Peer-graded Assignment: Statistical Inference Course Project - Part 1

Santiago Ruiz Navas

11/12/2019

OVERVIEW

This document consits in two sections, part 1 corresponding to the simualtion part of the Statistical Inference Course Project.

1. Simulations

1.1 calculating the theoretical and simulated means

The theoretical mean was calculated by 1/lambda with lambda = 0.2.

```
theoretical.lambda <- 0.2
theoretical.mean <- 1/theoretical.lambda
theoretical.mean
```

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

The simulated mean was calculated by calculating the mean of 1000 averages of 40 randomly generated samples of the exponential distribution with lambda = 0.2

```
set.seed(1)
mns  <- vector("numeric",1000)
for (i in 1 : 1000) mns[i] = mean(rexp(n = 40,rate = theoretical.lambda))
sample.mean     <- mean(mns)
sample.mean
```

```
## [1] 4.990025
```

1.2 Calculating the theoretical and simulated variance

The theoretical variance of an exponential distribution is calcualte with 1/lambda^2

```
theoretical.var <- 1/theoretical.lambda^2
theoretical.var
```

```
## [1] 25
```

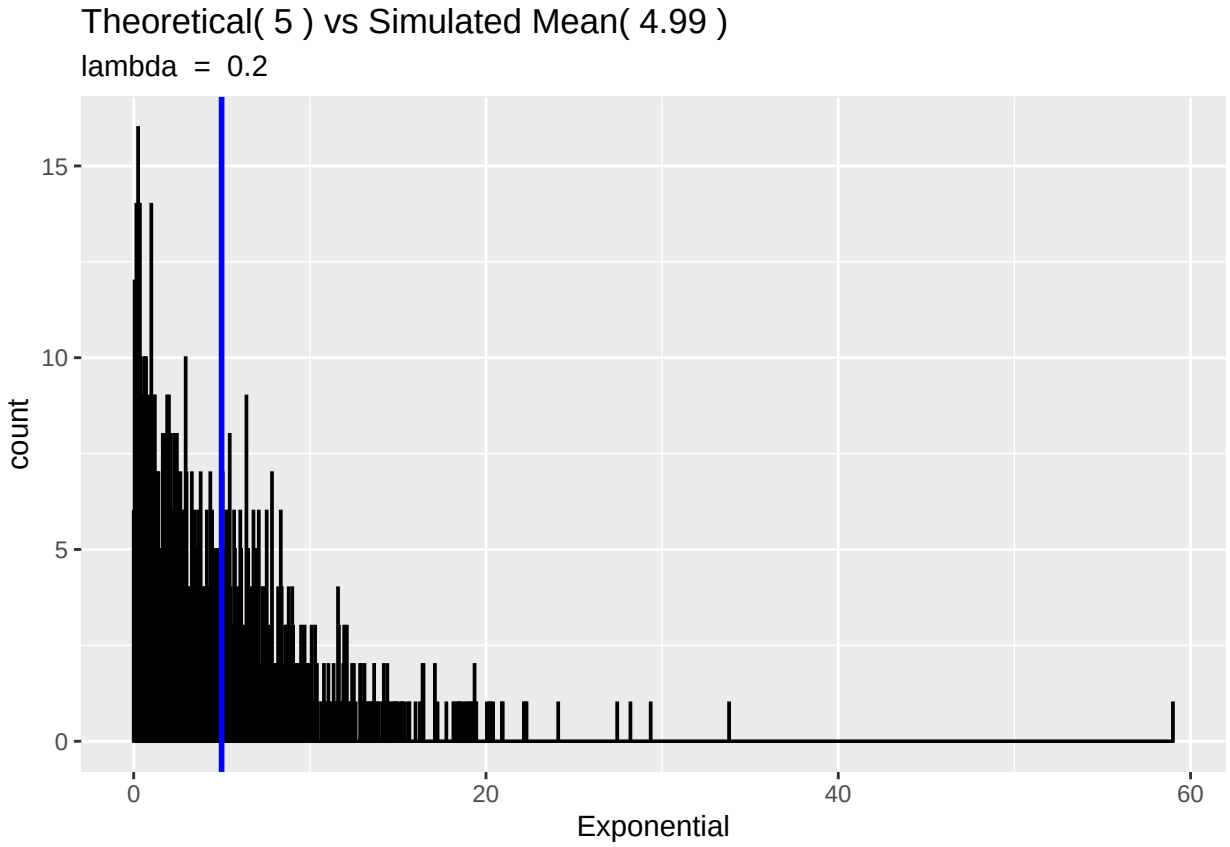
The simulated variance is calculated using the simualted mean obtained in point 1 and 1/lambda^2

```
sample.mean     <- mean(mns)
sample.lambda   <- 1/sample.mean
sample.var      <- 1/sample.lambda^2
sample.var
```

```
## [1] 24.90035
```

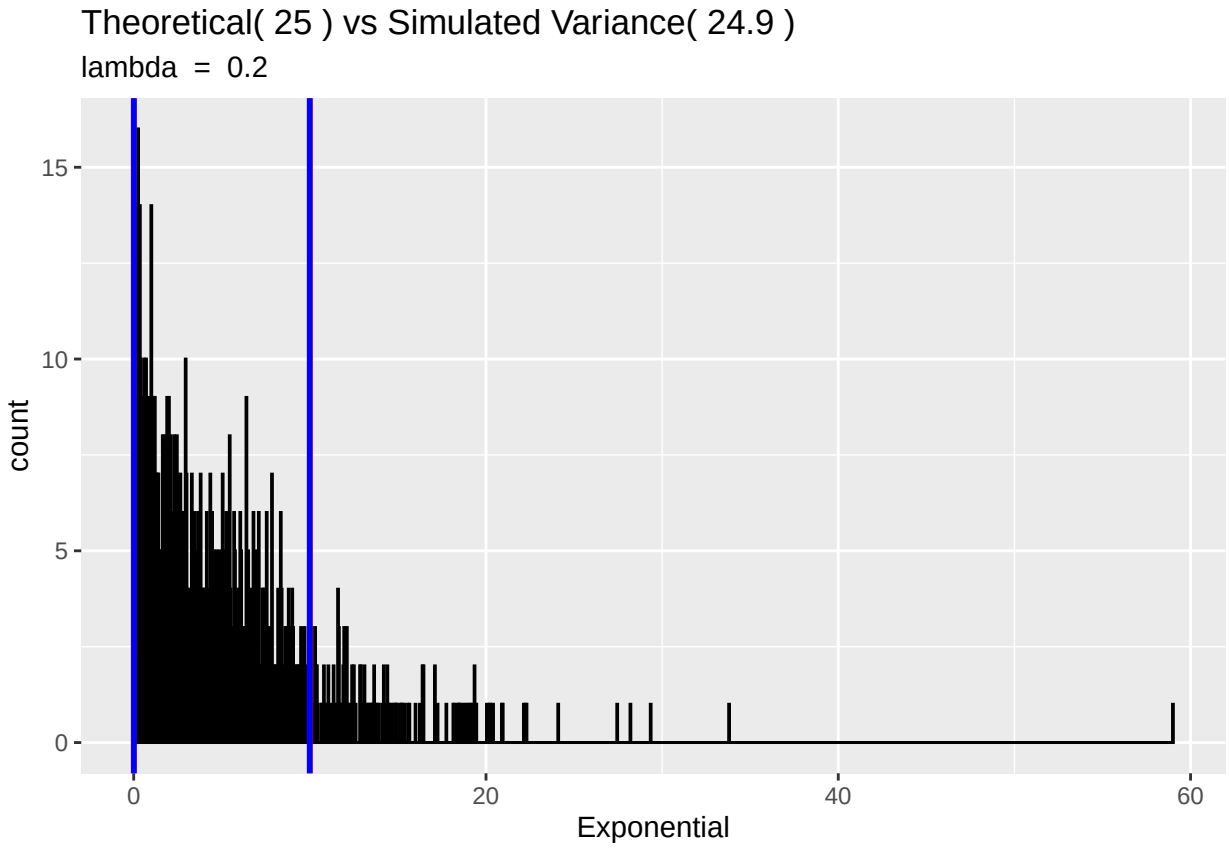
1.2.1 Simulated vs theoretical mean

I present in Figure bellow the distribution of the simulated data. Also two lines, in "blue" a vertical line showing the value of the theoretical mean (5) and in "red" the simulated mean



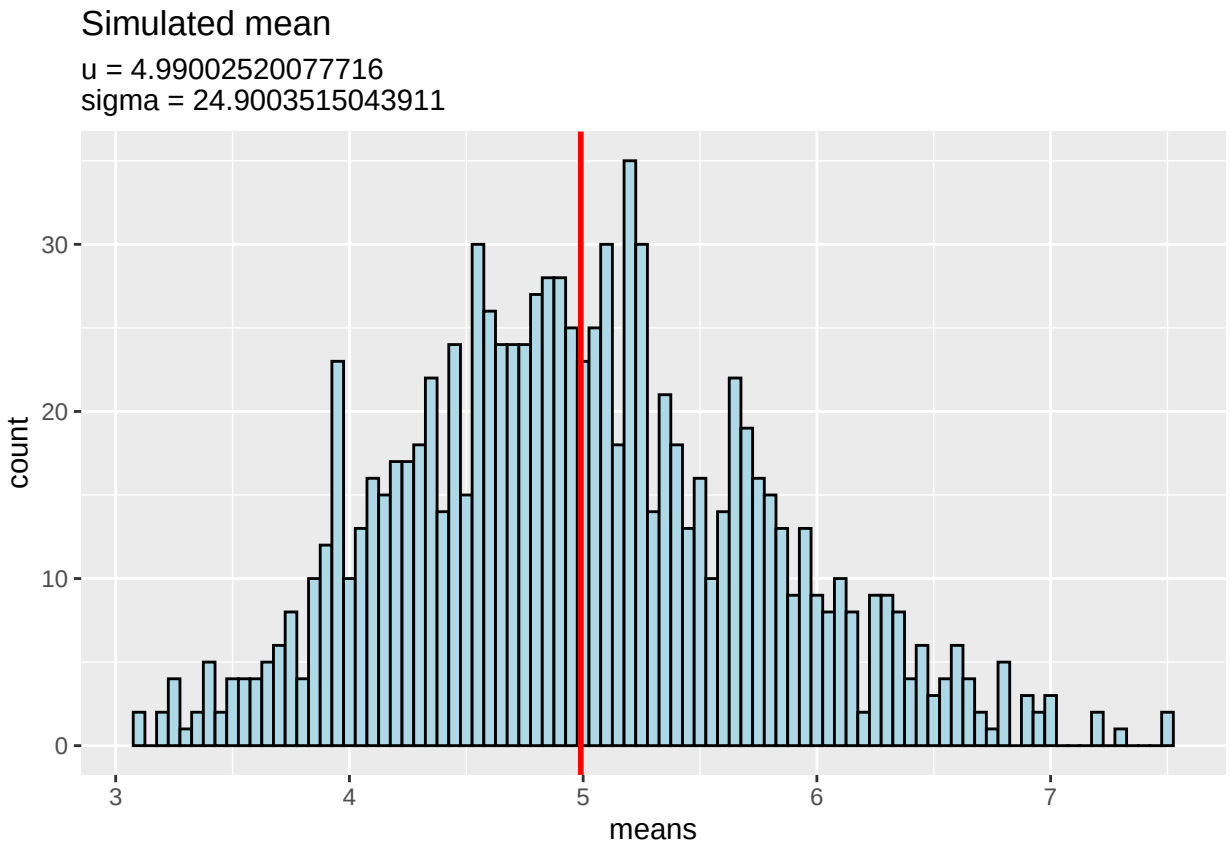
1.2.2 Simulated vs theoretical variance

Bellow it is possible to see the values of the theoretical and simulated square root of the variances



1.2.3 Distribution

In the figure bellow the histogram of the 1000 averages of 40 randomly generated samples of the exponential distribution with lambda = 0.2. The distribution shows a bell shaped curved resembling a normal distribu-



tion.

1.2.4 Interpretation

The differences between the theoretical and simulated means and variances was close to cero. Therefore, the simulated 40 averages of the exponential distributions drawn in this plot comply with the premises of the CTL.