Analysis of exponential distribution

Sayef Ishauqe

October 08, 2017

Overview:

This assignment aims to analyze a small sample of exponentials to determine its distribution. Primarily, computer simulation and resampling techniques are utilized to determine and compare the characterstics with normal distribution.

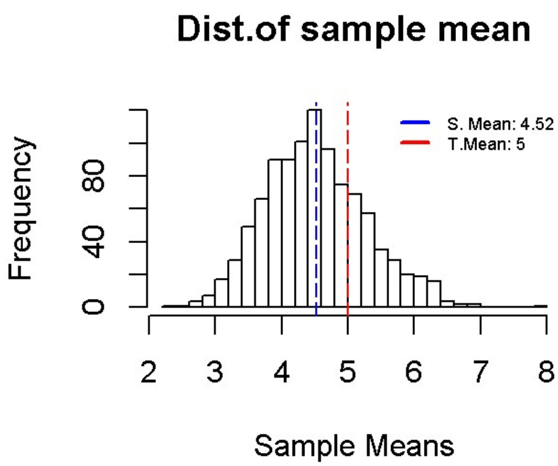
Simulations:

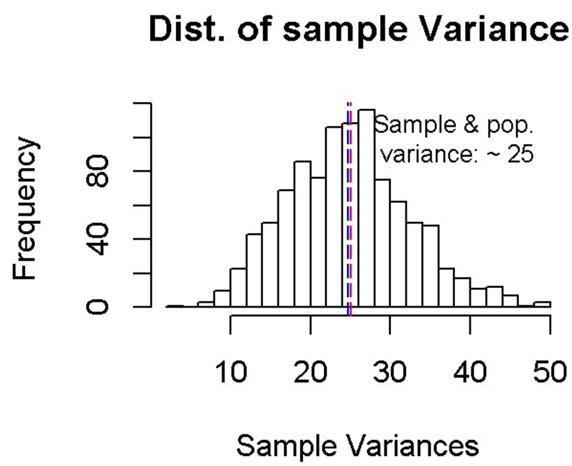
- Step 1: Generate sample data of exponential distribution
- Step 2: Examine the distribution of the sample data

```
n<-40;lambda<-0.2;set.seed(1000)
x<-rexp(n,lambda)
ms<-mean(x); sds<-sd(x); vars<-var(x)
# hist(x,breaks = 10,
#      main="Distribution of 40 exponetials (sample)")
# text(x=10,y=10,paste("Sample mean:",round(ms,2),"\nStandard Deviation:",round(sds,
2)))
```

- Step 3: Generate simulation data (1000 simulations, 40 exponentials)
    - Calculate mean and variance for each simulation
    - Plot histogram for mean and variance

```
sim<-1000
simData<-matrix(sample(x,n*sim,replace=T),sim,n)
means<-apply(simData,1,mean); vars<-apply(simData,1,var); sds<-apply(simData,1,sd)
par(mfrow=c(1,2))
hist(means,breaks = 20,
     main=paste("Dist.of sample mean"),
     xlab="Sample Means")
abline(v=mean(means),lty=5,col="blue")
abline(v=1/lambda,lty=5,col="red")
legend("topright", c(paste("S. Mean:",round(mean(means),2)),paste("T.Mean:",1/lambd
a)),
       cex=0.55, col=c("blue","red"), lty=5, lwd=2, bty="n");
hist(vars,breaks = 20,xlab="Sample Variances",
     main=paste("Dist. of sample Variance"))
abline(v=mean(vars),lty=5,col="blue")
text(38,100,cex=.8, paste("Sample & pop.\n variance: ~",round(mean(vars),0)))
abline(v=(1/lambda)^2,lty=5,col="red")
```

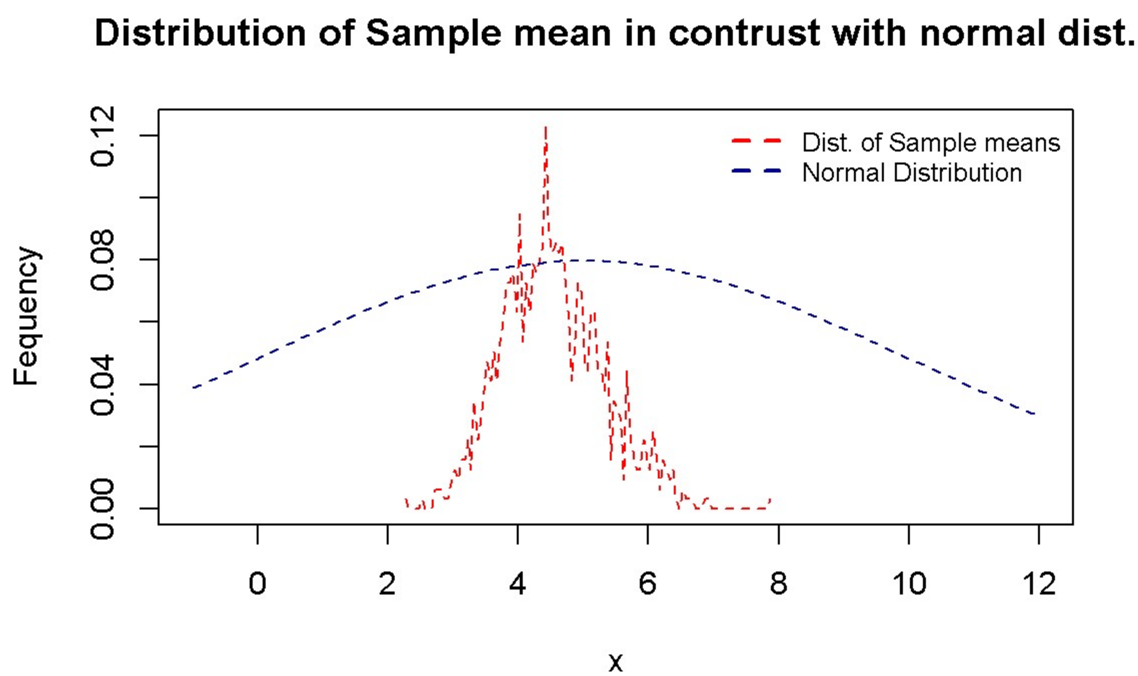




Findings - After 1000 simulation, average of sample means are very close to Theritical mean. The sample variance is also very close to theoritical population variance. *however, this is based on small sample size and number of simulations. Also, the result is highly dependent on random functions available in R / computer*

- Step 4: Compare the distribution of sample means with theoritical normal distribution curve
    - Calculate density of sample means
    - Plot density of mean with density distribution of normal distribution with same mean and st. deviation

```
hh<-h$density;xx<-h$mids
plot(xx,y=hh/sqrt(40),type ="l",lty=2,col="red",xlim=c(-1,12),
     main="Distribution of Sample mean in contrust with normal dist.",
     xlab="x",
     ylab="Fequency")
curve(dnorm(x,mean=1/lambda,sd=1/lambda), lty=2,
          col="darkblue",from=-1,to=12, n=100,add=TRUE)
legend("topright", c("Dist. of Sample means","Normal Distribution"),
       cex=0.8, col=c("red","darkblue"), lty=2, lwd=2, bty="n");
```



Findings - The Distribution of sample means are comparable to normal Distribution curve. However, sample means are more concentrated at the center (at least visually).

- Step 5: Revalidate findings through central limits
  - Calculate number of sample means within 1,2,3 standard error
  - Calculate density and compare it with p-values of normal distribution

```
u<-vector();v<-vector(); result<-data.frame()
result<-cbind(c("1 SD", "2 SD", "3 SD"))
for(i in 1:3)
{
     v[i]<-round(pnorm(i,mean = 0,sd=1),2)
     limit<-mean(means)+c(-1,1)*i*sqrt(var(vars))/sqrt(40)
     u[i]<-round(sum(limit[1]<means & limit[2]>means)/1000,2)
}
result<-cbind(result,v,u)
colnames(result)<-c("Limit      .","Dens.Normal      .", "Dens.s.means   .")
knitr::kable(result,format="html")
```

Limit .	Dens.Normal .	Dens.s.means .
1 SD	0.84	0.89
2 SD	0.98	1
3 SD	1	1

Findings - The density within 1,2,3 Standard deviation are comparable. So, it holds the characterstics of normal distribution at this point.