Stastical Inference - Assignment 1 - Part 1

DomR

Saturday, September 13, 2014

Part 1

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also also 1/lambda. Set lambda = 0.2 for all of the simulations. In this simulation, you will investigate the distribution of averages of 40 exponential(0.2)s. Note that you will need to do a thousand or so simulated averages of 40 exponentials.

- Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s. You should
    1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.
    2. Show how variable it is and compare it to the theoretical variance of the distribution.
    3. Show that the distribution is approximately normal.
    4. Evaluate the coverage of the confidence interval for 1/lambda: $\bar{X} \pm 1.96 S_n$. #### Data Preparation

```
set.seed(1000)
lambda <- 0.2
noOfSimulations <- 1000
noOfObservations <-  40
```

- Run 1000 simulations to calculate mean of 40 observations with a rate parameter of 0.2.

```
exponentialsMean <- data.frame(sapply(1:noOfSimulations,function(s) mean(rexp(noOfObservations,lambda))))
colnames(exponentialsMean)[1] =  c("means")
simulatedMean <- mean(exponentialsMean$means)
simulatedMean
```

```
## [1] 4.987
```

```
simulatedStdDeviation <- sd(exponentialsMean$means)
simulatedVariance <- simulatedStdDeviation^2
simulatedVariance
```

```
## [1] 0.6543
```

```
expectedMean <- 1/lambda
expectedMean
```

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

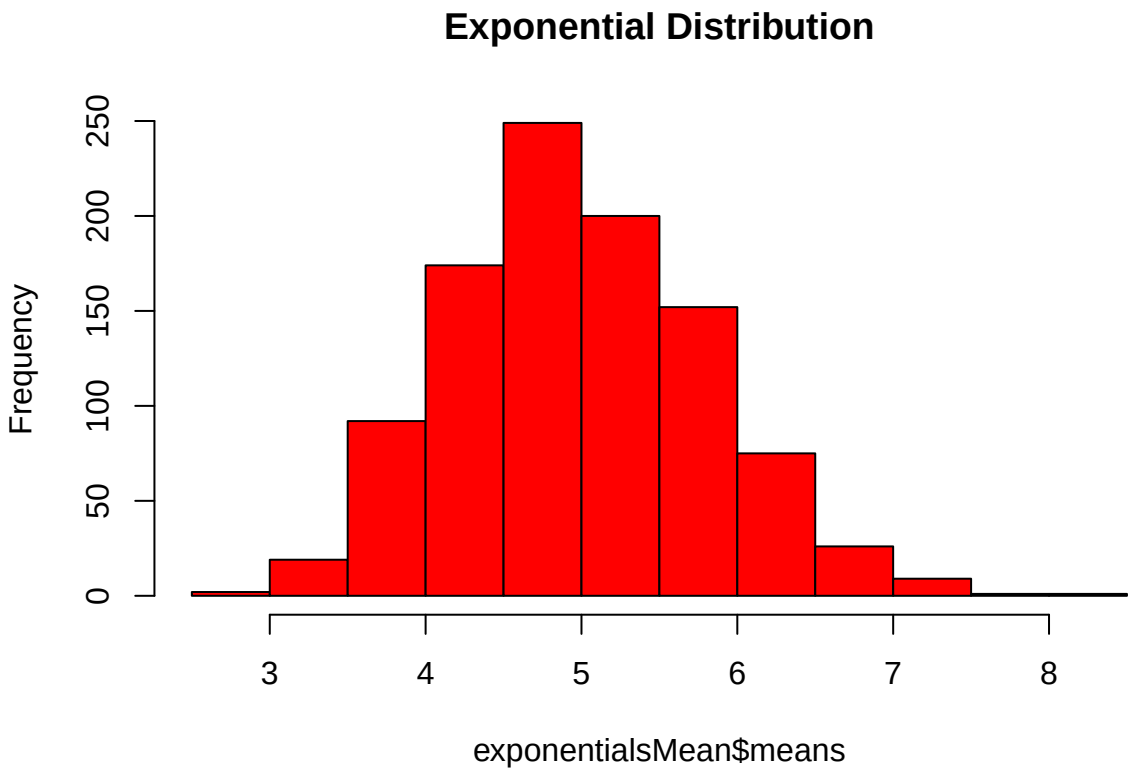
```
expectedStdDeviation <- 1/(lambda)/sqrt(noOfObservations)
expectedVariance <- expectedStdDeviation ^2
expectedVariance
```

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

Answers

1. Exponential distribution is centered at 4.987 which is close to expected mean of 5
2. Simulated variance of the distrubution is 0.6543 which is close to expected variance of 0.625
3. Showing historam of simulated means which shows the distribution is centered around 5

```
  hist(exponentialsMean$means, col="red",main="Exponential Distribution")
```



4. 95% Confidence Interval: simulatedMean +/- 1.96* std error

```
standardError <- simulatedStdDeviation/sqrt(noOfSimulations)
confidenceInterval <-   simulatedMean + c(-1,1) * 1.96 * standardError
confidenceInterval
```

```
## [1] 4.937 5.037
```