Comparing Sample exponential distribution with theoretical distribution

Han

Overview

In this project I will explore the exponential distribution with theoritical distribution (Central Limit Theorem). I will show as followings:

- show the sample mean and compare it to the theoretical mean of the distribution.
- show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.
- show that the distribution is approximately normal.

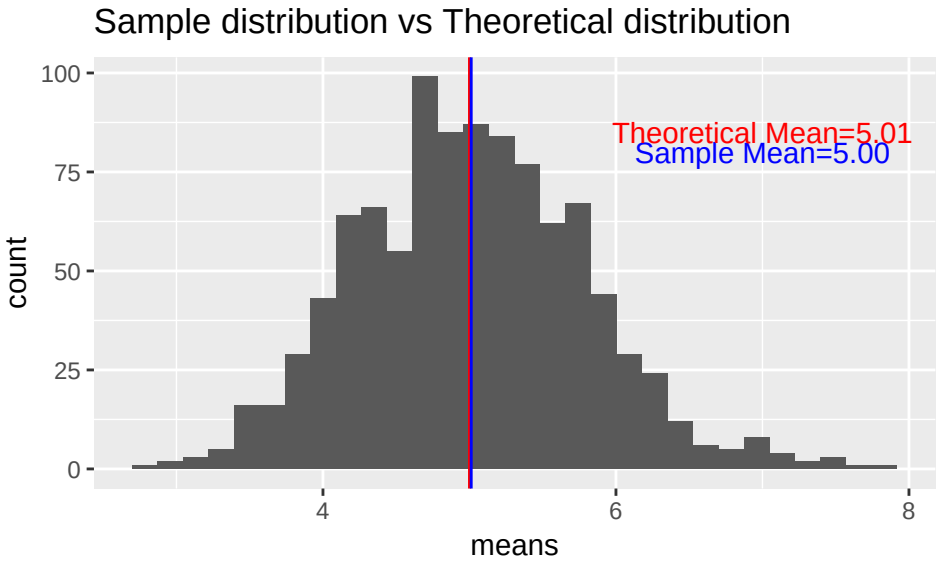
Simulations

I set lambda as 0.2, and the distribution of 1000 with n=40 exponential distribution. I used *rexp* R function in order to create exponential distribution. Also, I used *apply* R function for getting mean values of each exponential distribution. The following histogram is the result of the code, which create sample exponential mean distribution.

```
library(ggplot2)
lambda<-0.2
n<-40
number<-1000

set.seed(123)
exponential<-matrix(data=rexp(n*number,lambda),nrow=number)
exponential_mean<-data.frame(means=apply(exponential,1,mean))
mu<-1/lambda
mean1<-mean(exponential_mean$means)

g<-ggplot(data=exponential_mean,aes(x=means))
g+geom_histogram(bins=30)+
        labs(title="Sample distribution vs Theoretical distribution")+
        geom_vline(xintercept=mu,color='red')+
        geom_vline(xintercept=mean1,color='blue')+
        annotate('text',x=7,y=80,label='Sample Mean=5.00',color='blue')+
        annotate('text',x=7,y=85,label='Theoretical Mean=5.01',color='red')
```



Sample Mean versus Theoretical Mean

As we see in the above histogram, sample mean is similar to theoretical mean.

```
print(mu)
```

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

```
print(mean1)
```

```
## [1] 5.011911
```

Sample Variance versus Theoretical Variance

I checked both sample variance and theoretical variance. The values were similar like this:

```
sd1<-sd(exponential_mean$means)
var1<-var(exponential_mean$means)
sd<-1/lambda/sqrt(n)
var<-sd^2

print(var)
```

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

```
print(var1)
```

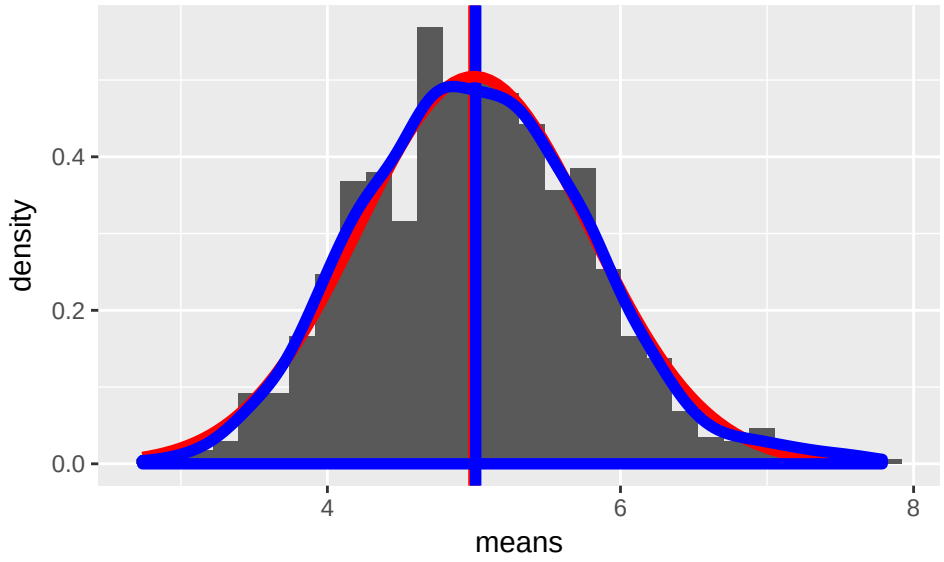
```
## [1] 0.6088292
```

Distribution

Finally, I show that the distribution is approximately normal.

```
## Sample Mean vs. Theoretical Mean
g+geom_histogram(bins=30,aes(y=..density..))+
        geom_vline(xintercept=mu,size=2,color="red")+
```

```
        geom_vline(xintercept=mean1,size=2,color="blue")+
        stat_function(fun=dnorm,args=list(mean=mu,sd=sd),size=2,color="red")+
        geom_density(size=2,color="blue")
```



The sample distribution (blue) seems to be similar to the normal distribution (red).