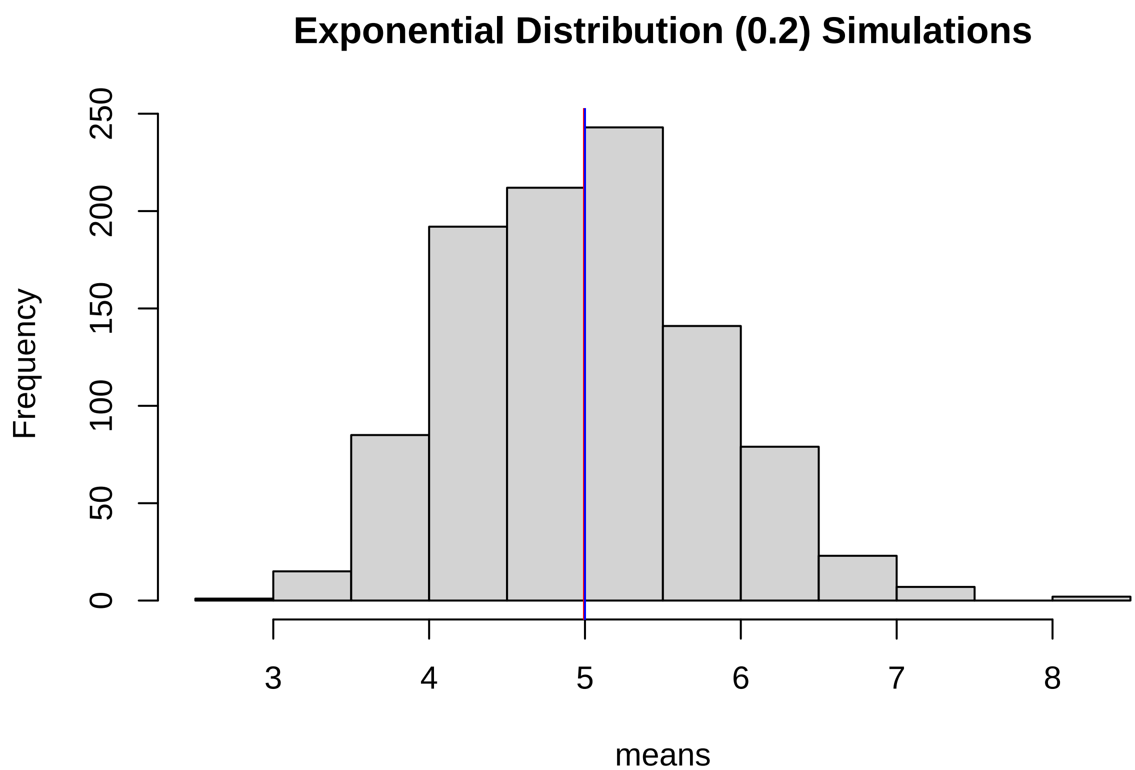
```
sims_mean <- mean(exponential_means)
sims_mean
```

```
## [1] 4.993867
```

```
theoretical_mean <- 1/lambda
theoretical_mean
```

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

```
hist(exponential_means, xlab = "means", main = "Exponential Distribution (0.2) Simulations")
abline(v = sims_mean, col = "red")
abline(v = theoretical_mean, col = "blue")
```



The simulated mean is 4.993867 and the theoretical mean is 5, showing that the simulated mean is almost exactly the theoretical mean.

Question 2

Next, the variance of the simulated distribution will be explored in order to compare to the theoretical variance of this exponential distribution.

```
stddev_sims <- sd(exponential_means)
stddev_sims
```

```
## [1] 0.7931608
```

```
stddev_theory <- (1/lambda)/sqrt(n)
stddev_theory
```

```
## [1] 0.7905694
```

```
var_sims <- stddev_sims^2
var_sims
```

```
## [1] 0.6291041
```

```
var_theory <- stddev_theory^2
var_theory
```

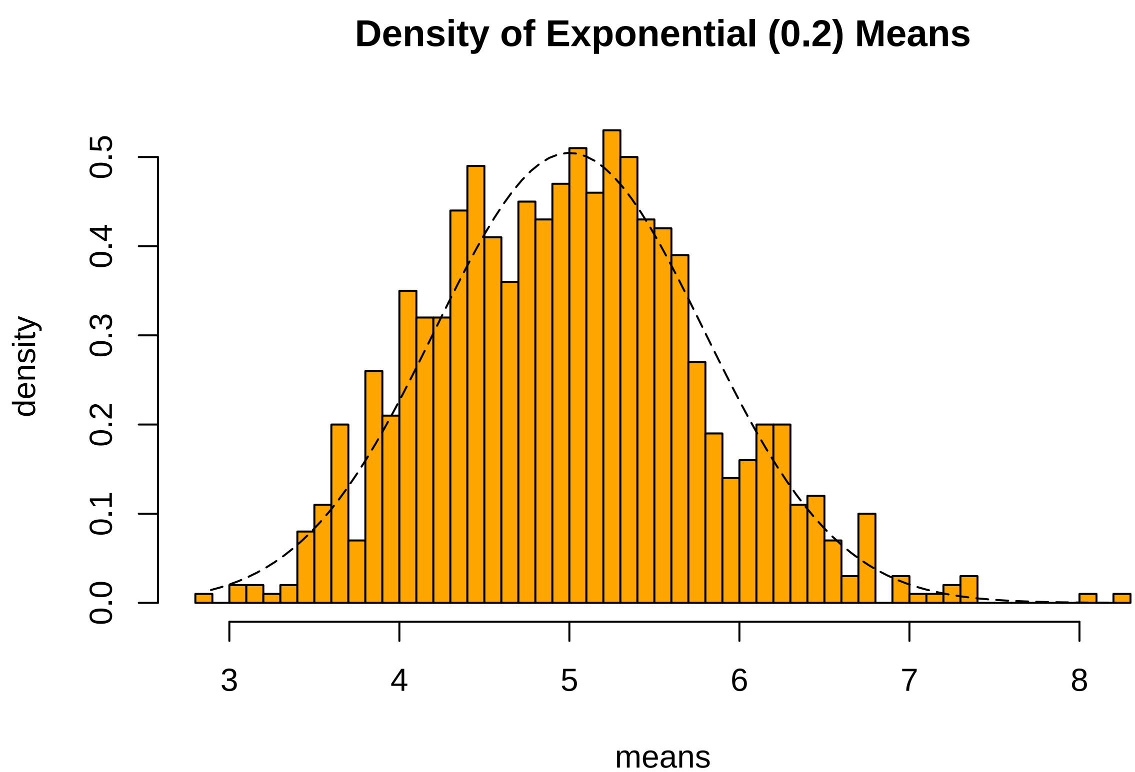
```
## [1] 0.625
```

The Simulated Standard Deviation is 0.7931608 with a variance of 0.629104, which is very close to the theoretical values of 0.790569 and 0.625 respectively.

Question 3

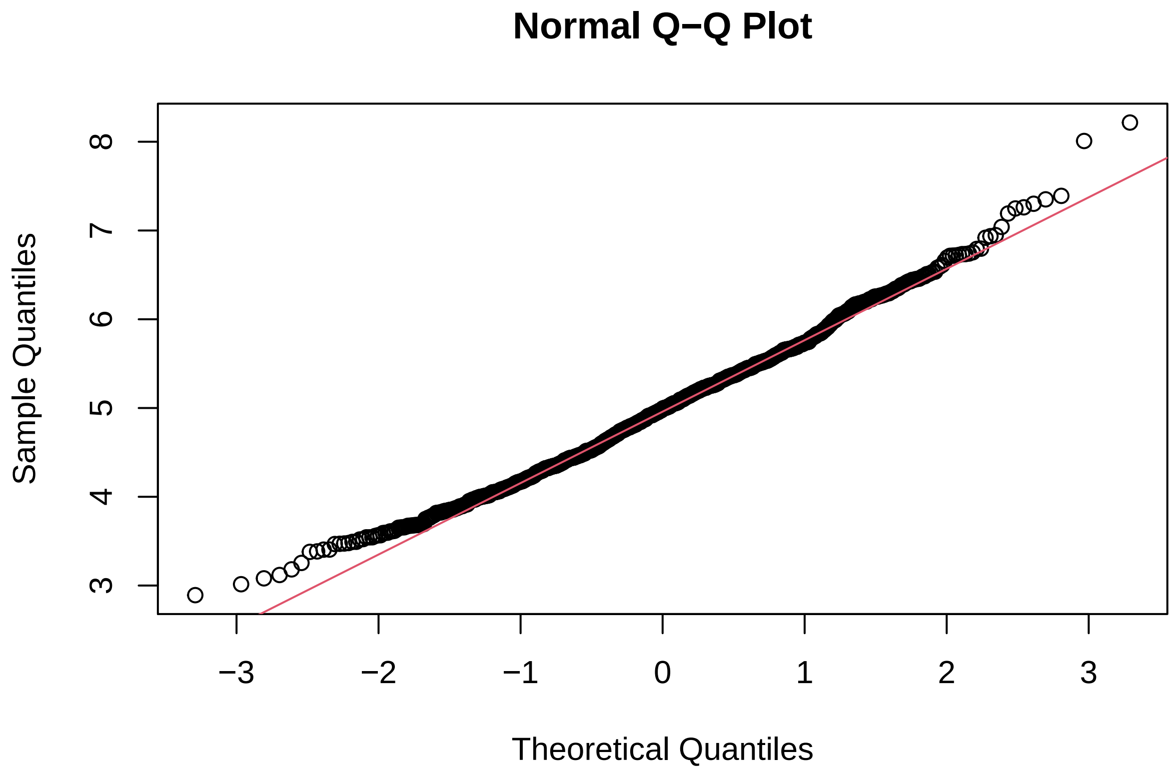
Using the simulated distribution, we will use the Central Limit Theorem to address the comparison to a normal distribution.

```
sim_fit <- seq(min(exponential_means), max( exponential_means), length = 100)

theory_fit <- dnorm(sim_fit, mean = 1/lambda, sd = (1/lambda)/sqrt(n))

hist(exponential_means, breaks = n, prob = T, col = "orange", xlab = "means",ylab = "density", main = "I

lines(sim_fit, theory_fit, pch = 22, col = "black", lty = 5)
```



Using quantile plots, we will show the comparison of the averages of 40 exponentials to a normal distribution.

```
qqnorm(exponential_means)
qqline(exponential_means, col = 2)
```



It can be seen in the Quantile-Quantile Plot that the Simulated Distribution of 40 Exponential (0.2) Distributions is very cose to a Normal Distribution