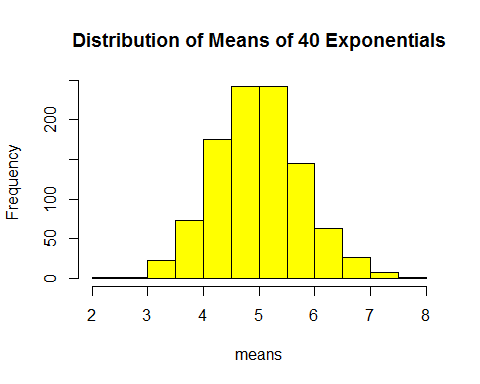
Investigate the exponential distribution in R

Aijing Gao

2015年12月23日

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

library(ggplot2)  
library(knitr)  
#Simulating exponential distribution  
set.seed(876)  
nsim=1000;n=40;lambda=0.2  
means=NULL  
for(i in 1:nsim) means=c(means,mean(rexp(n,lambda)))  
hist(means,col="yellow",main="Distribution of Means of 40 Exponentials")



#Sample means vs Theoretical means  
Sample\_mean=mean(means);Theoretical\_mean=1/lambda  
print(c(Sample\_mean,Theoretical\_mean))

## [1] 4.991893 5.000000

According to the result, sample mean derived from simulation is quite near to theoretical mean.

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

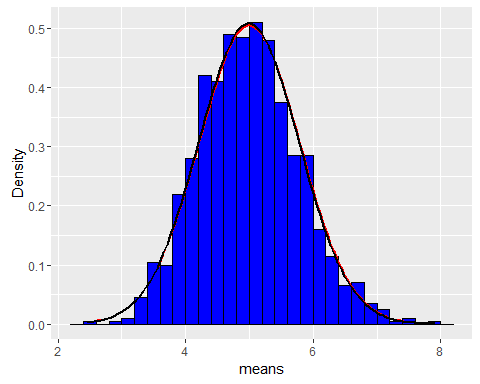
Sample\_variance=var(means)  
Theoretical\_variance=((1/lambda)^2)/n  
print(c(Sample\_variance,Theoretical\_variance))

## [1] 0.6168217 0.6250000

According to the above result, sample variance and theoretical variance are nearly equal.

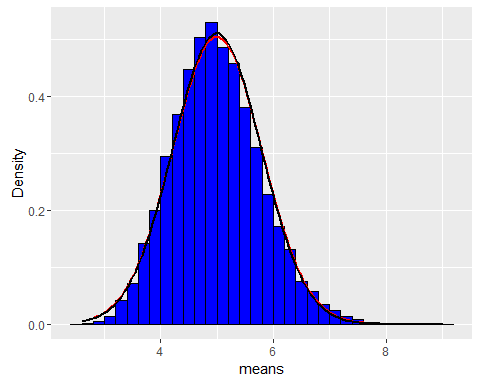
[3].Show that the distribution is approximately normal

g<-ggplot(data.frame(means),aes(x=means))  
g<-g+geom\_histogram(binwidth=0.2,aes(y=..density..),colour="black",fill="blue")  
g<-g+labs(x="means",y="Density")  
g<-g+stat\_function(fun=dnorm,args=list(mean=Theoretical\_mean, sd=sqrt(Theoretical\_variance)),color = "red", size = 1.0)  
g<-g+stat\_function(fun=dnorm,args=list(mean=Sample\_mean,sd=sqrt(Sample\_variance)),color="black",size=1.0)  
g

 Red line indicates the normal distribution of theoretical means, black line indicates the normal distribution of simulating 40 expontentials.

In order to show that the distribution is approximately normal, we could increase our simulating numbers to a large number (10,000).

set.seed(12345)  
nsim=10000;n=40;lambda=0.2  
#Recalculate  
means=NULL  
for(i in 1:nsim) means=c(means,mean(rexp(n,lambda)))  
Sample\_mean=mean(means);Theoretical\_mean=1/lambda  
Sample\_variance=var(means)  
Theoretical\_variance=((1/lambda)^2)/n  
g<-ggplot(data.frame(means),aes(x=means))  
g<-g+geom\_histogram(binwidth=0.2,aes(y=..density..),colour="black",fill="blue")  
g<-g+labs(x="means",y="Density")  
g<-g+stat\_function(fun=dnorm,args=list(mean=Theoretical\_mean, sd=sqrt(Theoretical\_variance)),color = "red", size = 1.0)  
g<-g+stat\_function(fun=dnorm,args=list(mean=Sample\_mean,sd=sqrt(Sample\_variance)),color="black",size=1.0)  
g

 As shown above, the simulating exponential distribution becomes quite close to normal distribution.