Coursera - Statistical Inference

Teddy Rusli

Friday, May 21, 2015

This project is meant to show how exponential distribution compared with the Central Limit Theorem. The exponential distribution is simulated with rexp(n,lambda) where lambda is the rate parameter. Lambda = 0.2 is set for all the simulations

Step 1: Show the sample mean and compare it to the theoretical mean of the distribution.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
lambda <- 0.2
n <- 40
simNumber <- 1:1000
set.seed(888)

sampleMean <- data.frame(
                x = sapply(simNumber,
                    function(x) {
                        mean(rexp(n,rate = lambda))
                    }
                    )
                )
```

```
simMean <- mean(sampleMean$x)
trueMean <- 1/lambda

simSD <- sd(sampleMean$x)
```

The simulation resulted in a sample mean of 4.984112 and theoretical mean is 5. They are very close to each other, less than one standard deviation (0.7992551)

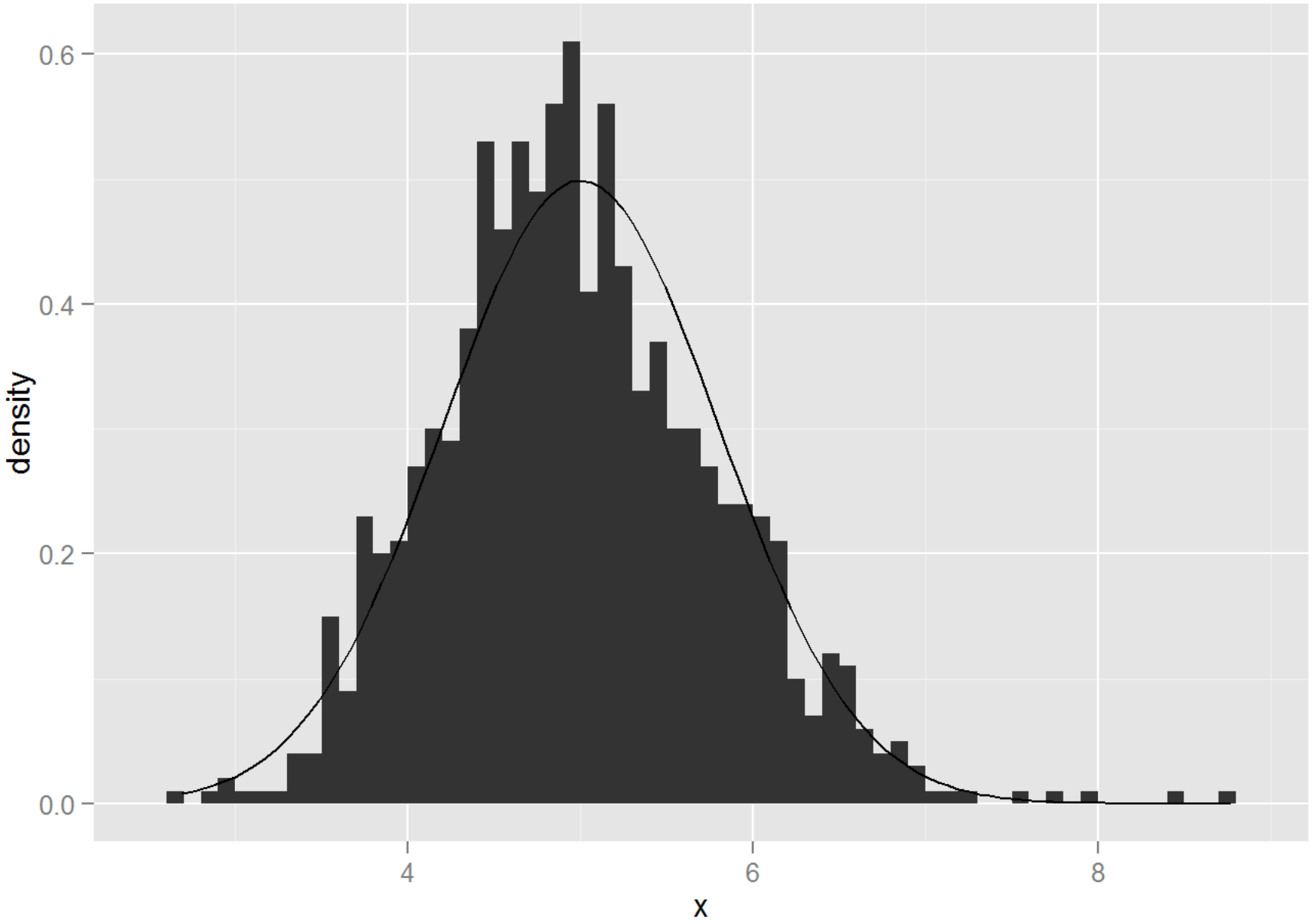
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

```
simVar <- var(sampleMean$x)
trueVar <- ((1/lambda)/sqrt(n))^2
```

The simulation resulted in a sample mean of 0.6388088 and theoratical mean is 0.625. Again, very close.

3. Show that the distribution is approximately normal.

```
ggplot(data = sampleMean, aes(x = x)) +
    geom_histogram(aes(y=..density..),binwidth = 0.10) +
    stat_function(fun = dnorm, arg = list(mean = trueMean, sd = sd(sampleMean$x)))
```



The plot compares the distribution of the means from the simulation against a normal distribution. We can conclude that the Central Limit Theorem applies to exponential distribution