Statistical Inference Assigment, Part A

Fabio Fornari

December 26, 2015

In this simulation, it will be investigated the distribution of averages of 40 exponentials. To do that we need to do a thousand simulated averages of 40 exponentials. 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. Lambda is set to 0.2 for all of the simulations. The code to do this is as follows.

```
set.seed(109)
nSim = 1000
n <- 40
lambda <- 0.2
simulations <- data.frame(x=replicate(nSim, mean(rexp(n, lambda))))
```

Questions 1

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

```
simulated_mean <- mean(simulations$x)
simulated_mean
```

```
## [1] 4.983461
```

The mean of the means of 1000 simulations of 40 exponential is 4.9835, which is very close to the expected mean (population mean) of 1/lambda = 5.0.

Questions 2

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

```
simulated_sd <- sd(simulations$x)
simulated_sd
```

```
## [1] 0.7923633
```

```
expected_sd <- (1/lambda)/sqrt(n)
expected_sd
```

```
## [1] 0.7905694
```

```
simulated_variance <- var(simulations$x)
simulated_variance
```

```
## [1] 0.6278395
```

```
expected_variance <- ((1/lambda)/sqrt(n))^2
expected_variance
```

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

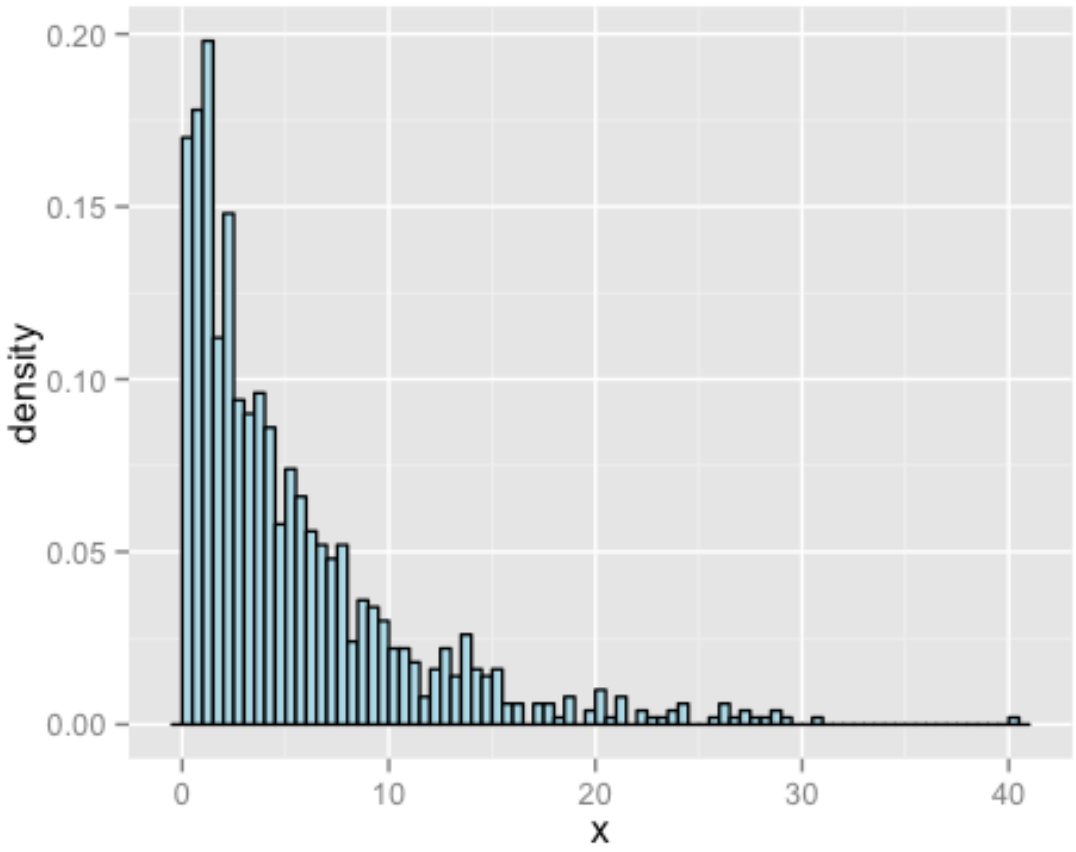
The simulated standard deviation is 0.7924, close to the expected standard deviation of 0.7906 The simulated variance of the distributions population is 0.628 close to the expetected variance of 0.625.

Question 3

Show that the distribution is approximately normal.

Below is a histogram plot of 1000 simulations of exponentials rexp(nSim, lambda). As can be seen this simulations is not normally distributed.

```
library(ggplot2)
set.seed(109)
simulations_2 <- data.frame(x=rexp(nSim, lambda))
g2 <- ggplot(data = simulations_2, aes(x = x)) +
      geom_histogram(aes(y=..density..), fill = 'lightblue',
                     binwidth = 0.5, color = 'black')
print(g2)
```



Below is a histogram plot of the means of the 1000 simulations of rexp(n, lambda). It is overlaid with a normal distribution with mean 5 and standard deviation 0.7906. As can be seen the simulations appears to be approximately normally distributed.

```
g <- ggplot(data = simulations, aes(x = x)) +
     geom_histogram(aes(y=..density..), fill = 'lightblue',
                    binwidth = 0.3, color = 'black')
g <- g + stat_function(fun = dnorm, arg = list(mean = 5, sd =
     expected_sd), size = 2, aes(colour = 'darkblue'))
g <- g + scale_colour_manual(name='Legend', values=c('darkblue'), labels='Normal
Distribution')
g <- g + theme(legend.position="bottom")
print(g)
```