Statistical Inference Course Project1

doyougnu

July 20, 2015

Overview

Overview to be writting here...

Simulations

```
#set seed
set.seed = 1234

#number of simulations to run
simnum <- 1000

#number of distributions to generate
exp_n <- 40

#exponential distribution parameters
exp_lambda <- 0.2
simMatrix <- matrix(rexp(exp_n * simnum, rate = exp_lambda), simnum, exp_n)
expMean <- rowMeans(simMatrix)
```

The simulation is done by....

Plots Comparing Simulation to Population

```
#load ggplot2
library(ggplot2)

#Compare Means
popMean <- 1 / exp_lambda
simMean <- mean(expMean)
popMean
```

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

```
simMean
```

```
## [1] 5.028078
```

The means are very close, the pop mean is ..., the sim mean is...

```r
#Compare Variances
popVariance <- (1 / exp_lambda) ^ 2 / exp_n
simVariance <- var(expMean)
popVariance
```

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

```r
simVariance
```

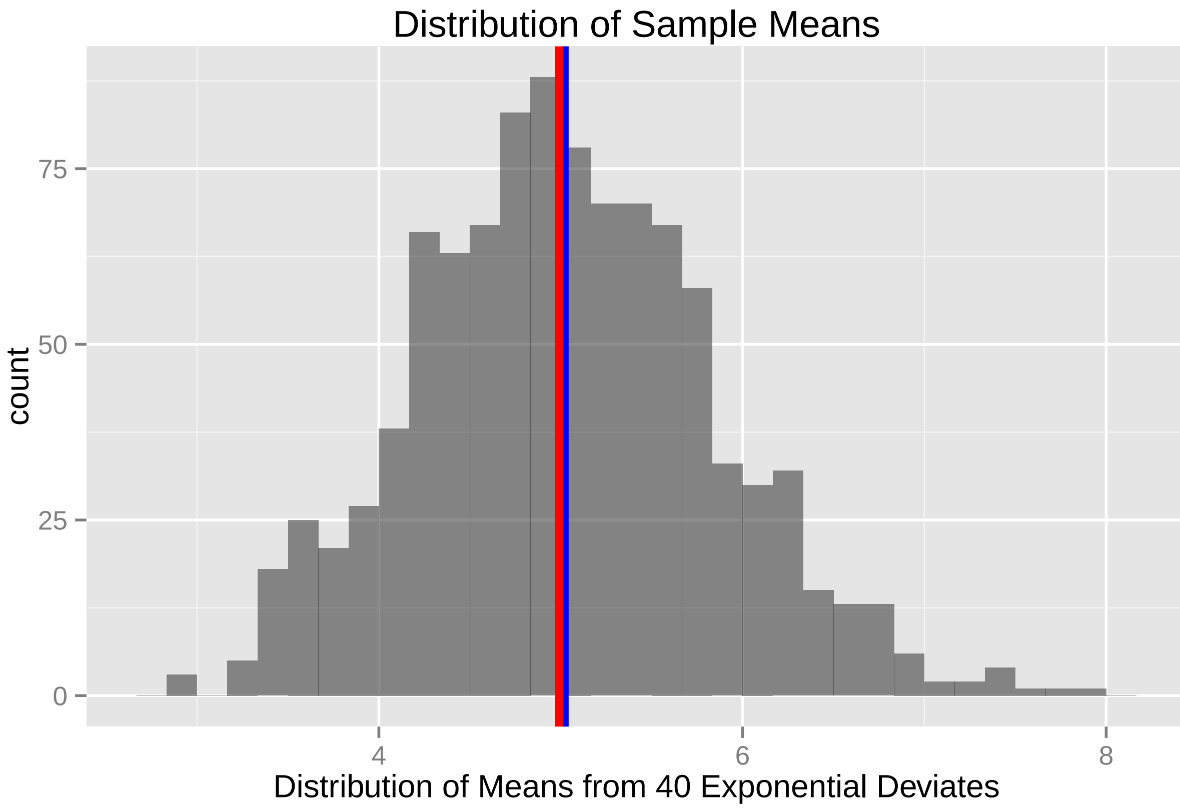
```
## [1] 0.6652346
```

Similarly for the Variances, the pop variance is . . . , the sim variance is . . .

Comparison between Random Variables and Sample Distributions

```r
#load libs
library(ggplot2)

#compare Variances visually
plt <- qplot(expMean, main = "Distribution of Sample Means", alpha = 0.3) +
    xlab("Distribution of Means from 40 Exponential Deviates") +
    geom_vline(size = 2, xintercept = popMean, colour = "red") +
    geom_vline(size = 1, xintercept = simMean, colour = "blue") +
    theme(legend.position = "none")
plt
```



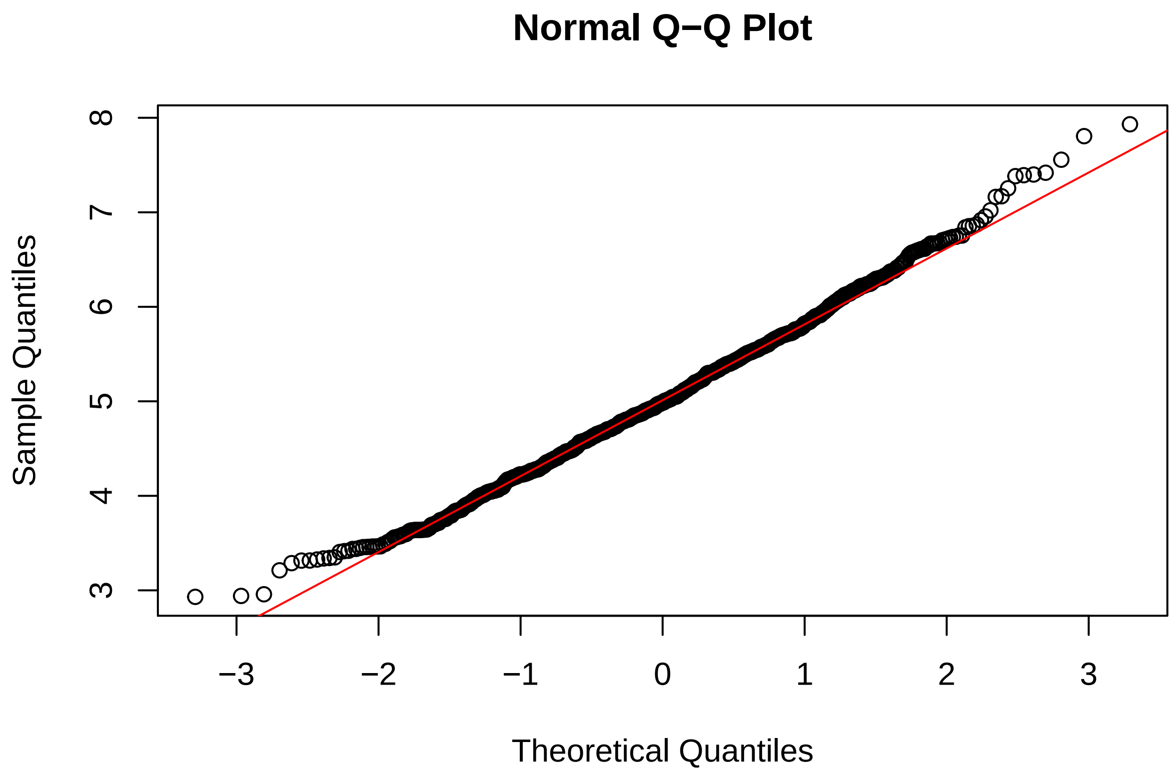
In this plot we can see the CLT in action . . .

Normality test

```r
#perform shapiro normality test
shapiro.test(expMean)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  expMean
## W = 0.9952, p-value = 0.003065
```

```r
#plot QQ plot
qqnorm(expMean); qqline(expMean, col = 2)
```


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
3
4
5
6
7
8
−3
−2
−1
0
1
2
3