Probability Assignment 5

Probability and statitics (National University of Computer and Emerging Sciences)

National University of Computer and Emerging Sciences

Assignment 5

MT-206 Probability and Statistics

Q1	Take an example of Binomial Distribution, hypergeometric Distribution, Poisson, Geometric, Negative Binomial Distribution and Normal distribution and Find.
i.	Construct the complete probability distribution.
ii.	Construct the Cumulative probability distribution.
iii.	Find the prob. of different events.
iv.	Draw probability mass function
v.	Cumulative probability function.
vi.	Discuss the different behavior of the line graph when n will be changed (only in case of disc. Prob. Dis.).
vii.	Also generate 50 random numbers from each distribution.
viii.	Find the First Second and Third Quartiles of each distribution under study.
ix.	Calculate the 50 th , 70 th , 95 th percentiles.
x.	Calculate the second decile and D6
Note	Clearly specified the complete procedure to execute each part in R

BINOMIAL DISTRIBUTION

Example 5.1: The probability that a certain kind of component will survive a shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

i) Construct the complete probability distribution.

R code:

```
n <- 4 ;  p<- ¾  ; x <- 0:n
y=dbinom(x,n,p)
tab=cbind(x,y)
tab
```

```
     x          y
[1,] 0 0.00390625
[2,] 1 0.04687500
[3,] 2 0.21093750
[4,] 3 0.42187500
[5,] 4 0.31640625
```

ii) Construct the Cumulative probability distribution.

R Code

```
n <- 4; p<- 0..75; x <- 0:n
y=pbinom(x,n,p)
tab= cbind(x,y)
tab
```

```
     x          y
[1,] 0 0.00390625
[2,] 1 0.05078125
[3,] 2 0.26171875
[4,] 3 0.68359375
[5,] 4 1.00000000
```

iii) Find the prob. of different events.

a) Let A be the event that exactly 2 of the next 4 components tested survived

R code
x<-2; n<-4; p<- 0.75;
y=pbinom(x,n,p)
y

```
> x<-2; n<-4; p<-0.75;
> y=dbinom(x,n,p)
> y
[1] 0.2109375
```

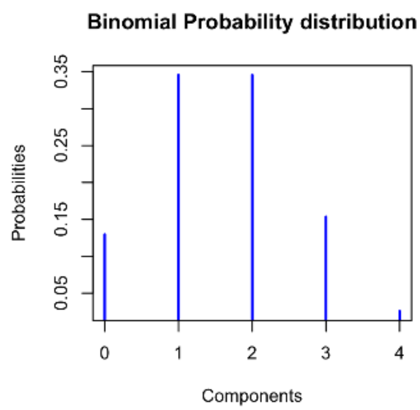
b) Let A be the event that exactly 3 of the next 4 components tested survived

R code
x<-3; n<-4; p<- 0.75;
y=pbinom(x,n,p)
y

```
> x<-3; n<-4; p<-0.75;
> y=dbinom(x,n,p)
> y
[1] 0.421875
```

iv) Draw probability mass function

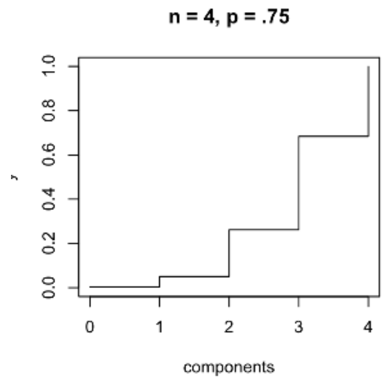
R Code
n <- 4; p<- 0.4; x <- 0:n
y=dbinom(x,n,p)
plot(x,y,type="l",xlab="Components", ylab="Probabilities",main="Binomial Probability distribution",lwd=2,col="red")



v) Cumulative probability function.

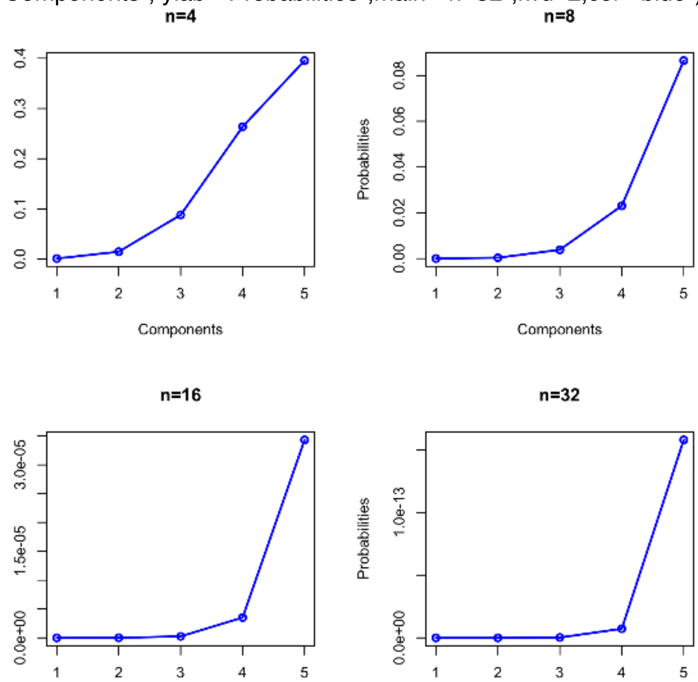
R Code

```
n <- 4; p<- 0.75; x <- 0:n
y=pbinom(x,n,p)
plot(x, y, xlab = "components",
ylab = "y", ylim = c(0, 1), type = "s", main = "n = 4, p = .75")
```



vi) Discuss the different behavior of the line graph when n will be changed

R code:
```
n <- 4;n2<- 8; n3<-16; n4 <-32; p<- 0.75;
x <- 0:n;
a=dbinom(x,n1,p)
b=dbinom(x,n2,p)
c=dbinom(x,n3,p)
d=dbinom(x,n4,p)
par(mfrow=c(2,2))
plot(a,type="l",xlab="Components", ylab="Probabilities",main="n=4",lwd=2,col="blue")
plot(b,type="l",xlab="Components", ylab="Probabilities",main="n=8",lwd=2,col="blue")
plot(c,type="l",xlab="Components", ylab="Probabilities",main="n=16",lwd=2,col="blue")
plot(d,type="l",xlab="Components", ylab="Probabilities",main="n=32",lwd=2,col="blue")
```



Observation: It is observed that when the number of components increases, the graph is inclined towards the right.

vii) Generate 50 random numbers from each distribution.

R Code
```
n <- 4; p<- 0.75;
rbinom(50,n,p)
```

```
 [1] 4 4 3 3 3 1 4 3 4 4 3 3 1 4 2 2 3 4 3 3 2 4 4 2 4 3 3 2 2 4 4 3 4 3 4 4 2
[38] 4 3 3 2 3 2 3 4 3 3 2 3 4
```

viii.) Find the First Second and Third Quartiles of each distribution under study.

R code:
```
n <- 4; p<- 0.75;
vec=c(.25,.5,.75)
qbinom(vec,n,p)
Output:
[1] 2 3 4
```

ix.) Calculate the 50th , 70th , 95 th percentiles.

R Code:
```
n <- 4; p<- 0.75;
vec=c(.5,.7,.95)
qbinom(vec,n,p)
Output:
[1] 3 4 4
```

x.) Calculate the second decile and D6

R Code
```
n <- 4; p<- 0.75;
vec=c(.2,.6)
qbinom(vec,n,p)
Output
[1] 2 3
```

HYPERGEOMETRIC DISTRIBUTION

Example 5.8: A particular part that is used as an injection device is sold in lots of 10. The producer deems a lot acceptable if no more than one defective is in the lot. A sampling plan involves random sampling and testing 3 of the parts out of 10. If none of the 3 is defective, the lot is accepted. Comment on the utility of this plan.

i. Construct the complete probability distribution.

R-code:

```
n <- 3; x<-0:n;  k<-2 ; N<-10 ;N1=N-k
y=dhyper(x,k,N1,n)
tab=cbind(x,y)
tab
```

Output --->

ii. Construct the Cumulative probability distribution.

```
n<-3; x<-0:n; k<-2; N<-10; N1= N-k
y=phyper(x,k,N1,n)
tab=cbind(x,y)
tab
```

Output --->

iii. Find the prob. of different events.

“If none of the 3 is defective, the lot is accepted. Comment on the utility of this plan.”

R-code:

```
n <- 3; x<-0;  k<-2 ; N<-10 ;N1=N-k
y=dhyper(x,k,N1,n)
tab=cbind(x,y)
tab
```

Output --->

iv. Draw probability mass function

```
n<-3; x<-0:n; k<-2; N<-10; N1= N-k
y=dhyper(x,k,N1,n)
par(mfrow=c(2,2))
plot(x,y,type="l",xlab="Outcomes", ylab="Probabilities",main="HyperGeometric Probability
distribution",lwd=2,col="red")
```

v. Cumulative probability function.

`n<-3; x<-0:n; k<-2; N<-10; N1= N-k`
`y=dhyper(x,k,N1,n)`
`par(mfrow=c(2,2))`
`plot(x,y,type="s",xlab="Outcomes", ylab="Probabilities",main="HyperGeometric cdf",lwd=2,col="red")`

vi. Discuss the different behavior of the line graph when n will be changed.

`n1<- 5; n2<-3; n3<-6; n4<-9;`
`x1<-0:n1; x2<-0:n2; x3<-0:n3; x4<-0:n4;`
`k<-2; N<-10; N1= N-k`
`a=dhyper(x1,k,N1,n1)`
`b=dhyper(x2,k,N1,n2)`
`c=dhyper(x3,k,N1,n3)`
`d=dhyper(x4,k,N1,n4)`
`par(mfrow=c(2,2))`
`plot(a,type="l",xlab="Outcomes", ylab="Probabilities",main="n=5",lwd=2,col="red")`
`plot(b,type="l",xlab="Outcomes", ylab="Probabilities",main="n=3",lwd=2,col="red")`
`plot(c,type="l",xlab="Outcomes", ylab="Probabilities",main="n=6",lwd=2,col="red")`
`plot(d,type="l",xlab="Outcomes", ylab="Probabilities",main="n=9",lwd=2,col="red")`
 	
We observe that the bigger the n is, the higher the peak is

vii. Also generate 50 random numbers from each distribution.

`n<-3; x<-50; k<-2; N<-10; N1= N-k`
`rhyper(x,k,N1,n)`
 	
Output : [1] 0 0 0 0 1 1 1 1 1 1 0 1 1 1 0 1 1 1 1 0 2 1 1 1 1 0 1 1 0 0 1 0 1 1 1 0 1
[38] 0 0 0 0 0 1 0 0 1 0 0 0 0

viii. Find the First Second and Third Quartiles of each distribution under study.
n<-3; x<-50; k<-2; N<-10; N1= N-k vec=c(.25,.5,.75) qhyper(vec,k,N1,n) Output ---> [1] 0 1 1

ix. Calculate the 50 th , 70 th , 95 th percentiles.
n<-3; x<-50; k<-2; N<-10; N1= N-k vec=c(.5,.7,.95) qhyper(vec,k,N1,n) Output ---> [1] 1 1 2

x. Calculate the second decile and D6
n<-3; x<-50; k<-2; N<-10; N1= N-k vec=c(.2,.6) qhyper(vec,k,N1,n) Output ---> [1] 0 1

POISSON DISTRIBUTION

Example : In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

(a) What is the probability that in any given period of 400 days there will be an accident on one day?
(b) What is the probability that there are at most three days with an accident?

i. Construct the complete probability distribution.

```
x=0:20;n<-400; p<-0.005; lambdha=n*p;
y=dpois(x,lambdha)
tab=cbind(x,y)
tab
```

ii. Construct the Cumulative probability distribution.

```
x=0:20;n<-400; p<-0.005; lambdha=n*p;
y=ppois(x,lambdha)
tab=cbind(x,y)
tab
```

iii. Find the prob. of different events.

a) P(X = 1)

R Code

```
x=1;n<-400; p<-0.005; la=n*p;
dpois(x,la)
```

Output ---> [1] 0.2706706

b) P(X ≤ 3)

R Code

```
x=3;n<-400; p<-0.005; la=n*p;
ppois(x,la)
```

Output --->[1] 0.8571235

iv. Draw probability mass function

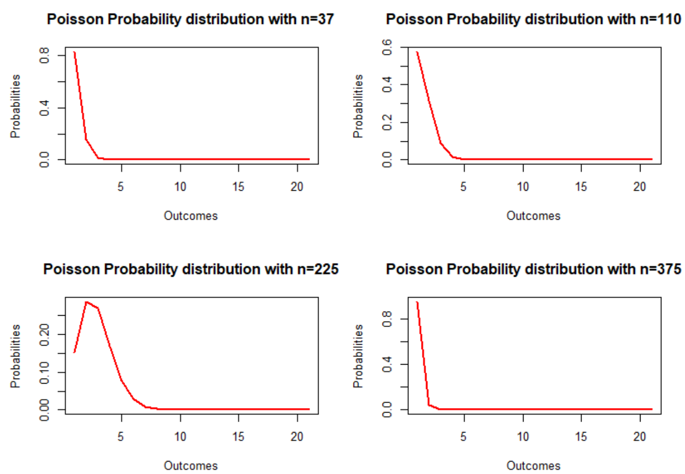
```
x=0:20;n<-400; p<-0.005; la=n*p;
y=dpois(x,la)
plot(x,y,type="l",xlab="Outcomes", ylab="Probabilities",main="Poisson Probability distribution",lwd=2,col="red"
```

v. Cumulative probability function.

```
x=0:20;n<-400; p<-0.005; la=n*p;
y=ppois(x,la)
plot(x,y,type="l",xlab="Outcomes", ylab="Cumulative Probabilities",main="Poisson Cumulative
Probability distribution",lwd=2,col="red")
```

vi. Discuss the different behavior of the line graph when n will be changed.

```
x=0:20; p<-0.005;
n1<-37 ; n2<-110; n3<-225; n3<-375;
la1=n1*p; la2=n2*p; la3=n3*p; la4=n4*p;
a=dpois(x,la1)
b=dpois(x,la2)
c=dpois(x,la3)
d=dpois(x,la4)
par(mfrow=c(2,2))
plot(a,type="l",xlab="Outcomes", ylab="Probabilities",main="n=37",lwd=2,col="red")
plot(b,type="l",xlab="Outcomes", ylab="Probabilities",main="n=110",lwd=2,col="red")
plot(c,type="l",xlab="Outcomes", ylab="Probabilities",main="n=225",lwd=2,col="red")
plot(d,type="l",xlab="Outcomes", ylab="Probabilities",main="n=375",lwd=2,col="red")
```



With a larger n, the probability tends to a constant value

vii. Also generate 50 random numbers from each distribution.

```
x=50; p<-0.005;
n<-400; la=n*p;
rpois(x,la)
```

Output ---> [1] 1 2 2 3 4 0 2 2 5 2 1 1 3 4 2 1 1 3 2 2 1 3 2 1 4 3 4 3 0 2 2 1 3 2 0 4 2 0
[39] 3 0 3 3 2 1 0 1 3 4 1 2

viii. Find the First Second and Third Quartiles of each distribution under study.

```
x=c(.25,.5,.75); p<-0.005;
n<-400; la=n*p;
qpois(x,la)
```

Output ---> [1] 1 2 3

ix. Calculate the 50 th , 70 th , 95 th percentiles.

```
x=c(.5,.7,.95); p<-0.005;
n<-400; la=n*p;
qpois(x,la)
```

Output ---> [1] 2 3 5

x. Calculate the second decile and D6

```
x=c(.2,.6); p<-0.005;
n<-400; la=n*p;
qpois(x,la)
```

Output ---> [1] 1 2

GEOMETRIC DISTRIBUTION

Example : For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?

i. Construct the complete probability distribution.

```
x=0:20;p=0.01
y=dgeom(x,p)
tab=cbind(x,y)
tab
```

ii. Construct the Cumulative probability distribution.

```
x=0:20;p=0.01
y=pgeom(x,p)
tab=cbind(x,y)
tab
```

iii. Find the prob. of different events.

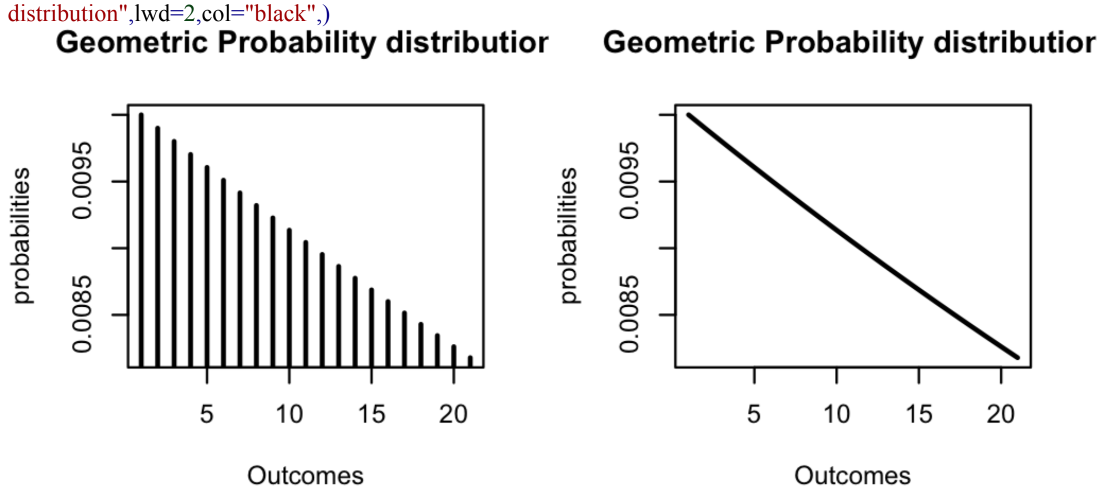
R Code

```
x=5;p=0.01
pgeom(x,p)
```

Output ---> [1] 0.0095099

iv. Draw probability mass function

```
x=0:20;p=0.01
y=dgeom(x,p)
par(mfrow=c(3,3))
plot(y,type="h",xlab="Outcomes", ylab="probabilities",main="Geometric Probability
distribution",lwd=2,col="black",pch=22)
plot(y,type="l",xlab="Outcomes", ylab="probabilities",main="Geometric Probability
distribution",lwd=2,col="black",)
```

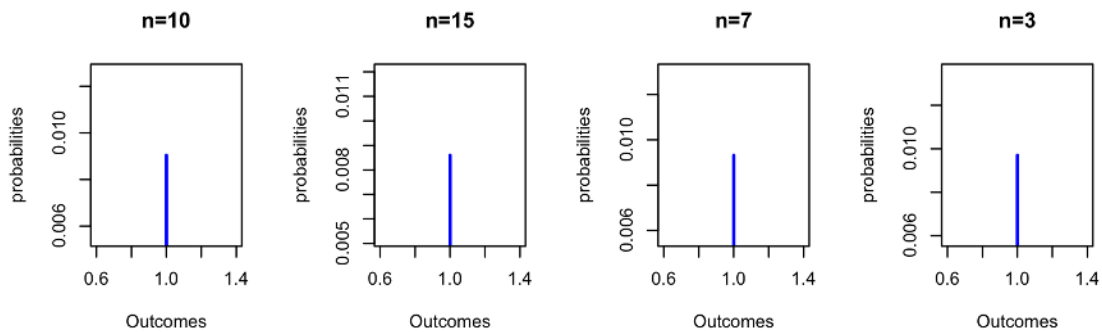


v. Cumulative probability function.

```
x=0:20;p=0.01
y=pgeom(x,p)
par(mfrow=c(3,3))
plot(y,type="s",xlab="Outcomes", ylab="probabilities",main="Cumulative Probability
distribution",lwd=2,col="red")
```

vi. Discuss the different behavior of the line graph when n will be changed.

```
x1=10; x2=15; x3=7 ; x4=3 ;p=0.01
a=dgeom(x1,p)
b=dgeom(x2,p)
c=dgeom(x3,p)
d=dgeom(x4,p)
par(mfrow=c(2,2))
plot(a,type="h",xlab="Outcomes", ylab="probabilities",main="n=10",lwd=2,col="blue")
plot(b,type="h",xlab="Outcomes", ylab="probabilities",main="n=15",lwd=2,col="blue")
plot(c,type="h",xlab="Outcomes", ylab="probabilities",main="n=7",lwd=2,col="blue")
plot(d,type="h",xlab="Outcomes", ylab="probabilities",main="n=3",lwd=2,col="blue")
```



vii. Also generate 50 random numbers from each distribution.

R Code

```
x=50;p=0.01
rgeom(x,p)
Output ---> [1]  36  31 169  82  35 168 123  41 215 114   0  80 254   6 250 113 204  15   0
[20] 174  41  48  30 132   1  22  22  10  14 171  29 153   2 170  31 139 233   1
[39]  61  12 266 231 354   9  38  41 205  26 229  31
```

viii. Find the First Second and Third Quartiles of each distribution under study.

```
x=c(.25,.5,.75);p=0.01
qgeom(x,p)
Output ---> [1]  28  68 137
```

ix. Calculate the 50 th , 70 th , 95 th percentiles.

```
x=c(.5,.7,.95);p=0.01
qgeom(x,p)
Output ---> [1]  68 119 298
```

x. Calculate the second decile and D6

```
x=c(.2,.6);p=0.01
qgeom(x,p)
Output ---> [1] 22 91
```

NEGATIVE BINOMIAL DISTRIBUTION

Example : In an NBA (National Basketball Association) championship series, the team that wins four games out of seven is the winner. Suppose that teams A and B face each other in the championship games and that team A has probability 0.55 of winning a game over team B.
What is the probability that team A will win the series in 6 games?

i. Construct the complete probability distribution.

```
x=0:6;n=x-k;k=4;p=0.55;
y=dnbinom(x=n,k,p)
tab=cbind(x,y)
tab
```

ii. Construct the Cumulative probability distribution.

```
x=0:7;n=x-k;k=4;p=0.55;
y=pnbinom(n,k,p)
tab=cbind(x,y)
tab
```

iii. Find the prob. of different events.

```
x=6;n=x-k;k=4;p=0.55;
dnbinom(x=n,k,p)
Output ---> [1] 0.1853002
```

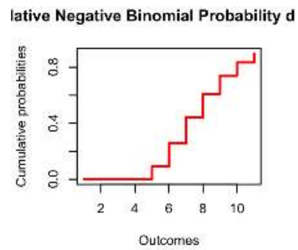
iv. Draw probability mass function

R Code

```
x=0:10;n=x-k;k=4;p=0.55;
y=dnbinom(x=n,k,p)
par(mfrow=c(1,2))
plot(y,type="h",xlab="Outcomes", ylab="probabilities",main="Negative Binomial Probability distribution",lwd=2,col="red")
```

v. Cumulative probability function.

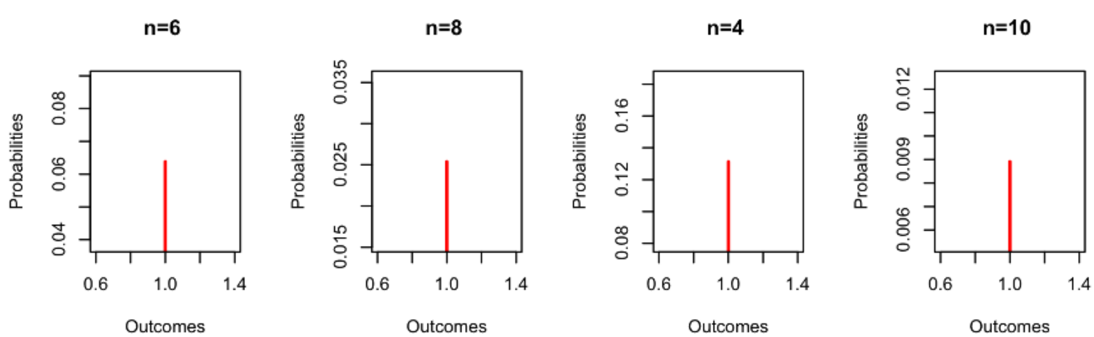
```
x=0:10;n=x-k;k=4;p=0.55;
y=pnbinom(n,k,p)
par(mfrow=c(3,3))
plot(y,type="s",xlab="Outcomes", ylab="Cumulative probabilities",main="Cumulative Negative Binomial
Probability distribution",lwd=2,col="red")
```



vi. Discuss the different behavior of the line graph when n will be changed.

R Code

```
x1=10;x2=12;x3=8;x4=14; k=4;
n1=x1-k; n2=x2-k; n3=x3-k; n4=x4-k;
p=0.55;
y1=dnbinom(n1,k,p)
y2=dnbinom(n2,k,p)
y3=dnbinom(n3,k,p)
y4=dnbinom(n4,k,p)
par(mfrow=c(3,4))
plot(y1,type="h",xlab="Outcomes", ylab="Probabilities",main="n=6
",lwd=2,col="red")
plot(y2,type="h",xlab="Outcomes", ylab="Probabilities",main="n=8
",lwd=2,col="red")
plot(y3,type="h",xlab="Outcomes", ylab="Probabilities",main="n=4
",lwd=2,col="red")
plot(y4,type="h",xlab="Outcomes", ylab="Probabilities",main="n=10
",lwd=2,col="red")
```



vii. Also generate 50 random numbers from each distribution.
x=54;n=x-k;k=4;p=0.55; rnbinom(n,k,p) [1] 2 3 5 1 5 6 4 6 2 10 2 3 2 2 2 0 1 2 3 4 5 1 5 0 4 [26] 3 3 6 6 2 2 1 1 0 5 5 4 4 3 5 0 0 1 4 2 3 6 1 3 0

viii. Find the First Second and Third Quartiles of each distribution under study.
x=c(.25,.5,.75);n=x-k;k=4;p=0.55; qnbinom(x,k,p) Output ---> [1] 1 3 5

ix. Calculate the 50 th , 70 th , 95 th percentiles.
x=c(.5,.7,.95);n=x-k;k=4;p=0.55; qnbinom(x,k,p Output ---> [1] 3 4 8

x. Calculate the second decile and D6
x=c(.2,.6);n=x-k;k=4;p=0.55; qnbinom(x,k,p) Output ---> [1] 1 3

NORMAL DISTRIBUTION

Example : An average light bulb manufactured by the Acme Corporation lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that an Acme light bulb will last at most 365 days?

i. Construct the complete probability distribution.

```
x=seq(100,400,by=50)
y=dnorm(x,mean=300,sd=50)
pr.distn<-data.frame(x,y)
Pr.distn
```

ii. Construct the Cumulative probability distribution.

```
x <- seq(100,400,by = 50)

y <- pnorm(x, mean = 300, sd = 50)

pr.distn<-data.frame(x,y)

pr.distnpr.distnpr.distn1<-transform(pr.distn, Fx=cumsum(y))

pr.distn1
```

iii. Find the prob. of different events.

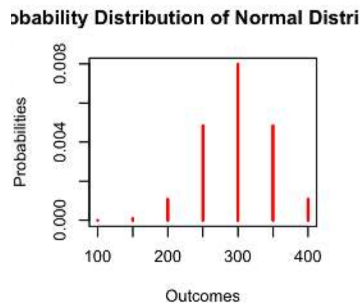
P(X<365)
R Code

```
x<-365;
pnorm(x,mean=300,sd=50,lower.tail=TRUE)
```

Output ---> [1] 0.9031995

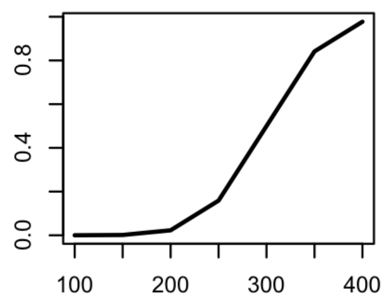
iv. Draw probability mass function

```
x=seq(100,400,by=50)
y=dnorm(x,mean=300,sd=50)
pr.distn<-data.frame(x,y)
par(mfrow=c(3,3))
plot(x,y,type="h",lwd=2,col="red",xlab="Outcomes", ylab="Probabilities", main="Probability Distribution of Normal Distribution")
```



v. Cumulative probability function.

```
x <- seq(100,400,by = 50)
plot(x, pnorm(x, mean=300, sd=50), xlab = "Diameter of Trunks", ylab = "P(X = x)", type = "l",lwd = "2")
 points(x, pnorm(x, mean=300, sd= 50),pch=19)
```



vi. Discuss the different behavior of the line graph when n will be changed.

```
rnorm(50, mean = 300, sd = 50)
```

vii. Also generate 50 random numbers from each distribution.

R Code

```
rnorm(50,mean=300,sd=50)
```

Output

```
> rnorm(50,mean=300,sd=50)
 [1] 260.8292 372.2065 342.2250 391.2505 335.3145 277.4920 283.1231 284.0512
 [9] 336.1504 427.5247 271.3364 304.9629 363.8846 291.7837 278.0754 228.0943
[17] 358.3166 354.1909 243.6309 328.8976 291.2688 227.9252 294.0452 301.0368
[25] 371.6805 287.3115 376.0972 166.4565 365.7307 235.0726 286.6787 314.2967
[33] 296.6799 330.9675 238.8761 307.6804 279.3104 355.5701 222.7099 254.7970
[41] 316.5530 248.9653 378.3781 258.6003 380.1350 329.3749 271.4360 390.8171
[49] 339.4292 245.5682
```

viii. Find the First Second and Third Quartiles of each distribution under study.

R Code

```
x=c(.25,.5,.75)
qnorm(x,mean=300,sd=50)
```

Output ---> [1] 266.2755 300.0000 333.7245

ix. Calculate the 50th , 70th , 95 th percentiles.

R Code

```
x=c(.5,.7,.95)
qnorm(x,mean=300,sd=50)
```

Output ---> [1] 300.0000 326.2200 382.2427

x. Calculate the second decile and D6

R Code

```
x=c(.2,.6)
qnorm(x,mean=300,sd=50)
```

Output ---> [1] 257.9189 312.6674