Course6Week4Assignment-YR-Part1

Yonatan Rafael

==========

Overview

Part 1 of this report illustrates a simulation of the exponential distribution and it's proximity to the theoretical distribution, along with proof that the distribution is approximately normal.

Sample versus Theoretical Mean

Run 1000 simulations of an exponential distribution. Each simulation contains 40 observations and a lamba = .2. The output returns the sample mean of our simulation.

```
n <- 40
lambda <- .2
nosim <- 1000
df <- data.frame(ncol = 2, nrow = nosim)
colnames(df) <- c("index", "simulation")
set.seed(1000)
for (i in 1:nosim) {
        df[i, 1] <- i
        df[i, 2] <- mean(rexp(n, lambda))
}
mean(df$simulation)
```

```
## [1] 4.986963
```

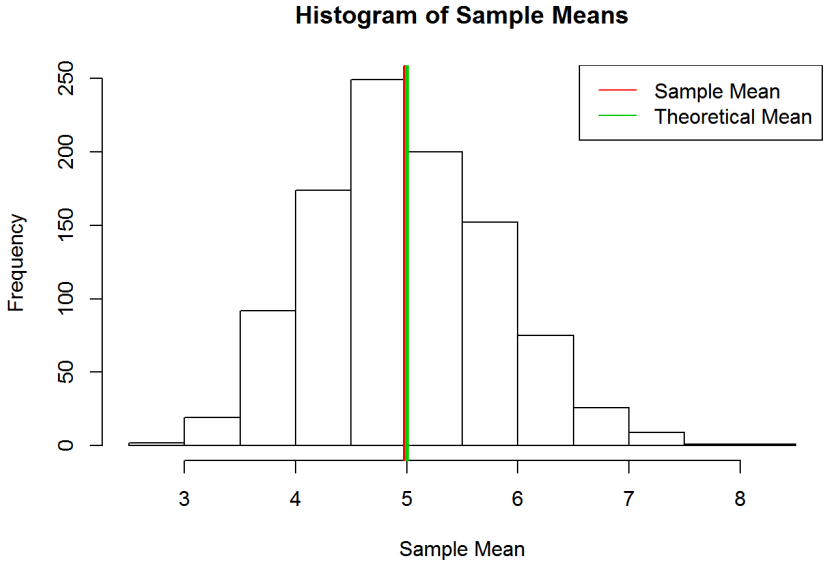
The theoretical mean is calculated below.

```
1/ lambda
```

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

Below is a visual representation of the simulation distribution, with lines showing the proximity between the sample mean (4.99) and theoretical mean (5).

```
hist(df$simulation, xlab = "Sample Mean", main = "Histogram of Sample Means")
abline(v = mean(df$simulation), col = 2, lwd = 3)
abline(v = 1 / lambda, col = 3, lwd = 3)
legend('topright', c("Sample Mean", "Theoretical Mean"),  col = c(col = 2, col = 3), lty = c(1,1))
```



Sample versus Theoretical Variance

Using the simulation presented above, below is the sample variance.

```
var(df$simulation)
```

```
## [1] 0.654343
```

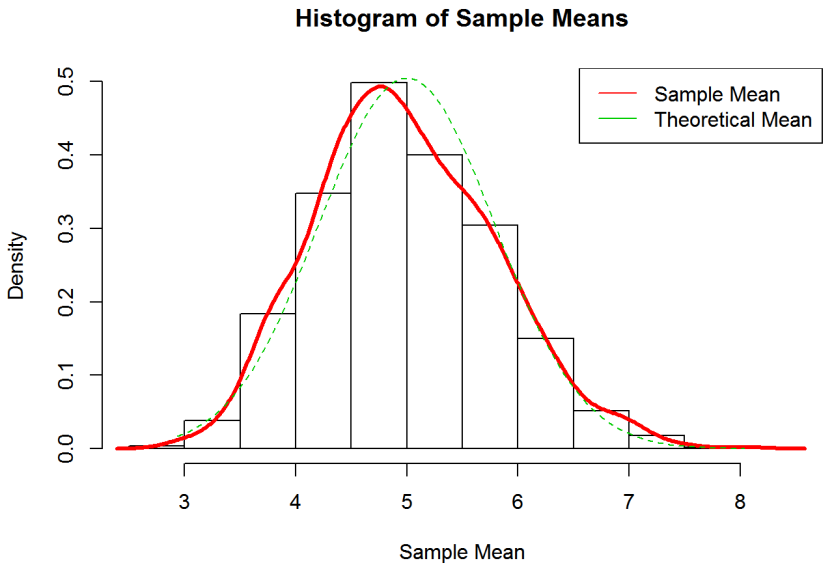
The theoretical variance is calculated below.

```
(1/ (lambda ** 2))/n
```

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

Below is another visual representation of the simulation distribution, with lines showing the proximity between the sampel variance (.654) and theoretical variance (.625).

```
hist(df$simulation, xlab = "Sample Mean", main = "Histogram of Sample Means", prob = TRUE)
lines(density(df$simulation),col = 2,lwd = 3)
xfit <- seq(min(df$simulation), max(df$simulation), length = 100)
yfit <- dnorm(xfit, mean = 1/lambda, sd = (1/lambda/sqrt(n)))
lines(xfit, yfit, col = 3, lty = 2)
legend('topright', c("Sample Mean", "Theoretical Mean"),  col = c(col = 2, col = 3), lty = c(1,1))
```



Proving Normal Distribution

The chart above starts to show that the distribution resembles a normal distribution. That said, let's use a QQ-plot to prove normality.

```
qqnorm(df$simulation, col = "red")
qqline(df$simulation)
```