Statistical Inference CP1 Part1

Simulation exercise

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.

For this simulation I will use 500, 1000 & 1500 simulated averages of 40 exponentials to show that the distribuion will be normally distributed based on the Centreal Limit Theorm.

This project will illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s.

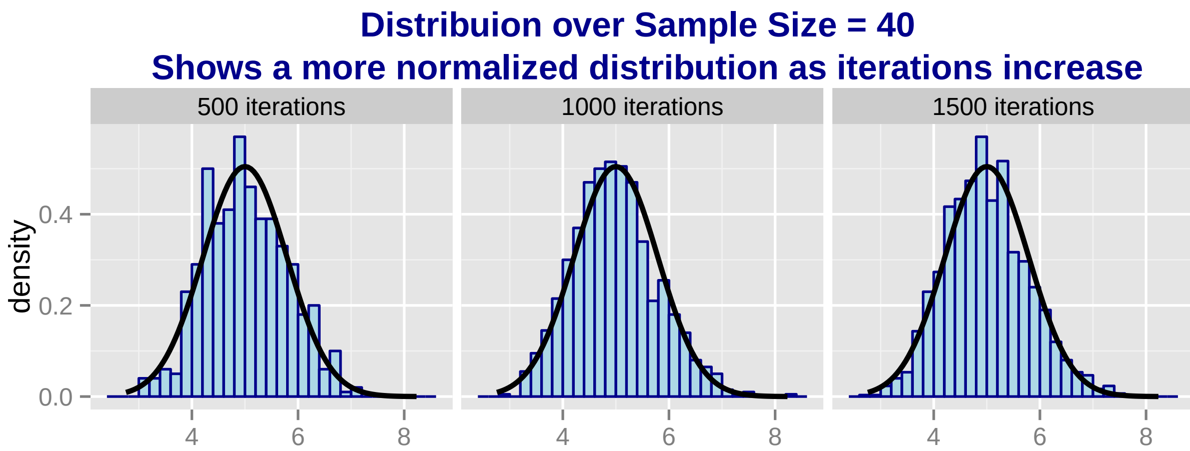
```
set.seed(1313)
lambda <- 0.2
mean <- 1/lambda
std <- 1/lambda
Sims <- c(500, 1000, 1500)
SampleSize <- 40
#Create 3 sets of samples with 500, 1000, and 1500 simulations
datamean <- lapply(Sims, function(x)
          {apply(matrix(rexp(x*SampleSize, rate = lambda),
               x, SampleSize), 1, mean)})
datastd <- lapply(Sims, function(x)
          {apply(matrix(rexp(x*SampleSize, rate = lambda),
               x, SampleSize), 1, sd)})

Sim1000Mean <- sapply(datamean[2], mean)
Sim1000var <- sapply(datamean[2],  var)
Sim1000Std <- sapply(datamean[2],  sd)
```

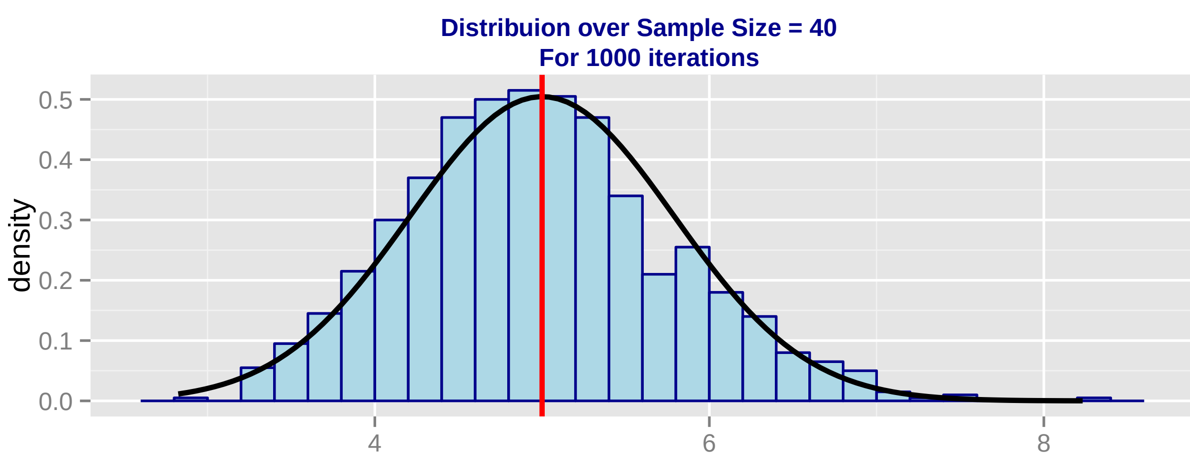
Show where the distribution is centered at and compare it to the theoretical center of the distribution. The distribuion of our 1000 Simulation = 4.979 which is be very close to the close to the theoretical mean of $1/\lambda = 1/0.2 = 5$.

Show how variable it is and compare it to the theoretical variance of the distribution. Based on the 1000 Simulation parameters, our variance = 0.6257. If you compare this result to the theoretical variance where the Standard Diviation = $1/\lambda = 5$ and the theoretical variance = std^2/ Sample Size = 5^2 / 40 = 0.625

Show that the distribution is approximately normal. To show how the Central Limit Theorm played a role in selecting 1000 simulations, the next graph will display 500, 1000 and 1500 simulations. As the graph shows, the more simulations that are completed, the more the distribution becomes normalized.



Taking a closer look a the 1000 iterations, you can see that the distribuion is very close to a normalized distribution.



Evaluate the coverage of the confidence interval for 1/lambda: X¯±1.96Sn???n

An approximation of a 95% conference interval for teh mean is calculated by $\bar{X} \pm 1.96\frac{S}{\sqrt{n}}$. The coverage is calculated by created a confidence interval for each of the 1000 simulations and determining how meny of them contain the true mean of 1/lambda = 5.

```
dmean <- matrix(unlist(datamean[2]), ncol = 1, byrow = TRUE)
dstd <- matrix(unlist(datastd[2]), ncol = 1, byrow = TRUE)

Coverage<-function(lambda) {
LowerLimit <- dmean - 1.96*dstd/sqrt(SampleSize)
UpperLimit <- dmean + 1.96*dstd/sqrt(SampleSize)
mean(LowerLimit < 1/lambda & UpperLimit > 1/lambda)}
par(Coverage(lambda))
```

Based on the above 1000 simulations, 93% of the samples contain the true mean of 1/lambda = 5.