Statistical Inference - Simulation Project

Joe DeMaro

December 2018

Set Parameters

```
setwd("c:\\rprograms\\statinference")
n <- 1000
group <- 40
lambda <- 0.2
```

Calculate Theoretical Mean and SD

```
theoreticalmean <- 1 / lambda
theoreticalsd <- (1 / (lambda * lambda)) / group
```

Run Simulation

```
# run the simulation
data  <- rexp(n * group, lambda )
matrixdata <- matrix(data, n, group)
matrixmean <- apply(matrixdata,1,mean)
```

Tabulate Simulation Data

```
simmean <- mean(matrixmean)
simsd <- sd(matrixmean)
simvar <- var(matrixmean)

#compare to CLT
simse = simsd / sqrt(group)
#calculate confidence
low <- simmean - 1.96 * simse
high <- simmean + 1.96 * simse

library(knitr)

theorecticalData <- c(theoreticalmean,theoreticalsd)

simulationdata <- c(simmean, simsd)

df = data.frame(theorecticalData,simulationdata)
colnames(df) <- c("Theorectical", "Simulated")
row.names(df) <- c("Mean", "SD")

kable(df, format="markdown", caption="Comparison between Theorecital and
Simulated Data")
```

	Theorectical	Simulated
Mean	5.000	5.0138069
SD	0.625	0.7903444

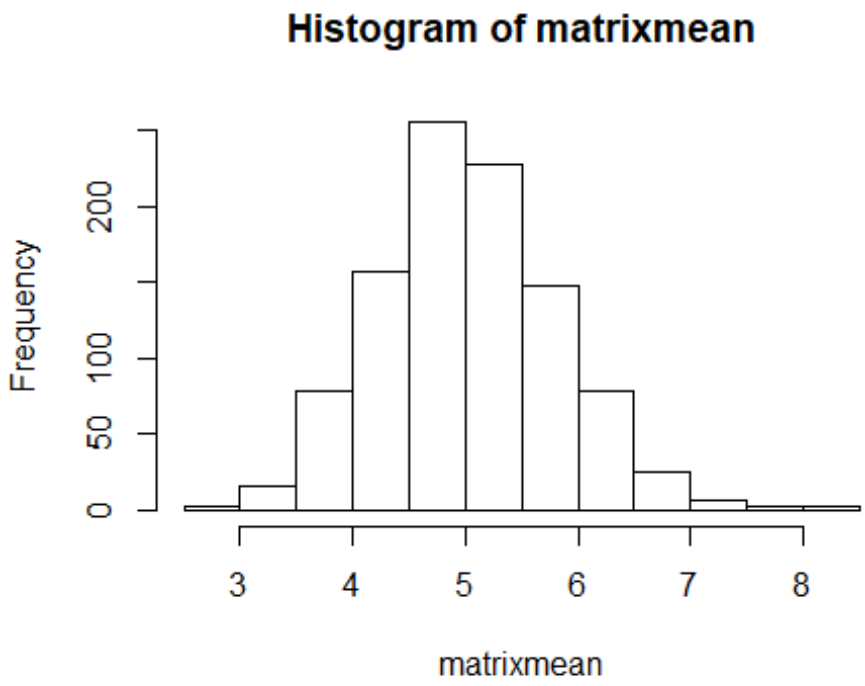
Distribution of Simulated Data around the mean (95% CI)

```
ciData <- c(low, high)
ciDF <- data.frame(ciData)
row.names(ciDF) <- c("Low", "High")
kable(ciDF, format="markdown", caption="CI for Simulated Data",
title="Simulated Data CI")
```

	ciData
Low	4.768877
High	5.258737

Histogram of Means from Simulation

```
hist(matrixmean)
```



Analysis

The null hypothesis Ho is the mean = 5; making the alternate hypothesis Ha the mean != 5. With a 95% confidence level, the simulated data is generated to test. The histogram shows the data are normally distrubtedd and the data show the simulated mean falls with in the 95% level. Therefore you can not reject the Ho.