Statistical Inference Course Project

Fan Ouyang

Overview

This project investigates the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution is 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.

Part 1:Simulation Exercise Instructions

Q1 Show the sample mean and compare it to the theoretical mean of the distribution.

Results show the simulated mean is 5.01, compared to the calculated theoretical mean of 5.

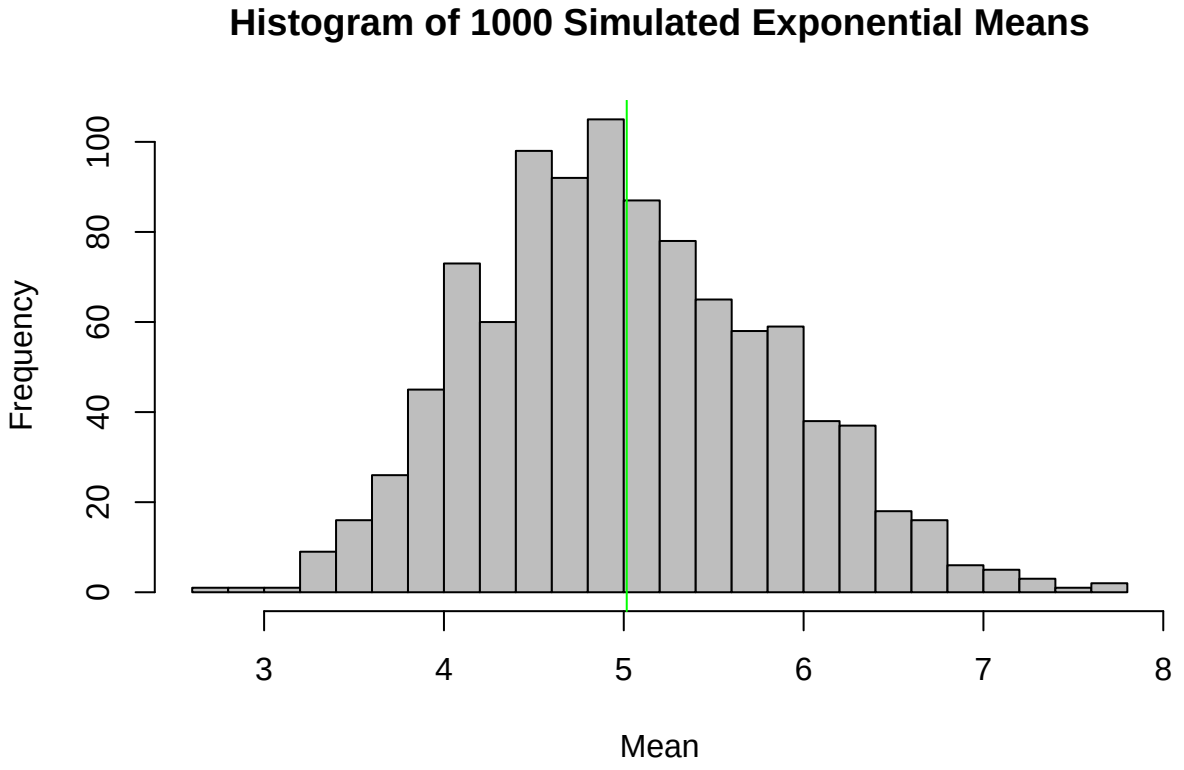
```
set.seed(2)
lambda <- 0.2
sim <- 1000
n <- 40

# Simulate and take mean
sim.exp <- replicate(sim, rexp(n,lambda))
mean.exp <- colMeans(sim.exp)

# Simulated mean
s.mean <- mean(mean.exp)
print(s.mean)
```

```
## [1] 5.016356
```

```
# Theoretical mean
t.mean <- 1/lambda
print(t.mean)
```

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

```
# plot
hist(mean.exp, breaks = 20, xlab = "Mean", main = "Histogram of 1000 Simulated Exponential Means", col =
abline(v = s.mean, col = "green")
```



Q2 Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

Results show the simulated variance is 0.6691, compared to the theoretical variance of 0.625.

```
# Simulated Standard Deviation and Variance
sstdev <- sd(mean.exp)
svar <- sstdev^2
print(svar)
```

```
## [1] 0.6691305
```

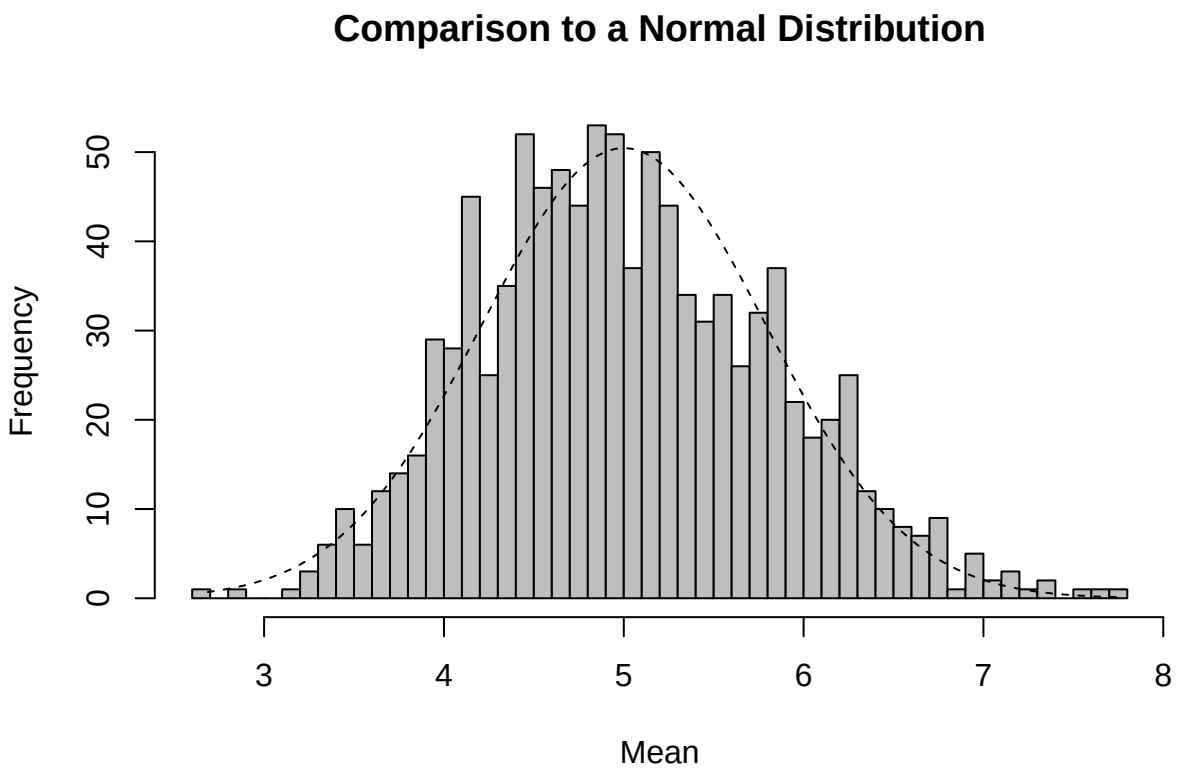
```
# Theoretical Standard Deviation and Variance
tstdev <- (1/lambda)/sqrt(n)
tvar <- tstdev^2
print(tvar)
```

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

Q3 Show that the distribution is approximately normal.

As the plot shows, the distribution is approximately matched with a normal distribution.

```
# Plot Histogram
hist(mean.exp, breaks = 40, xlab = "Mean", main = "Comparison to a Normal Distribution", col = "grey")

# Add the Theoretical Normal Distribution Line
xfit <- seq(min(mean.exp), max(mean.exp), length = 100)
yfit <- dnorm(xfit, mean = 1/lambda, sd = 1/lambda/sqrt(n))
lines(xfit, yfit*100, lty=2)
```


Comparison to a Normal Distribution
Frequency
0
10
20
30
40
50
3
4
5
6
7
8
Mean