Probability and Statistics
Lab n^o 1 Report – Cyril Naves

1 Random Variables

1. Generate 3 random vectors of size 10000 from different distributions :

   - A uniform distribution between 0 and 1.
   - A normal distribution $N(0,10)$
   - A exponential distribution of parameter $\lambda = 2$

   (a) What is the number of bins to be used to represent the corresponding histograms according to Sturge's rule?

   Ans:

   By Sturges Rule:

   No of bins:

```
> 1+log2(10000)
[1] 14.28771
```

   Uniform Distribution

```
> nclass.Sturges(runif(10000,0,1))
[1] 15
```

   No of bins:15

   Normal Distribution

```
> nclass.Sturges(rnorm(10000,0,sqrt(10)))
[1] 15
>
```

   No of bins:15

   Exponential distribution

```
> nclass.Sturges(rexp(10000,2))
[1] 15
```

   No of bins:15

(b) What is the bin size according to the Normal Reference rule?

Ans:

Uniform Distribution:

```
> sdrunif<-runif(10000,0,1)
> sd<-sd(runif(10000,0,1))
> 3.5*sd*10000^-(1/3)
[1] 0.04714702
```

Bin Size: 0.04714702

Normal Distribution:

```
> sd<-sd(rnorm(10000,0,sqrt(10)));3.5*sd*10000^-(1/3)
[1] 0.5145692
```

Bin Size: 0.5145692

Exponential Distribution:

```
> sd<-sd(rexp(10000,2))
> 3.5*sd*10000^-(1/3)
[1] 0.08206662
```

Bin Size:0.0820662

(c) What is the number of bins for each sample vector you have generated according to the Normal Reference Rule ?

Ans:

No of bins for Uniform Distribution:

No of Samples/BinSize= 10000/0.04714702 =212102.5

```
> 10000/0.04714702
[1] 212102.5
```

No of bins for Normal Distribution:

No of Samples/BinSize= 10000/1.616809 = 19433.73

```
> 10000/0.5145692
[1] 19433.73
```

No of bins for Exponential Distribution:

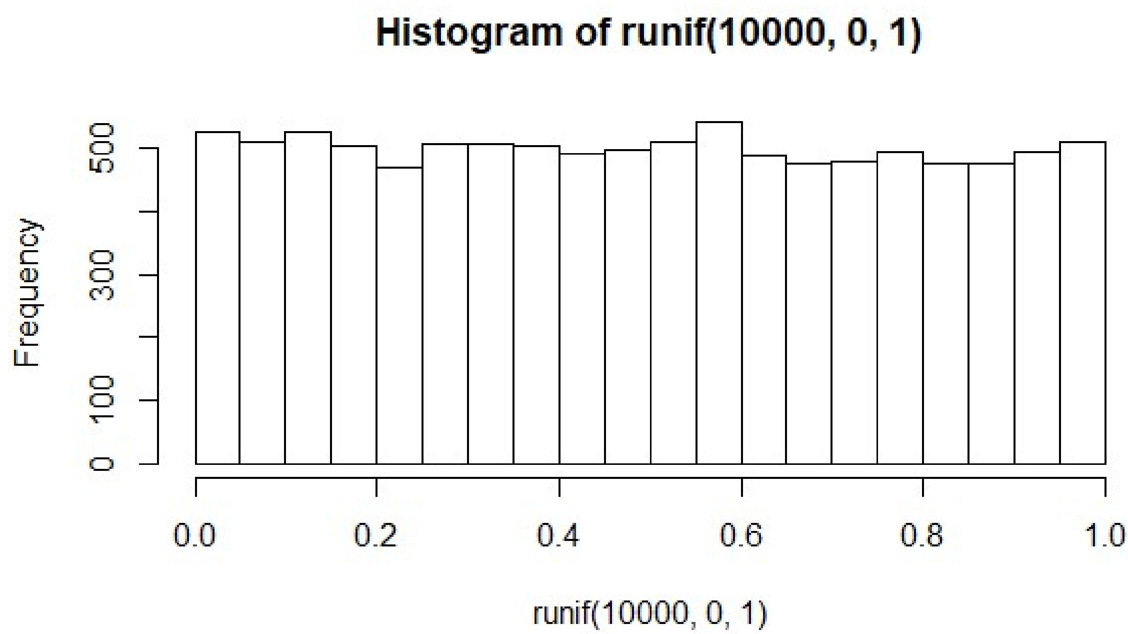
No of Samples/BinSize= 10000/0.0820662=121852.8

```
> 10000/0.0820662
[1] 121852.8
```

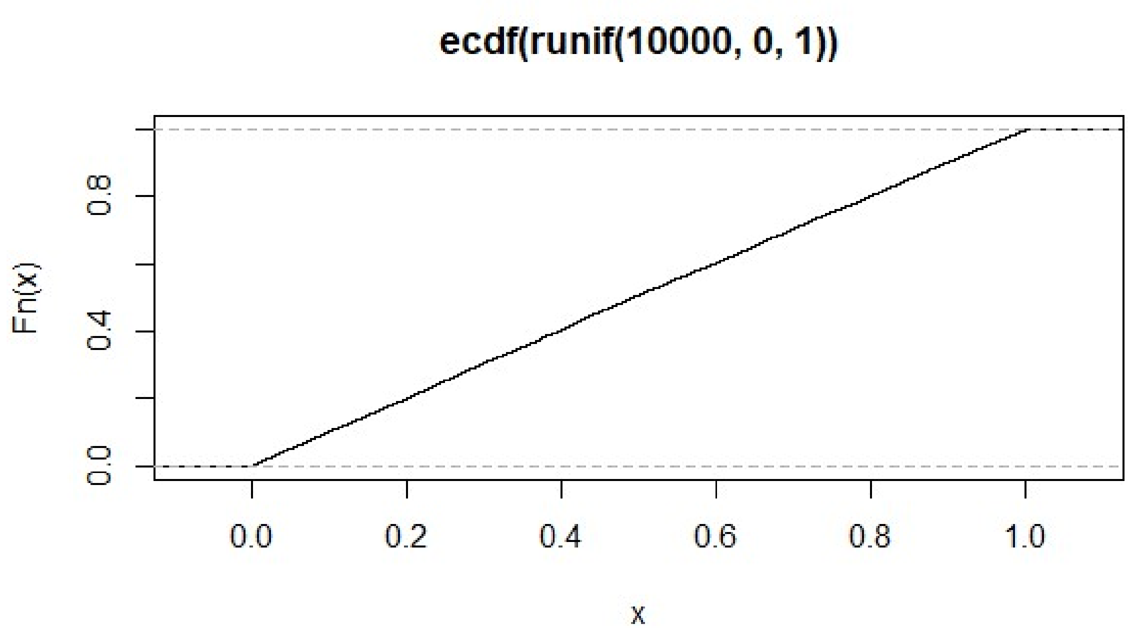
(d) Represent the histograms (R is using Sturge's rule with improvements, hence you can just use hist(X)) , cdfs and boxplots of each random vector.

Uniform Distribution:

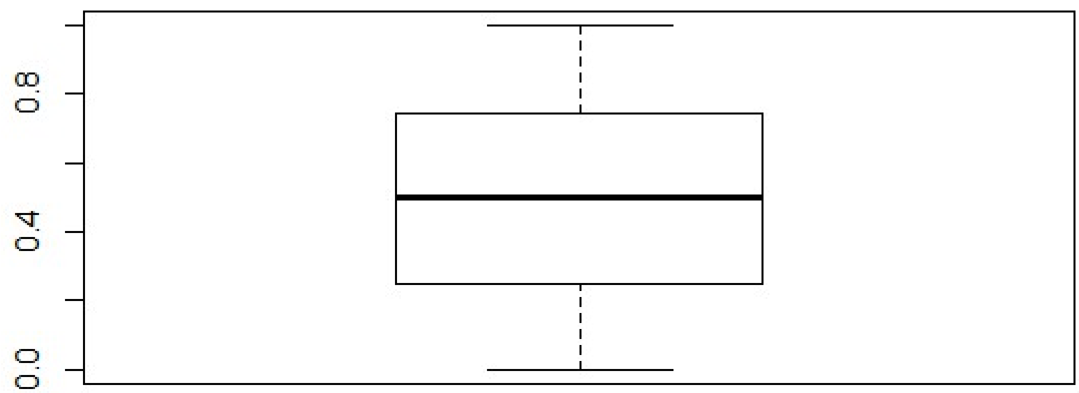
```
> hist(runif(10000,0,1))
```




```
> plot(ecdf(runif(10000,0,1)))
```

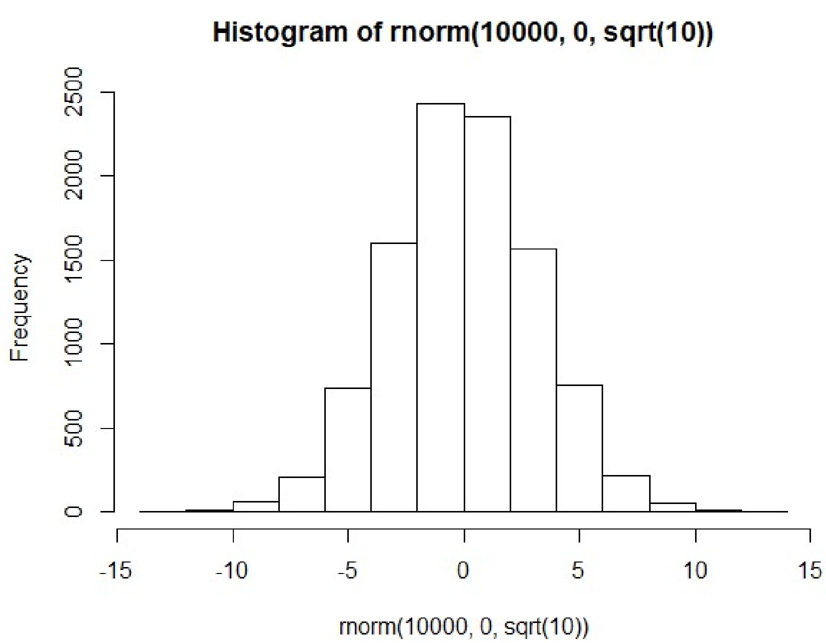


```
> boxplot(runif(10000,0,1))
```

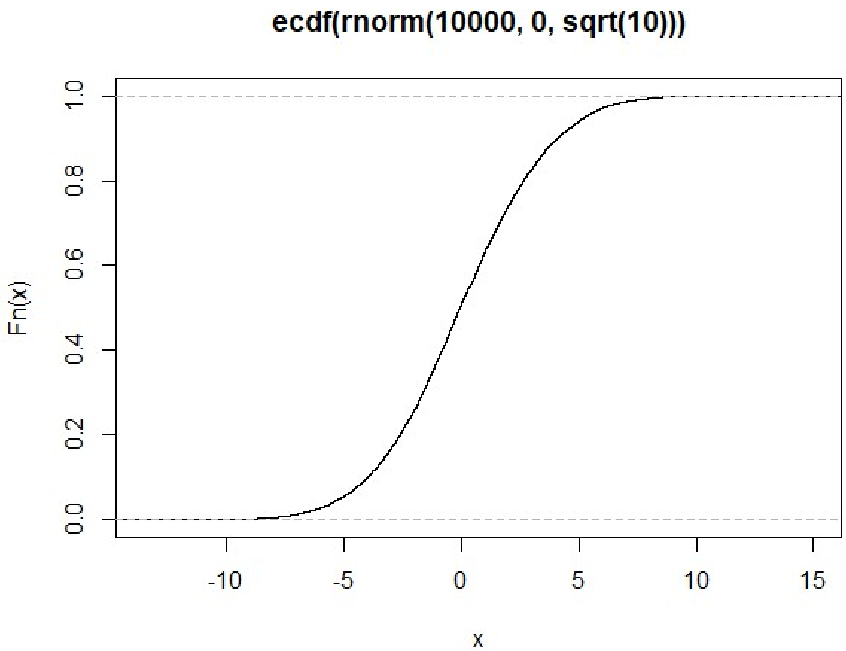


Normal Distribution:

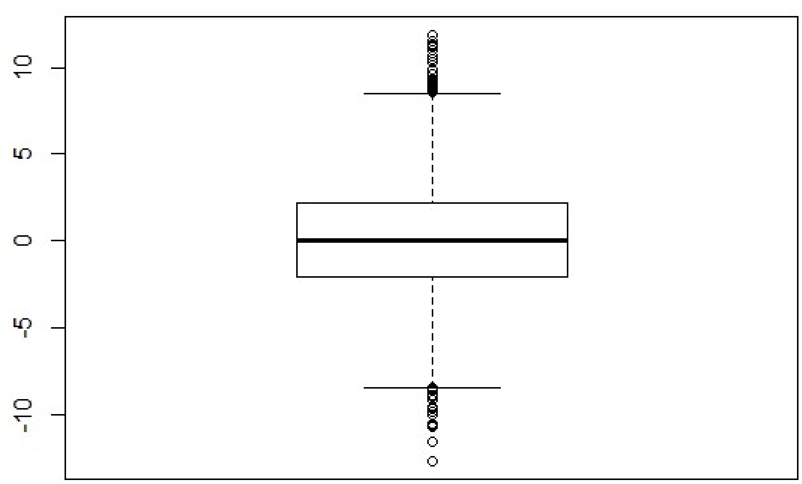
```
> hist(rnorm(10000,0,sqrt(10)))
```



```
> plot(ecdf(rnorm(10000,0,sqrt(10))))
```

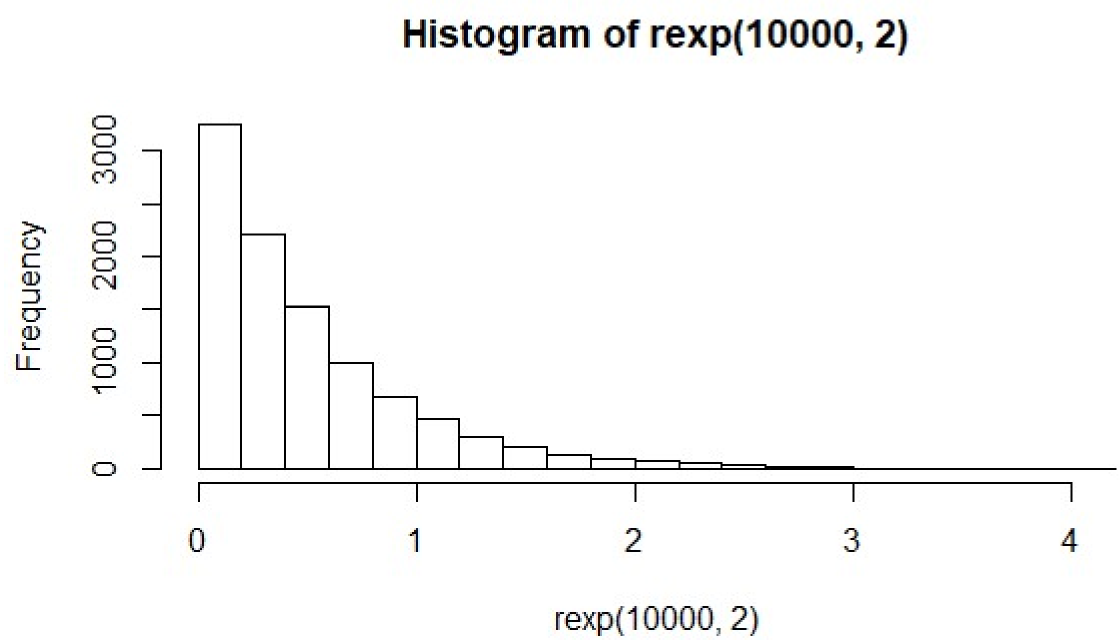



```
> boxplot(rnorm(10000,0,sqrt(10)))
```

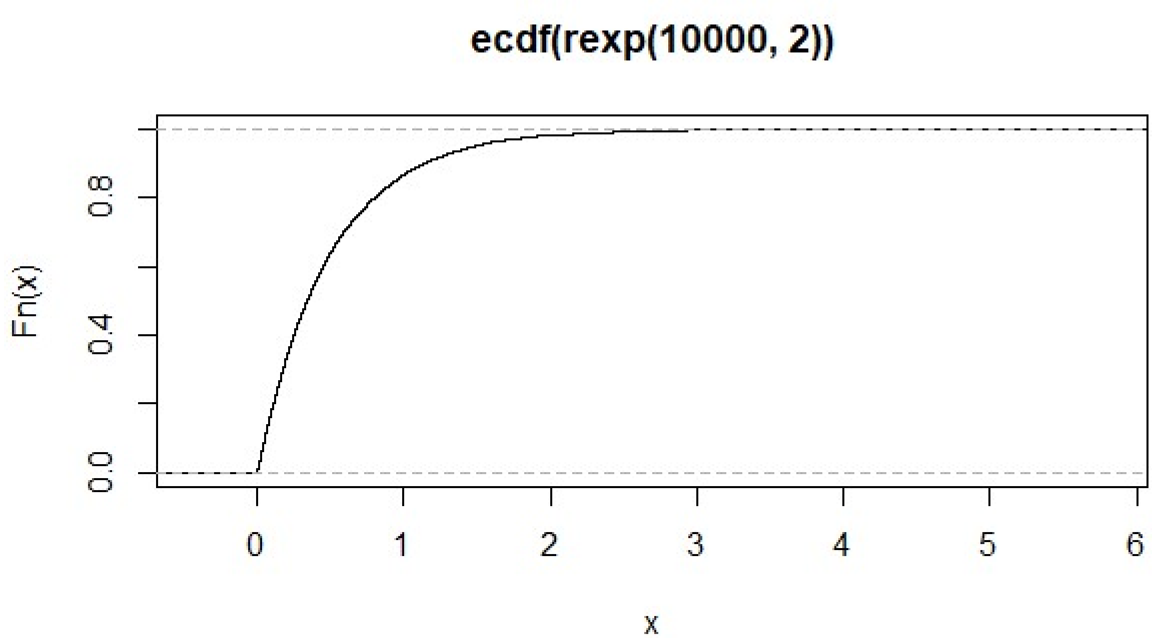


Exponential Distribution

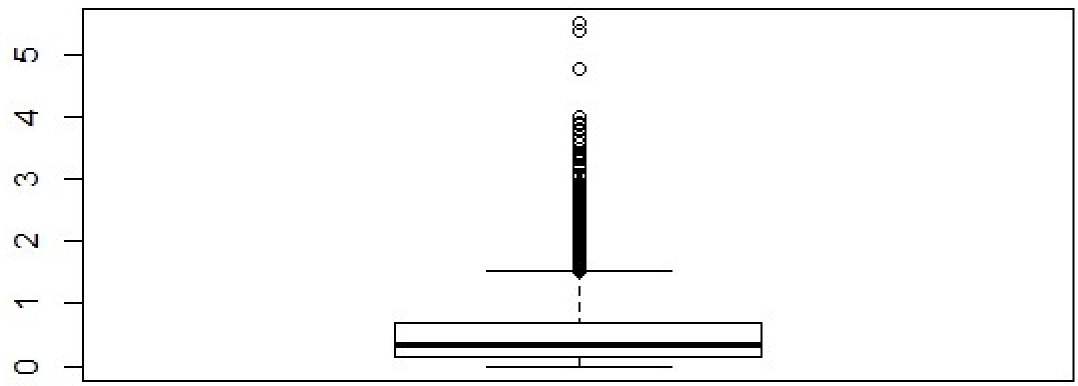
```
> hist(rexp(10000,2))
```




```
> plot(ecdf(rexp(10000,2)))
```



```
> boxplot(rexp(10000,2))
```



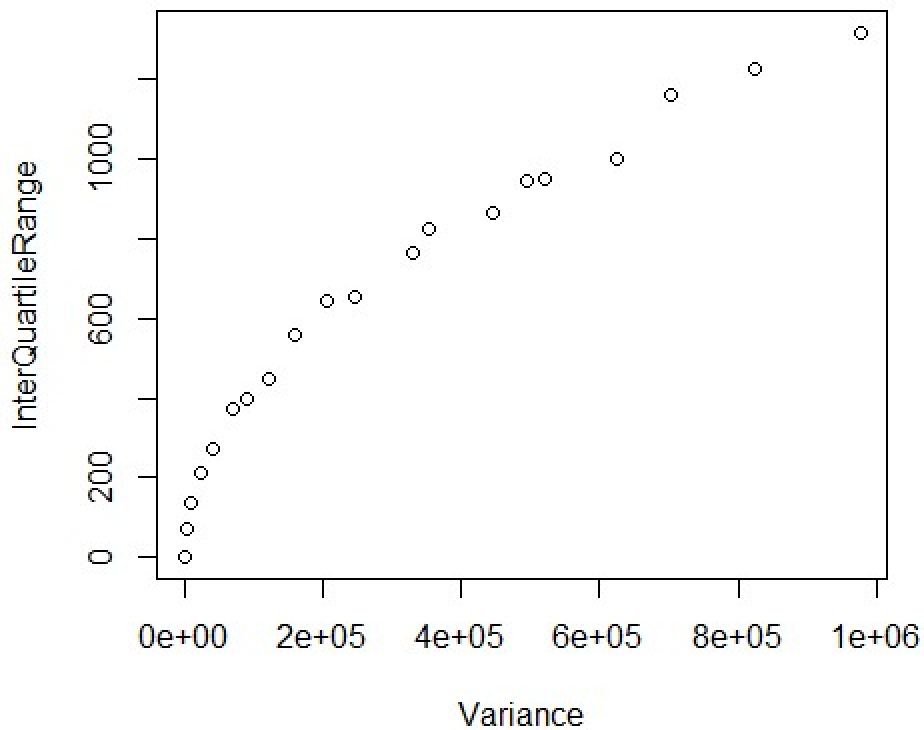
2. Consider one random vector of size 1000 for each normal distribution $\mathrm{N}(0,v)$ for $v \in 1 \ldots 1000$ with steps of 50. For each random vector, compute the empirical variance and the empirical IQR and plot those pairs in a graph.

   The objective here was to highlight the fact that both the IQRs and the variances can be used to measure the variability of a random variable.

Ans:

Code:

```
I=seq(1,1000,by=50);
Variance=NULL;
InterQuartileRange=NULL;
for(i in I)
{
y=rnorm(1000,0,i);
Variance=c(Variance,var(y));
InterQuartileRange=c(InterQuartileRange,IQR(y));
}
plot(Variance,InterQuartileRange);
```



This plot signifies the relationship between variance and the InterQuartile Range which is almost a linear relation in which both are proportionally variable.

2 $E[1/X]$ vs. $1/E[X]$

We observed in course/recitation that the discrepancy between $E[1/X]$ vs. $1/E[X]$ is a function of the variance of X. This is what we are going to illustrate.

Let us consider the family of uniform distributions in the interval $[100-v, 100+v]$ for $v > 0$

1. What are the mean/variance of the family?

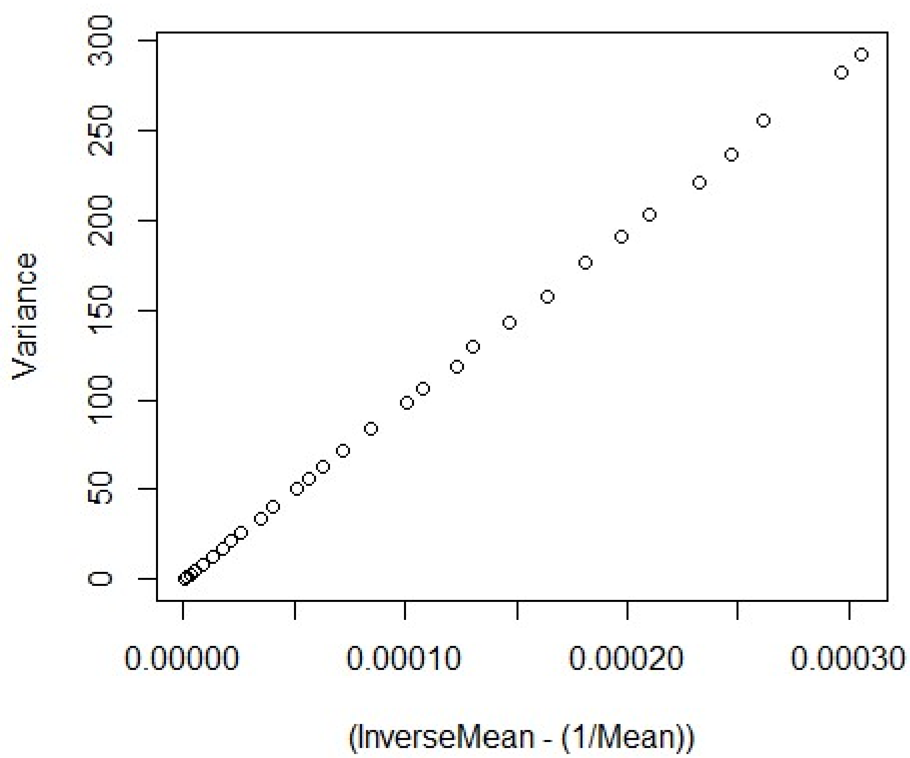
Ans:
A=100-v
B=100+v
Mean:
E[x]=(a+b)/2
= (100-v+100+v)/2
=100
-Variance=((b-a)^2)/12)
= (4(v^2))/12
= (v^2)/3

2. For each $v \in \{1,2,\ldots 30\}$, draw a random vector of size 1000, compute its empirical variance $v[X]$ as well as $E[1/X]$ (simply mean(1/x) in R). Plot the pairs $(E[1/X] - 1/E[X], v[X])$ and comment.

Ans:

Code:

```
V=seq(1,30,by=1);
Mean=NULL;
InverseMean=NULL;
Variance=NULL
for(v in V)
{
y=runif(1000,100-v,100+v);
Mean=c(Mean,mean(y));
InverseMean=c(InverseMean,mean(1/y));
Variance=c(Variance,var(y));
}
plot((InverseMean-(1/Mean)),Variance);
```



Comment:

$E[1/X] - 1/E[X], v[X]$ pairs of plot represent a straight line , we know that $E[1/X]$ is not equal to $1/E[X]$ and which implies there the variance which is how far the value is spread out from the mean of x is equal to the difference between $E[1/X] - 1/E[X]$ *and is a constant straight line increasing proportionally.*

3 Dependence vs. similar distribution

1. Draw a random variable X and a random variable Y (both of size 10000) from the same exponential distribution of parameter $\lambda = 2$. Plot the qqplot and the scatterplot of X and Y. The scatterplot is simply obtained by plot(X,Y). In the scatterplot, it might be useful to zoom in where the mass is. You can adjust the x-axis (resp. y-axis) between the 10-th and 90-th quantiles of X (resp. Y) with the command :

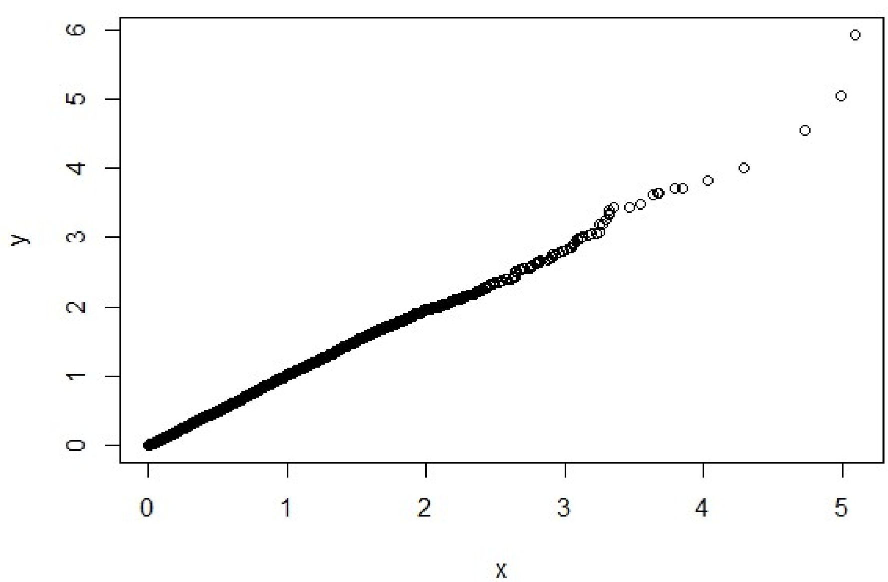
   plot(X,Y,xlim=cbind(quantile(X,0.1),quantile(X,0.9)),
   ylim=cbind(quantile(Y,0.1),quantile(Y,0.9)))

   Comment the results

   Ans:

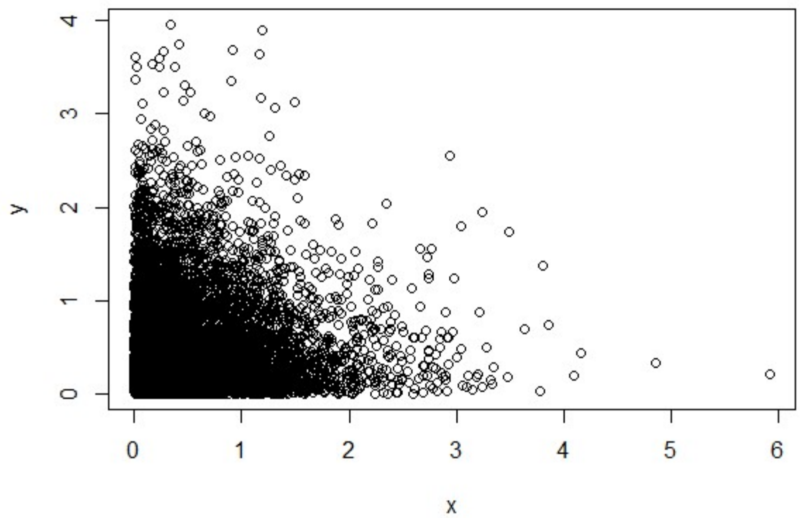
   Qqplot:

```
> x=rexp(10000,2);
> y=rexp(10000,2);
> qqplot(x,y)
```



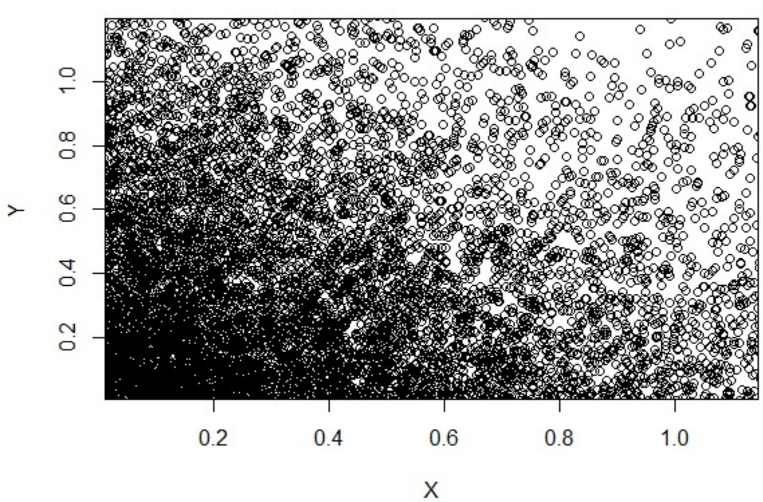
ScatterPlot:

```
> x=rexp(10000,2);
> y=rexp(10000,2);
> plot(x,y)
```



Scatterplot after adjusting the value of X between the 10th and 90th quantile of X with respect to Y:

```
X=rexp(10000,2);
Y=rexp(10000,2);
plot(X,Y,xlim=cbind(quantile(X,0.1),quantile(X,0.9)),ylim=cbind(quantile(Y,0.1),quantile(Y,0.9)));
```



Here QQ Plot signifies the specific relation and it portrays that is from the same distribution a straight line which is from exponential distribution for the quantiles.

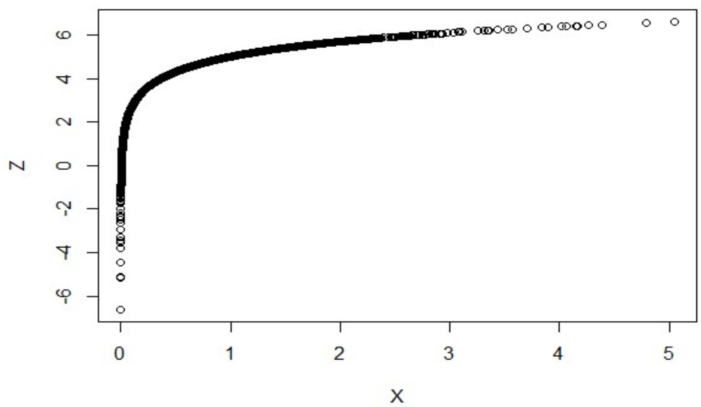
But In for a normal ScatterPlot , the temporal relation between the data is quite dependent for initial values and then becomes slightly independent.

But after adjusting the value of X between 10th and 90th Quantile of X with respect to Y signifies that the values are closer in the 10th Quantile range and then disperses as it near the 90th Quantile range.

Let $Z = log(X) + 5$. Plot the qqplot and the scatterplot of X and Z. Comment the results

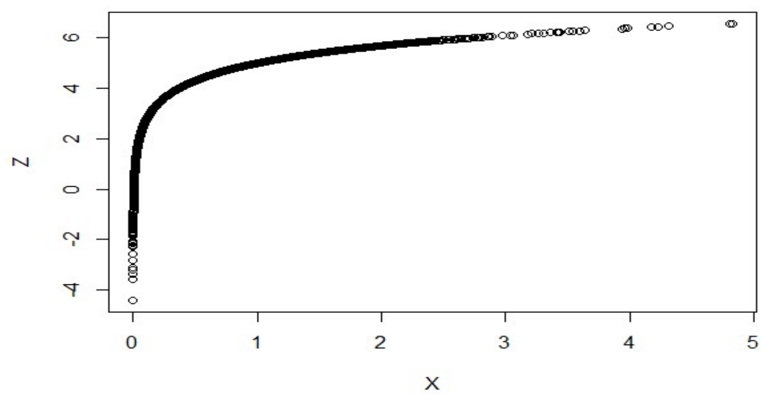
QQPlot:

```
X=rexp(10000,2);
Z=log(X)+5;
qqplot(X,Z);
```



ScatterPlot:

```
> X=rexp(10000,2);
> Z=log(X)+5;
> plot(X,Z)
```



QQ Plot signifies the value of the quantile doesn't come from a same distribution sine the value of the plot is not a straight line for X and Z. With Respect to the scatterplot they seem to be dependent values ,they seem to have a temporal relation which remains constant for some time showing a horizontal plot.

4 Loss events

The characterization of the loss process in the Internet is an highly debated topic in the research community (including ISPs).

In the present lab, we focus on losses observed by a sender machine S, involved in a BitTorrent session. In a BT session, a machine connects to and is contacted by other machines with whom it exchanges data over TCP connections.

The 2 files IP address _loss.txt contain the time intervals between consecutive losses for the traffic sent by S to two different hosts.

4.1 Data Cleaning

We are interested in consecutive loss periods in a TCP stream of packets. Loss periods are not exactly loss events: if ever two packets sent back-to-back are lost, we count only a single period of loss. Thus, a loss period is a period where one or more consecutive losses occur. Due to the way BitTorrent and TCP are working, data needs to be cleaned in two ways: (i) removing too small values and (ii) too large values. The reason why we need to get rid of too large and too small values are the following:

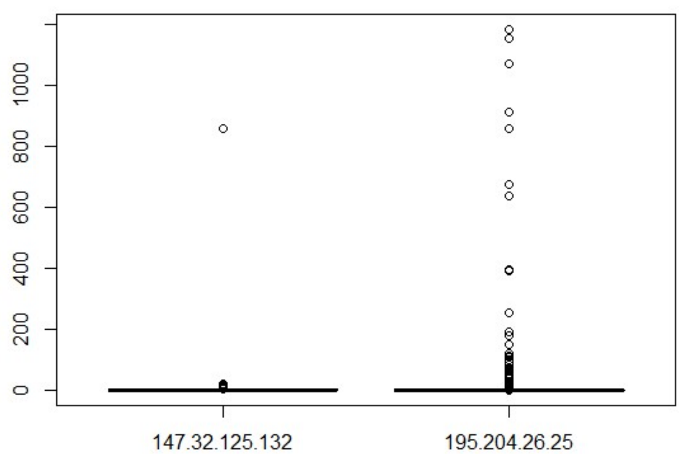
- Too small values might reveal a high degree of dependence among the packets that were lost. To put it differently, losses that occur in the same time window should be discarded (counted only once, as we concentrate on loss periods and not on loss events).

- Too large values might/can be due to the way BT is working and not to the absence of losses. Indeed, BT can generate "long" idle periods of time during which hosts do not exchange data. This means that a BT transfer looks like an ON/OFF FTP process, see Figure 1.

1. Clean the traces by using the 10-th and 90-th quantiles (i.e.,keep only the values in between those two quantiles) of the distributions and show how the boxplots of data evolves before and after cleaning for the two TCP transfers.

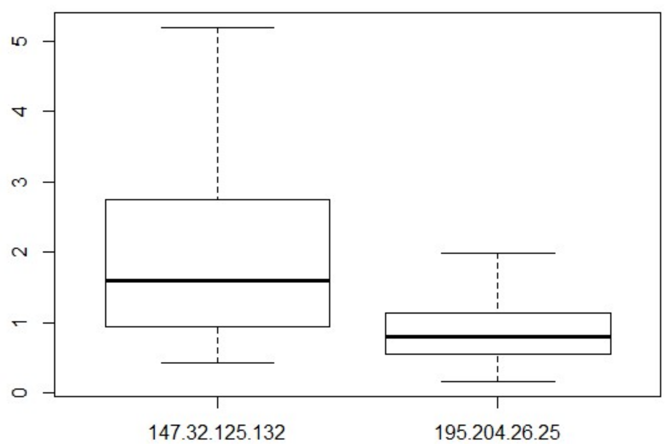
   Ans:
   For 147.32.125.132 & 195.204.26.25
   Before Cleaning:

```
> lossx<-read.table("147.32.125.132.loss.txt");
> lossx=lossx[,1];
> lossy<-read.table("195.204.26.25.loss.txt");
> lossy=lossy[,1];
> boxplot(lossx,lossy,names=c("147.32.125.132","195.204.26.25"));
```



After Cleaning:

```
> lossx<-read.table("147.32.125.132.loss.txt");
> lossx=lossx[,1];
> lossy<-read.table("195.204.26.25.loss.txt");
> lossy=lossy[,1];
> Ix=which(lossx>quantile(lossx,0.1)&lossx<quantile(lossx,0.9));
> Iy=which(lossy>quantile(lossy,0.1)&lossy<quantile(lossy,0.9));
> boxplot(lossx[Ix],lossy[Iy],names=c("147.32.125.132","195.204.26.25"));
```



4.2 Assessing the exponential hypothesis

Some researchers have observed during large measurement campaigns that inter-arrival times are exponentially distributed. We investigate here if it is a reasonable assumption for our two traces.

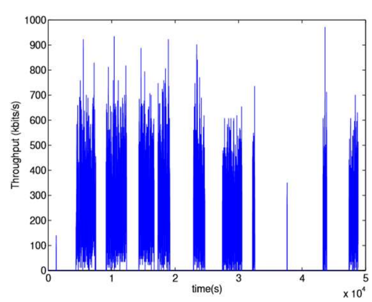


Figure 1: Example BT connection between two hosts

1. For each of the 2 connections (the cleaned versions obtained from the previous question), estimate the parameter of the exponential distribution that should model it.

For 147.32.125.132

Parameter 1/mean obtained by:

```
> lossx<-read.table("147.32.125.132.loss.txt");
> lossx=lossx[,1];
> Ix=which(lossx>quantile(lossx,0.1)&lossx<quantile(lossx,0.9));
> 1/mean(lossx[Ix]);
[1] 0.5085034
```

For 195.204.26.25

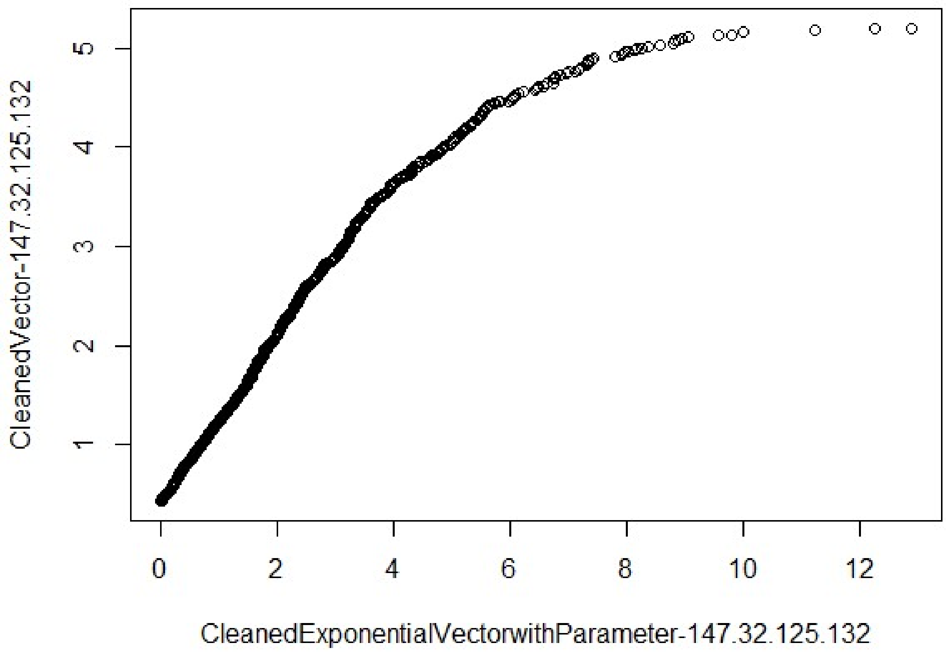
Parameter 1/mean obtained by:

```
> lossy<-read.table("195.204.26.25.loss.txt");
> lossy=lossy[,1];
> Iy=which(lossy>quantile(lossy,0.1)&lossy<quantile(lossy,0.9));
> 1/mean(lossy[Iy]);
[1] 1.150088
```

For each of the 2 connections, generate a random vector following the exponential distribution of size 1000, represent the qqplot of each vector and the corresponding trace. Comment.

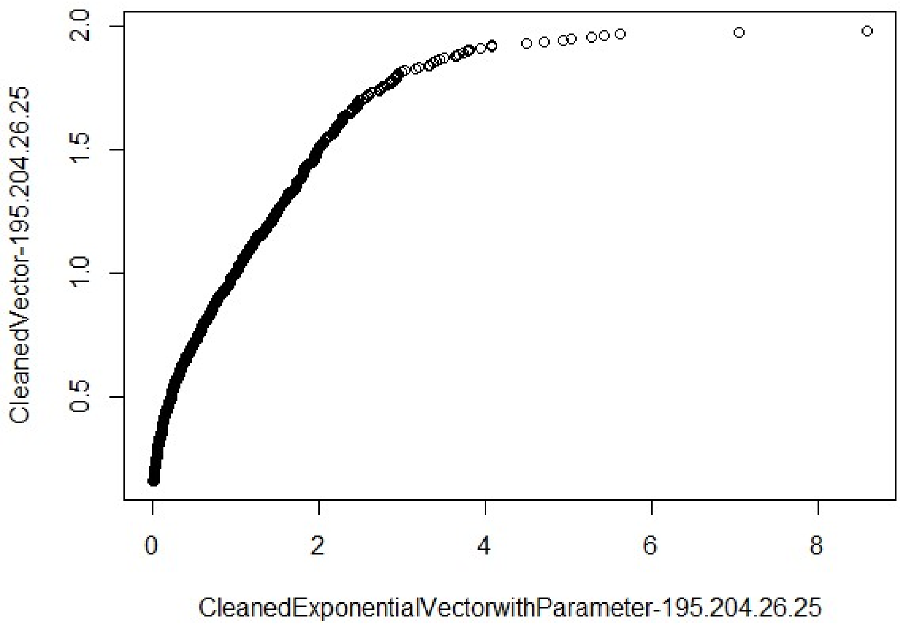
For 147.32.125.132

```
> lossx<-read.table("147.32.125.132.loss.txt");
> lossx=lossx[,1];
> Ix=which(lossx>quantile(lossx,0.1)&lossx<quantile(lossx,0.9));
> lambdalossx=1/mean(lossx[Ix]);
> xvector=rexp(1000,lambdalossx)
> qqplot(xvector,lossx[Ix],xlab = "CleanedExponentialVectorwithParameter-14
7.32.125.132",ylab = "CleanedVector-147.32.125.132");
```



For 195.204.26.25

```
> lossx<-read.table("195.204.26.25.loss.txt");
> lossx=lossx[,1];
> Ix=which(lossx>quantile(lossx,0.1)&lossx<quantile(lossx,0.9));
> lambdalossx=1/mean(lossx[Ix]);
> xvector=rexp(1000,lambdalossx)
> qqplot(xvector,lossx[Ix],xlab = "CleanedExponentialVectorwithParameter-19
5.204.26.25",ylab = "CleanedVector-195.204.26.25");
```



Comment: The QQPlot of both the files 147.32.125.132 and `195.204.26.25 represent that they are from the same distribution almost since they are in almost straight line`

5 Central limit theorem

We are going to illustrate here the central limit theorem, which can formulated as follows (source wikipedia.org):

Let $X_1,\ldots,X_n$ be a random sample of size n – that is, a sequence of independent and identically distributed random variables drawn from distributions of expected values given by μ and finite variances given by σ^2. Suppose we are interested in the sample average

$$S_n := \frac{X_1 + \cdots + X_n}{n}$$

of these random variables. By the law of large numbers, the sample averages converge in probability and almost surely to the expected value μ as $n \to \infty$. The classical central limit theorem describes the size and the distributional form of the stochastic fluctuations around the deterministic number μ during this convergence. More precisely, it states that as n gets larger, the distribution of the difference between the

sample average S_n and its limit μ, when multiplied by the factor $\sqrt{n}$ (that is $\sqrt{n}(S_n-\mu)$), approximates the normal distribution with mean 0 and variance σ^2. For large enough n, the distribution of S_n is close to the normal distribution with mean μ and $\frac{\sigma^2}{n}$ variance. The usefulness of the theorem is that the distri-

√_

bution of $n(S_n - \mu))$ approaches normality regardless of the shape of the distribution of the individual X_i's.

Lindeberg-Levy CLT: Suppose´ $\{Xi\}i \in$ N is a sequence of i.i.d. random variables with $E[Xi] = \mu$ and $V ar[Xi] = \sigma^2 < \infty$. Then as n approaches infinity, the random variables $\mathrm{p}(n)(Sn - \mu)$ converge in distribution to a normal $\mathrm{N}(0,\sigma^2)$:

$$\sqrt{n}\left(S_n - \mu\right) \xrightarrow{d} \mathcal{N}(0,\ \sigma^2).$$

where

$$S_n := (X_1 + \cdots + X_n)/n$$

To illustrate the theorem, let us generate 3 random vectors of size 1000 from different distributions :

- A uniform distribution between 0 and 1.
- A normal distribution N(0,10)
- A exponential distribution of parameter $\lambda = 2$

Questions:

1. Report in a table the empirical (resp. theoretical) mean and standard deviation for each randomvector (resp. random variable).

```
> unif=runif(1000,0,1);
> norm=rnorm(1000,0,sqrt(10))
> exp=rexp(1000,2);
> meanunif=mean(unif);
> sdunif=sd(unif);
> meannorm=mean(norm);
> sdnorm=sd(norm);
> meanexp=mean(exp);
> sdexp=sd(exp);
> print(c("Mean for Uniform Distribution=",meanunif));
[1] "Mean for Uniform Distribution=" "0.5030549839288"
> print(c("SD for Uniform Distribution=",sdunif));
[1] "SD for Uniform Distribution=" "0.283301630883049"
> print(c("Mean for Normal Distribution=",meannorm));
[1] "Mean for Normal Distribution=" "0.139470659865493"
> print(c("SD for Normal Distribution=",sdnorm));
[1] "SD for Normal Distribution=" "3.10932690759593"
> print(c("Mean for Exponential Distribution=",meanexp));
[1] "Mean for Exponential Distribution=" "0.492471688425875"

> print(c("SD for Exponential Distribution=",sdexp));
[1] "SD for Exponential Distribution=" "0.51"
```

Distribution	Empirical Mean	Empirical Standard Deviation
Uniform Distribution	0.5030549839288	0.283301630883049
Normal Distribution	0.139470659865493	3.10932690759593
Exponential Distribution	0.492471688425875	0.51

Distribution	Theoretical Mean	Theoretical-Standard Deviation
Uniform Distribution	=(a+b)/2=(1+0)/2=>0.5	(b-a)^2/12=(1-0)^12=>0.08
Normal Distribution	0	3.16
Exponential Distribution	0.5	0.5

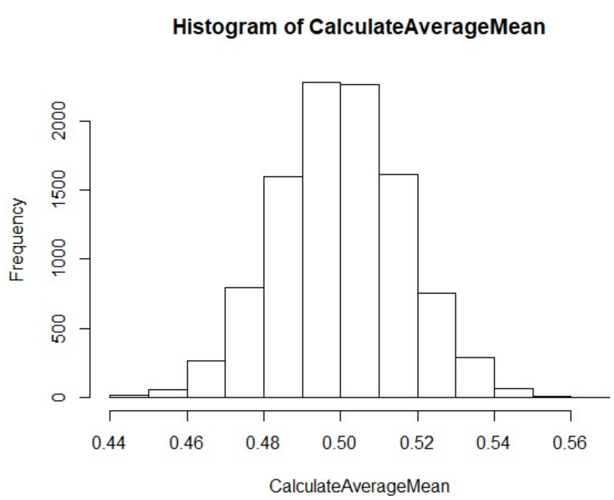
2. Prove that we are in the conditions of the theorem for each vector.

Uniform Distribution:

1) Distribution of Averages

We calculate the distribution of 10000 iteration averages with for uniform distribution with **n=300**,min=0,max=1 which is given by :

```
> iteration=10000;
> CalculateAverageMean=NULL;
> for(i in 1:iteration){
+     CalculateAverageMean=c(CalculateAverageMean,mean(runif(300,0,1)));
+ }
> hist(CalculateAverageMean);
```



This histogram almost represents a gaussian or normal distribution as per central Limit theorem

2) Sample Mean vs Theoretical Mean:

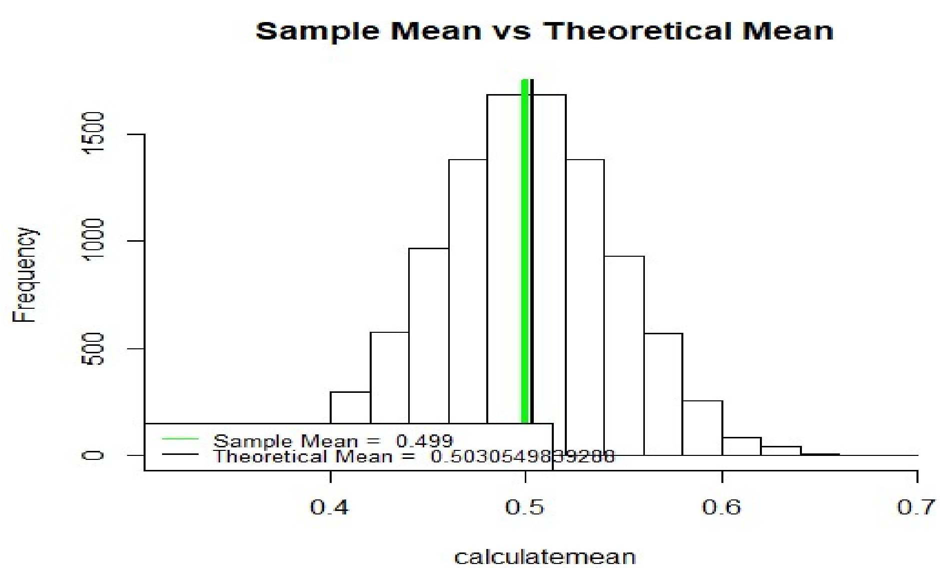
Theoretical Mean as already derived from previous question: 0.5030549839288

Sample Mean given by:

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(runif(300,0,1)));
+ }
> mean(calculatemean);
[1] 0.4998975
```

Sample Mean: 0.4998975 which is close to the theoretical mean derived.

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(runif(300,0,1)));
+ }
> samplemean=mean(calculatemean);
> theoreticalmean=0.5030549839288;
> h<-hist(calculatemean,main="Sample Mean vs Theoretical Mean");
> abline(v = samplemean, lwd="4", col="green");
> abline(v = theoreticalmean, lwd="2", col="black");
> legend("bottomleft", legend=c(paste("Sample Mean = ", round(samplemean,3)
), paste("Theoretical Mean = ", theoreticalmean)), col=c("green", "black"),
 lty=1, cex=0.8);
```



Here Sample Mean represented in Green is coinciding with the theoretical me
an.

3) **Sample SD vs Theoretical SD**

Theoretical SD is given by the previous problem which is 0.2833

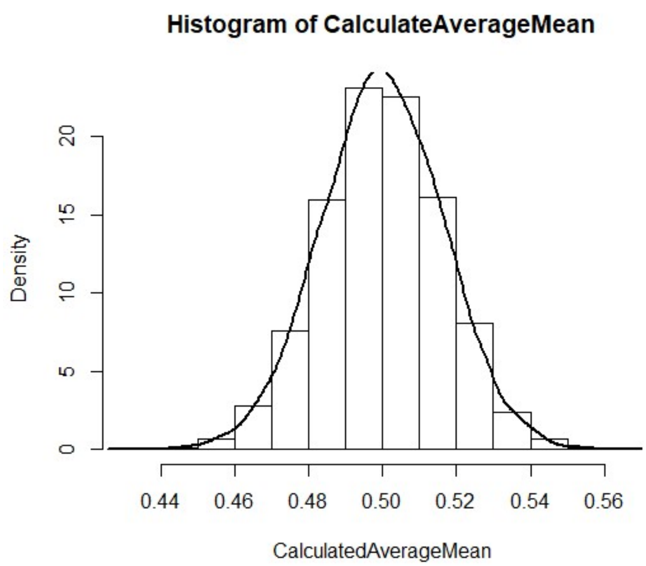
Sample SD given by:

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,sd(runif(300,0,1)));
+ }
> mean(calculatemean);
[1] 0.2885564
```

Theoretical SD is almost equal to Sample SD values are close to each other.

4) **Distribution:**

Finally we can prove that the uniform distribution follows the central limit theorem if the sample distribution follows the normal distribution. We use a density histogram to prove this.

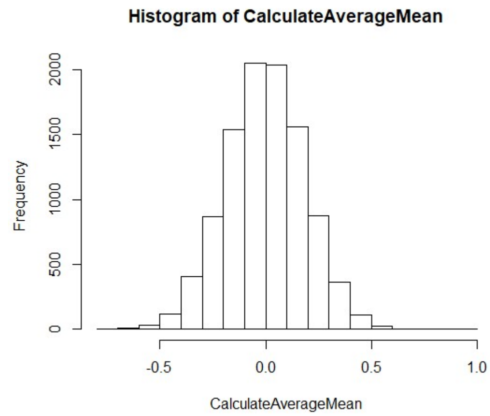


Normal Distribution:

1) **Distribution of Averages:**

We calculate the distribution of 10000 iteration averages with for uniform distribution with **n=300**,mean=0,SD=sqrt(10) which is given by :

```
> iteration=10000;
> CalculateAverageMean=NULL;
> for(i in 1:iteration){
+     CalculateAverageMean=c(CalculateAverageMean,mean(rnorm(300,0,sqrt(10))));
+ }
> hist(CalculateAverageMean);
```



This histogram almost represents a gaussian or normal distribution as per central Limit theorem

2) Sample Mean vs Theoretical Mean

Theoretical Mean as already derived from previous question: 0

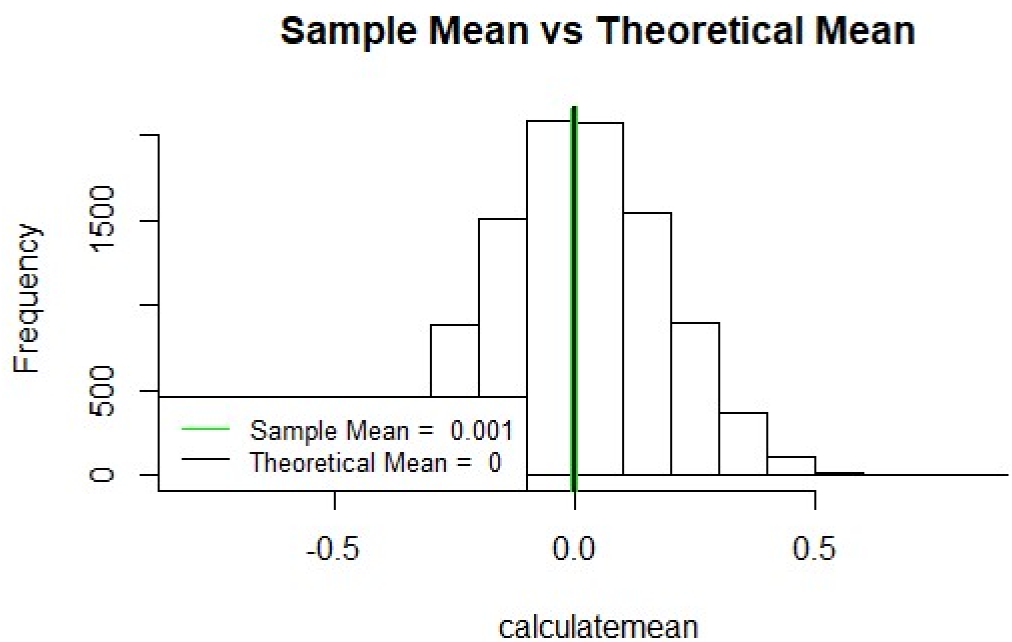
Sample Mean given by:

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(rnorm(300,0,sqrt(10))));
+ }
> mean(calculatemean);
[1] 0.001542606
```

Sample Mean: 0.0015 which is close to the theoretical mean derived.

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(rnorm(300,0,sqrt(10))));
+ }
> samplemean=mean(calculatemean);
> theoreticalmean=0;
> h<-hist(calculatemean,main="Sample Mean vs Theoretical Mean");
> abline(v = samplemean, lwd="4", col="green");
> abline(v = theoreticalmean, lwd="2", col="black");
> legend("bottomleft", legend=c(paste("Sample Mean = ", round(samplemean,3)), paste("Th
 Mean = ", theoreticalmean)), col=c("green", "black"), lty=1, cex=0.8);


>
```



Here Sample Mean represented in Green is coinciding with the theoretical mean.

3) Sample SD vs Theoretical SD

Theoretical SD is given by the previous problem which is 3.10

Sample SD given by:

```
> iteration=10000;
> calculatesd=NULL;
> for(i in 1:iteration){
+     calculatesd=c(calculatesd,sd(rnorm(300,0,sqrt(10))));
+ }
> mean(calculatesd);
[1] 3.159624
```

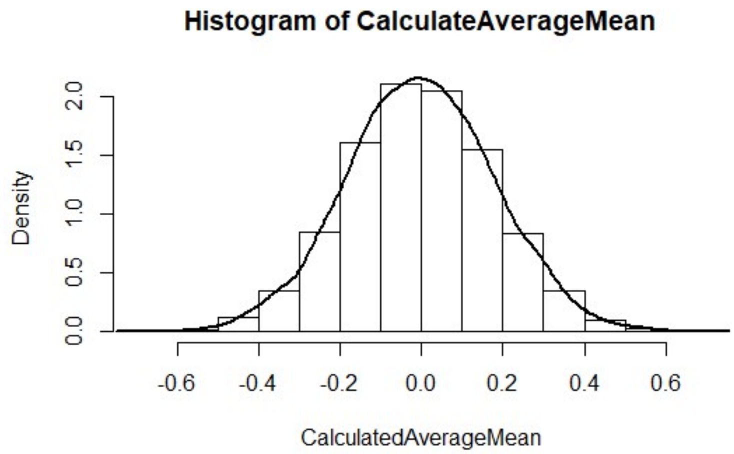
Theoretical SD is almost equal to Sample SD values are close to each other.

4) Distribution

Finally we can prove that the uniform distribution follows the central limit theorem if the sample distribution follows the normal distribution. We use a density histogram to prove this.

```
> iteration=10000;
> CalculateAverageMean=NULL;
> for(i in 1:iteration){
+     CalculateAverageMean=c(CalculateAverageMean,mean(rnorm(300,0,sqrt(10))));
+ }
> hist(CalculateAverageMean,prob="TRUE",xlab="CalculatedAverageMean",ylab="Density");
> lines(density(CalculateAverageMean),col="black",lwd=2);


>
```



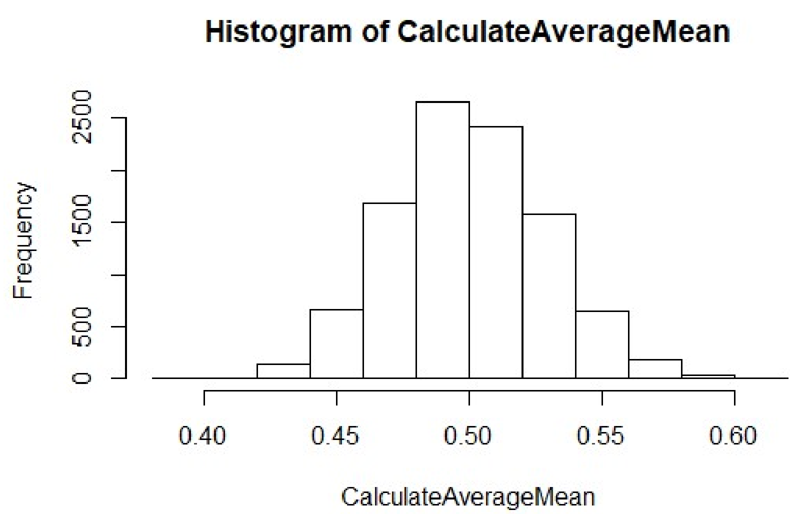
Exponential Distribution:

1) **Distribution of Averages**

   We calculate the distribution of 10000 iteration averages with for uniform distribution with **n=300**,lamda=2 which is given by :

```
> iteration=10000;
> CalculateAverageMean=NULL;
> for(i in 1:iteration){
+     CalculateAverageMean=c(CalculateAverageMean,mean(rexp(300,2)));
+ }
> hist(CalculateAverageMean);


>
```



This histogram almost represents a gaussian or normal distribution as per central Limit theorem

2) Sample Mean vs Theoretical Mean

Theoretical Mean as already derived from lambda which is 2: 0.5

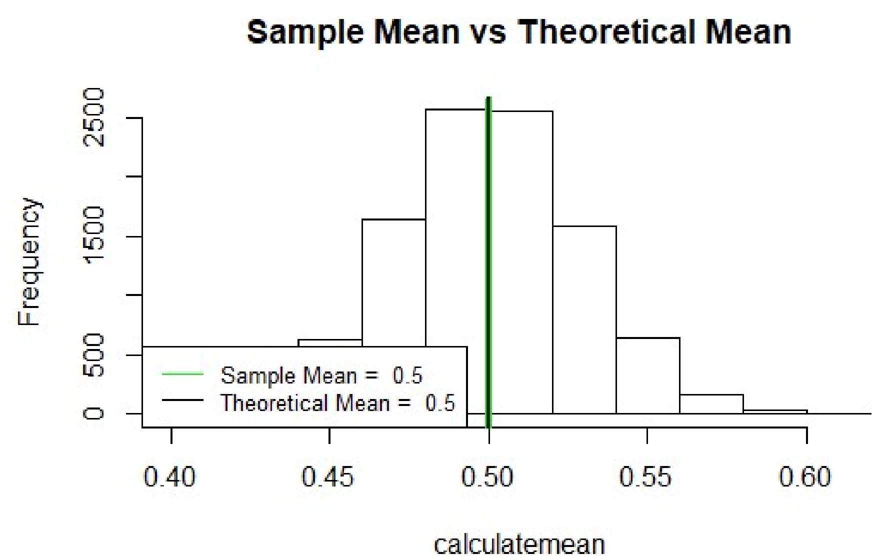
Sample Mean given by:

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(rexp(300,2)));
+ }
> mean(calculatemean);
[1] 0.500022
```

Sample Mean: 0.5000 which is close to the theoretical mean derived.

```
> iteration=10000;
> calculatemean=NULL;
> for(i in 1:iteration){
+     calculatemean=c(calculatemean,mean(rexp(300,2)));
+ }
> samplemean=mean(calculatemean);
> theoreticalmean=0.5;
> h<-hist(calculatemean,main="Sample Mean vs Theoretical Mean");
> abline(v = samplemean, lwd="4", col="green");
> abline(v = theoreticalmean, lwd="2", col="black");
> legend("bottomleft", legend=c(paste("Sample Mean = ", round(samplemean,3)), paste("Th
 Mean = ", theoreticalmean)), col=c("green", "black"), lty=1, cex=0.8);


>
```



Here Sample Mean represented in Green is coinciding with the theoretical mean.

3) **Sample SD vs Theoretical SD**

```
> iteration=10000;
> calculatesd=NULL;
> for(i in 1:iteration){
+     calculatesd=c(calculatesd,sd(rexp(300,2)));
+ }
> mean(calculatesd);
[1] 0.4984496
```

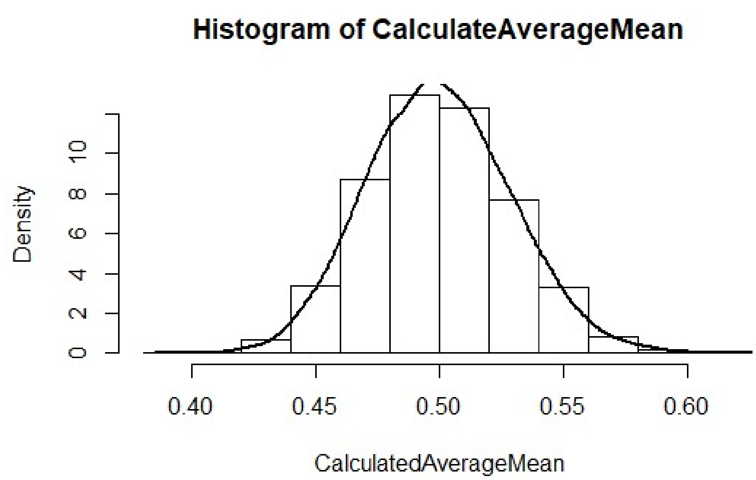
Theoretical SD:
Sqrt(1/lamda^2)=sqrt(1/2^2)=0.5

Theoretical SD is almost equal to Sample SD calculated

4) **Distribution**

Finally we can prove that the uniform distribution follows the central limit theorem if the sample distribution follows the normal distribution. We use a density histogram to prove this.

```
> iteration=10000;
> CalculateAverageMean=NULL;
> for(i in 1:iteration){
+     CalculateAverageMean=c(CalculateAverageMean,mean(rexp(300,2)));
+ }
> hist(CalculateAverageMean,prob="TRUE",xlab="CalculatedAverageMean",ylab="Density");
> lines(density(CalculateAverageMean),col="black",lwd=2);
```

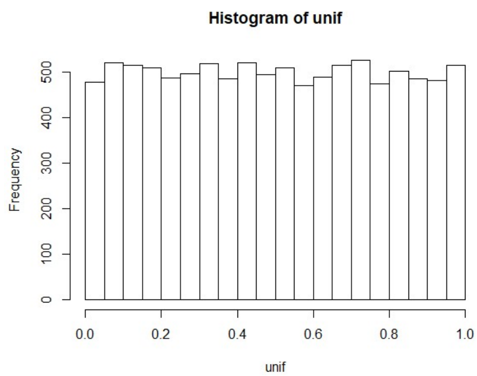
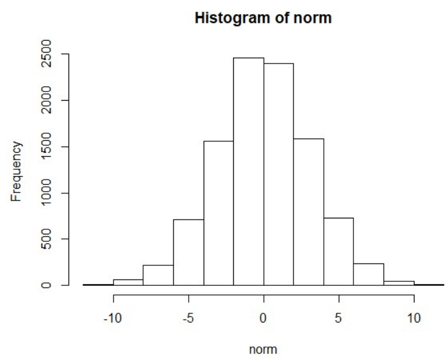
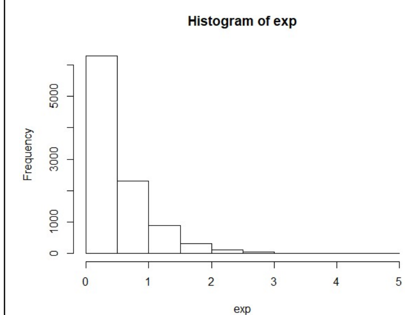
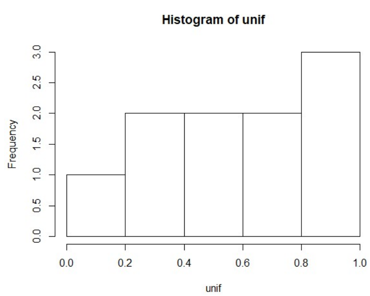
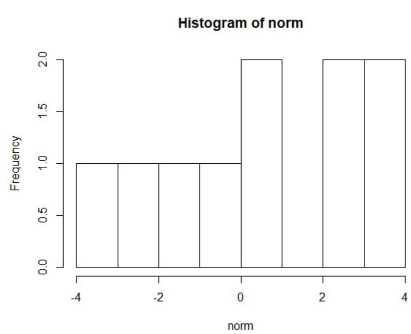
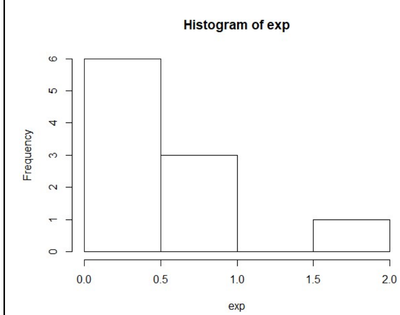


3. Towards which distribution should $\sqrt{(n)}(Sn - \mu)$ should converge in each case.

   Since it follows a central Limit theorem the $\sqrt{(n)}(Sn - \mu)$ should converge in each case to a normal distribution

4. Represent in a table with three columns (one for each original distribution) and two rows corresponding to:

   a) the histogram of the original distributions b) S_{10}

<table>
<tr><td>Uniform Distribution</td><td>Normal Distribution</td><td>Exponential Distribution</td></tr>
<tr>
<td>Original distribution
> unif=runif(10000,0,1);
> hist(unif);

</td>
<td>Original distribution
>norm=rnorm(10000,0,sqrt(10))
> hist(norm);

</td>
<td>Original distribution
> exp=rexp(10000,2);
> hist(exp);

</td>
</tr>
<tr>
<td>S_{10}
> unif=runif(10,0,1);
> hist(unif);

</td>
<td>S_{10}
> norm=rnorm(10,0,sqrt(10))
> hist(norm);

</td>
<td>S_{10}
> exp=rexp(10,2);
> hist(exp);

</td>
</tr>
</table>

5. Report also the empirical mean and standard deviation for S_{10} for all cases.

```
> unif=runif(10,0,1);
> norm=rnorm(10,0,sqrt(10))
> exp=rexp(10,2);
> meanunif=mean(unif);
> sdunif=sd(unif);
> meannorm=mean(norm);
> sdnorm=sd(norm);
> meanexp=mean(exp);
> sdexp=sd(exp);
> print(c("Mean for Uniform Distribution=",meanunif));
[1] "Mean for Uniform Distribution=" "0.389881837414578"
> print(c("SD for Uniform Distribution=",sdunif));
[1] "SD for Uniform Distribution=" "0.31149615833305"
> print(c("Mean for Normal Distribution=",meannorm));
[1] "Mean for Normal Distribution=" "1.20510892010718"
> print(c("SD for Normal Distribution=",sdnorm));
[1] "SD for Normal Distribution=" "3.31911304976232"
> print(c("Mean for Exponential Distribution=",meanexp));
[1] "Mean for Exponential Distribution=" "0.453684924261696"

> print(c("SD for Exponential Distribution=",sdexp));
[1] "SD for Exponential Distribution=" "0.361384264948565"
```

Distribution	**Mean**	**Standard Deviation**
Uniform Distribution	0.389881837414578	0.31149615833305
Normal Distribution	1.20510892010718	3.31911304976232
Exponential Distribution	0.453684924261696	0.361384264948565