Assignment: Statistical Inference Course Project 2 Part 1

John Allen

Thursday, January 28, 2016

Load from library

```
library(ggplot2)
```

Set parameters

```
set.seed(109759)
lambda <- .2
n <- 40
samples <- 1000
```

Show work and result

```
echo = TRUE
```

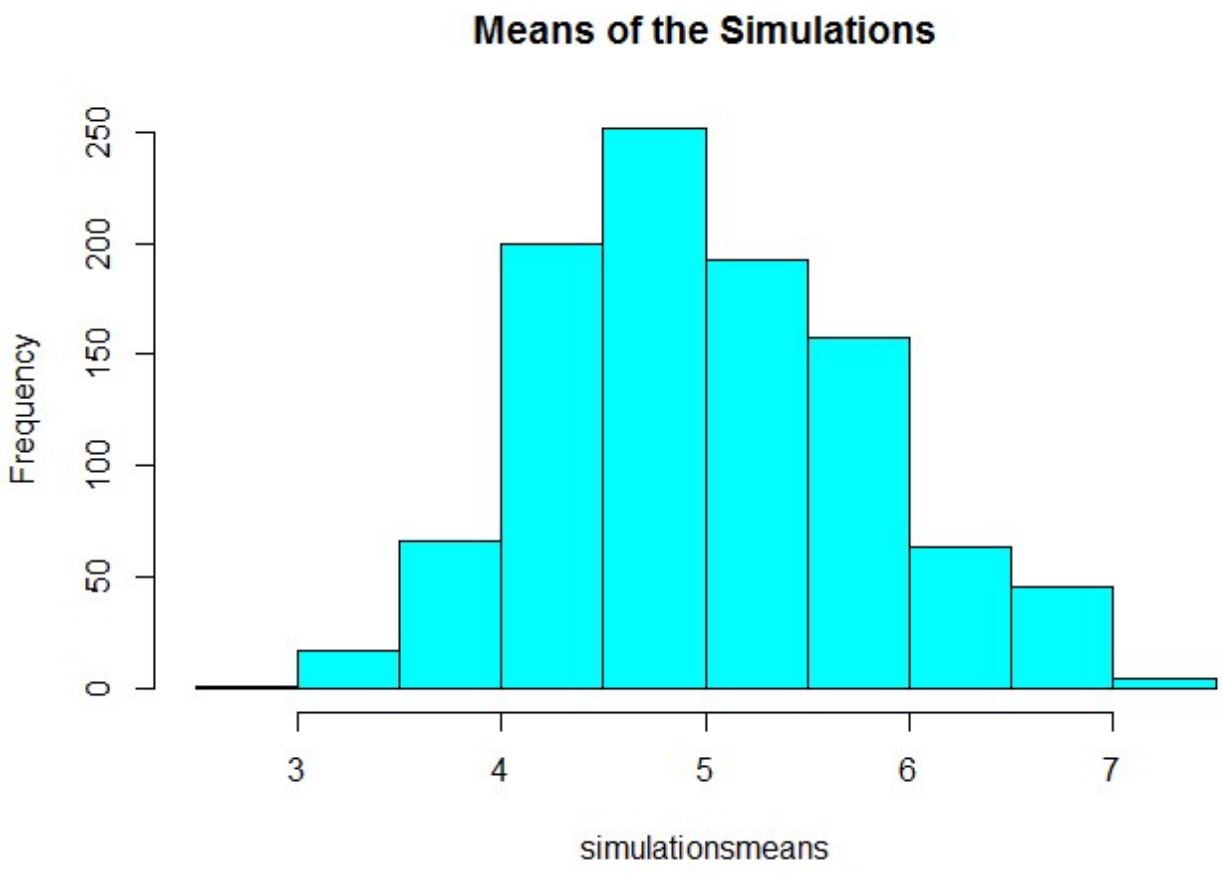
Create a matrix consisting of 1000 simulations of 40 exponentials

```
simulations <- matrix(rexp(n * samples, lambda), samples, n)
simulationsmeans <- rowMeans(simulations)
```

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

First lets plot the means

```
hist(main = "Means of the Simulations", simulationsmeans, col = "cyan")
```



Show the sample mean

```
MEAN <- mean(simulationsmeans)

MEAN
```

```
## [1] 5.011185
```

Show the theoretical mean (lambda^-1)

```
THEOMEAN <- lambda^-1

THEOMEAN
```

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

There is only a slight difference between the theoretical and sample means of the distribution

```
MEAN - THEOMEAN
```

```
## [1] 0.01118477
```

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

Show the sample variance

```
VARMEAN <- var(simulationsmeans)

VARMEAN
```

```
## [1] 0.6282926
```

Show the theoretical varicance (lambda * sqrt(n))^-2

```
VARTHEO <- (lambda * sqrt(n))^-2

VARTHEO
```

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

Again there is only a small difference between the sample and theoretical

```
VARMEAN - VARTHEO
```

```
## [1] 0.003292582
```

Show that the distribution is approximately normal

To accoplish this I will overlay a plot of a normal distribution to a histogram of the sample data

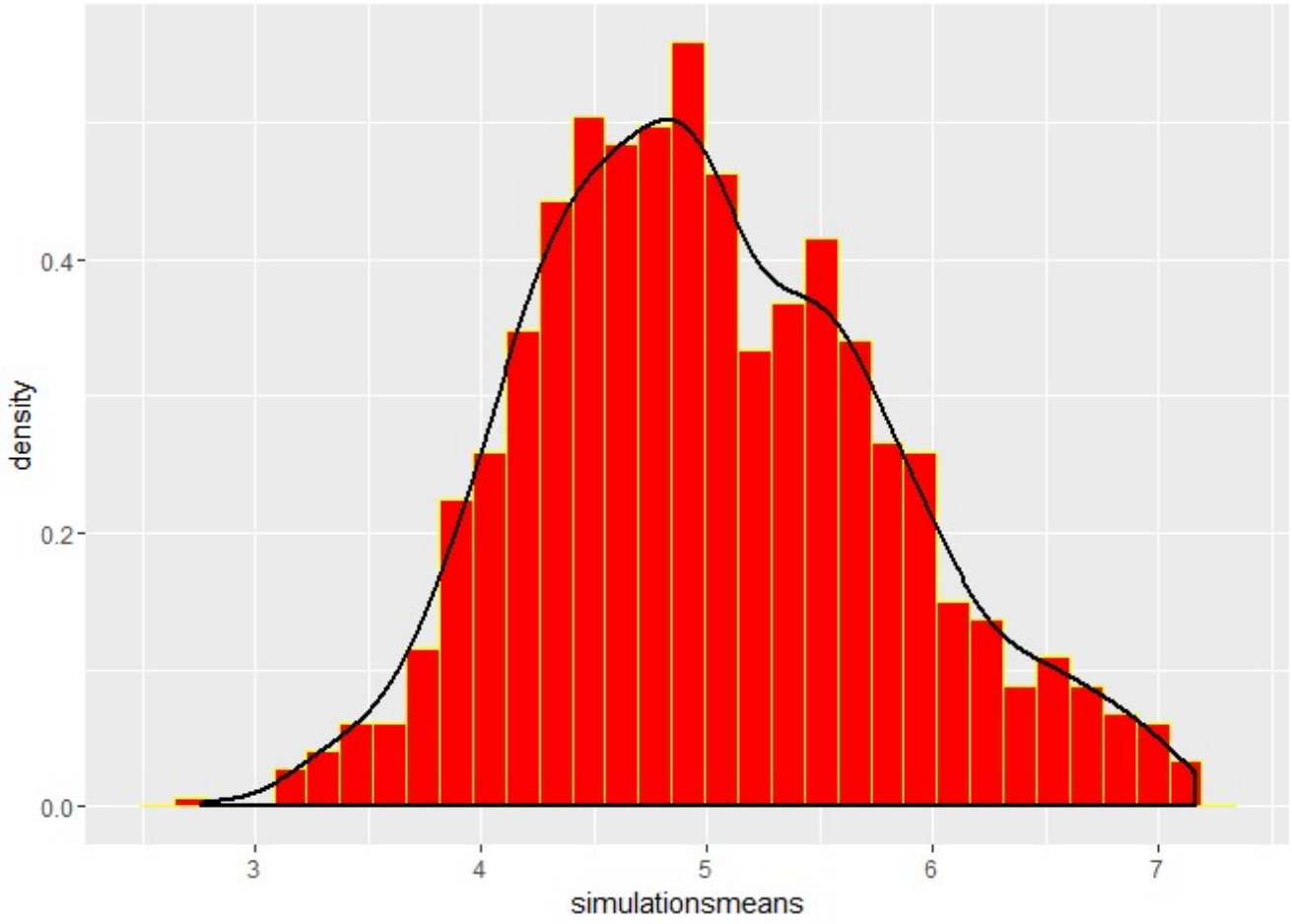
```
plotdata <- data.frame(simulationsmeans);

m <- ggplot(plotdata, aes(x =simulationsmeans))

m <- m + geom_histogram(aes(y=..density..), colour="yellow",
      fill = "red")

m + geom_density(colour="black", size=1);
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



This shows that the samples can be approximated with a normal distribution