

Statistical Inference Project

Fatemeh Abyarjoo

Saturday, April 25, 2015

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. 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 $1/\lambda$. Set `lambda = 0.2` for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

```
n=40
lambda=0.2
set.seed(500)
meanvector=NULL
for(i in 1:1000) {
  meanvector <- c(meanvector, mean(rexp(n, lambda)))
}
head(meanvector)
```

```
## [1] 4.958067 5.178396 5.683279 5.619003 4.433293 5.867299
```

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

Sample mean:

```
mean(meanvector)
```

```
## [1] 5.010562
```

Theoretical mean:

```
1/lambda
```

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

Sample mean and theoretical mean are close.

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

Variance of samples:

```
var(meanvector)
```

```
## [1] 0.6201215
```

Theoretical variance:

```
((1/lambda)/(sqrt(40)))^2
```

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

Two variances are close.

3. Show that the distribution is approximately normal.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
mydata <- as.data.frame(meanvector)
#hist(meanvector, col="Yellow", xlim=c(0,10), xlab="Sample Mean", main="Histogram of #Sample Means")

ggplot(data = mydata, aes(x = meanvector), xlim=c(0,10)) +
  geom_histogram(aes(y=..density..), fill = I("deeppink"),
    binwidth = 0.10, color = I('darkviolet')) +
  stat_function(fun = dnorm, arg = list(mean = 5, sd = sd(meanvector)))
```

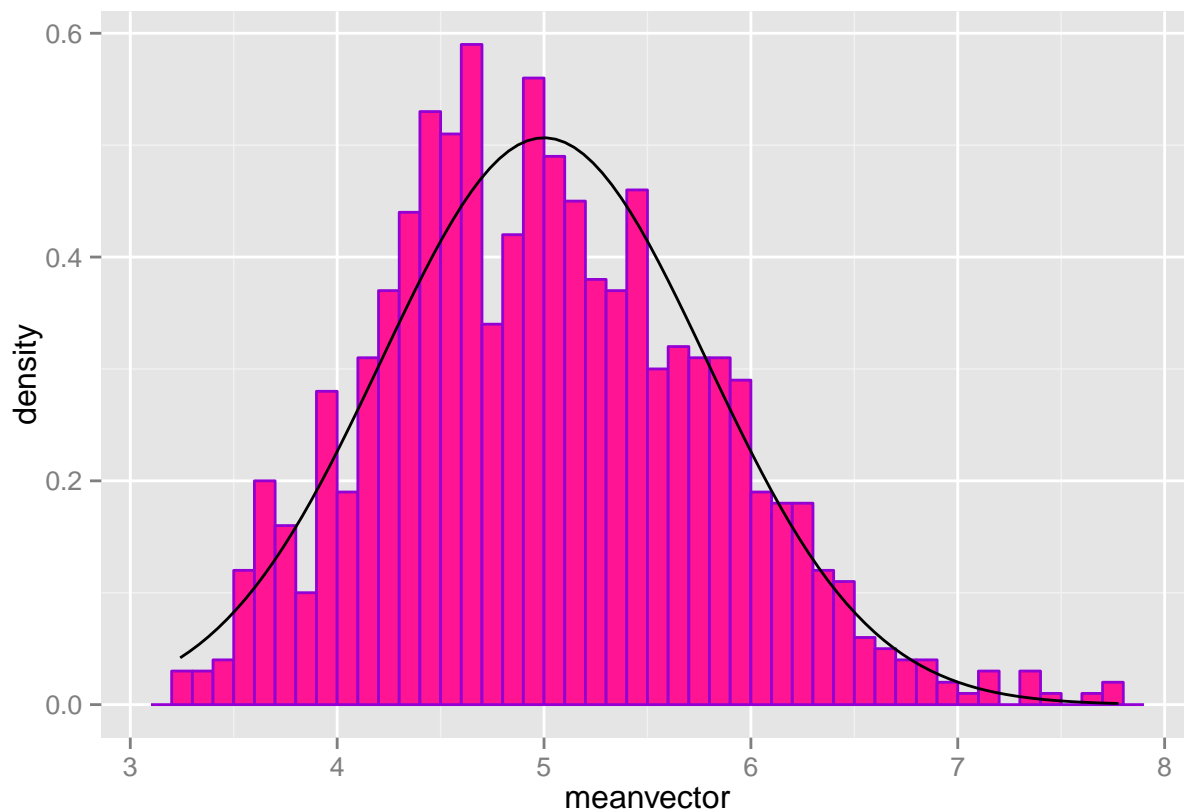


Figure is the histogram of distribution of means of random samples overlaid with normal distribution with mean equal to 5. It shows that the distribution is approximately normal.