Coursera Statistical Inference Project Part 1

Robert C Phillips

Sunday, September 21, 2014

Problem Statement

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.

We will use lambda = 0.2 for all of the simulations. In this simulation, we will investigate the distribution of averages.

```
lambda <- .2; nosim <- 1000; sampleSize <- 40; mu <- 1/lambda; sigma <- 1/lambda;
samples <- matrix(rexp(sampleSize * nosim, lambda), nosim)
```

Question 1

Show where the distribution is centered at and compare it to the theoretical center of the distribution.

```
samplesMean <- apply(samples, 1, mean)

#compute the mean of sample means, should be close to 1/lambda = 5
mresult <- data.frame(mean(samplesMean), mu)
colnames(mresult) <- c("sample mean","theoretical mean")
mresult
```

```
##   sample mean theoretical mean
## 1       4.996                5
```

Question 2

Show how variable it is and compare it to the theoretical variance of the distribution.

```
#compute the variance of sample means, should be close to sigma^2/sampleSize
vresult <- data.frame(var(samplesMean), sigma^2/sampleSize)
colnames(vresult) <- c("sample var","theoretical var")
vresult
```

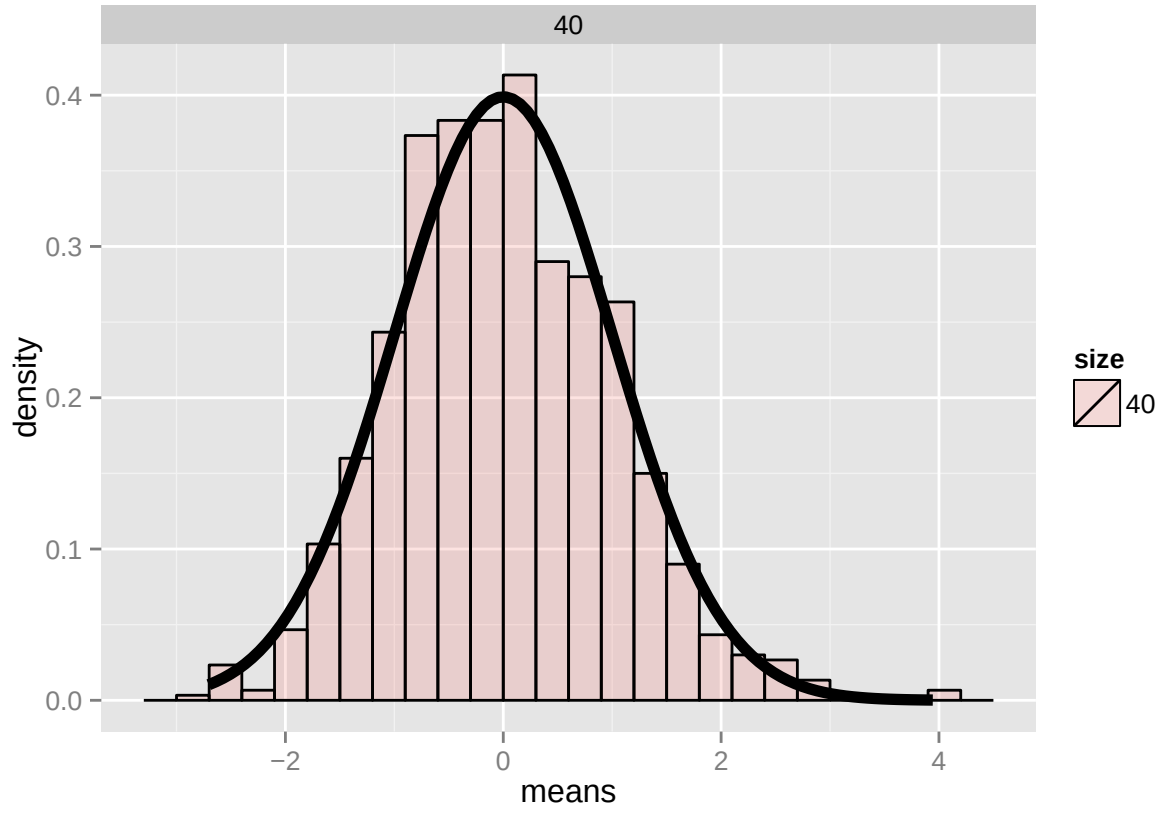
```
##   sample var theoretical var
## 1     0.6126           0.625
```

Question 3

Show that the distribution is approximately normal. . . normalize then plot.

```
cFun <- function(x,n) sqrt(n) * (mean(x) - mu)/sigma

dat <- data.frame(
   means = c(apply(samples, 1, cFun, sampleSize)),
   size = factor(rep(c(sampleSize), nosim)))
```



Question 4

Evaluate the coverage of the confidence interval for 1/lambda: $\bar{X} \pm 1.96\frac{S}{\sqrt{n}}$

```
cFun2 <- function(x,n) mean(x) + c(-1,1) * 1.96/sqrt(n)*sd(x)
samplesci <- t(apply(samples, 1, cFun2, sampleSize))
total <- sum(apply(samplesci, 1, function(i) if (mu >= i[1] && mu <= i[2]) 1 else 0))

cresult <- data.frame(total/nosim, .95)
colnames(cresult) <- c("sample ci","theoretical ci")
cresult
```

```
##   sample ci theoretical ci
## 1     0.927           0.95
```