PERFORMANCE

Term	Remarks	Staff Member's Signature
I		
II	Complete	A.M.

THAKUR COLLEGE OF SCIENCE & COMMERCE

NAAC
Accredited
with Grade "A"
(3rd Cycle)

THAKUR
TRUSTS

ISO
9001 : 2015
Certified

Degree College

Computer Journal

CERTIFICATE

SEMESTER II UID No. ______

Class F.Y CS Roll No. 1840 Year ~~1840~~ 2019-2020

This is to certify that the work entered in this journal is the work of Mst. / ~~Ms.~~ Pushpraj ojha

______

who has worked for the year 2019-2020 in the Computer Laboratory.

A. Kouli

Teacher In-Charge

Head of Department

Date : 4.3.20

Examiner

Sem - 2

★★ INDEX ★★

27/11/19

029

Practical No: 1 27/11/19

Basic of R Software

1) R is a Software for Data analysis & Statical computing.

2) this software is used for effective Data haldling & output Storage is possible.

3) It is capable of Graphical Display.

4) It a free Software

```
> 2^2 + sqrt(25)+35
[1] 44
> 2*5*3 + 62/5 + sqrt(49)
[1] 49.4
> sqrt (76 + (4*2) + (9/5))
[1] 9.262829
> 43+ abs(-10) + 7^2 +3*9
[1] 129
```

```
> x=20
> y=30
> x+y
[1] 50
```

> $x^2 + y^2$

[1] 1300

> sqrt $(y^3 - x^3)$

[1] 137.8405

```
> abs(x-y)
[1] 10

> c(2,3,4,5)^2
[1] 4 9 16 25

> c(4,5,6,8)*3
[1] 12 15 18 24

> c(2,3,5,7)* c(-2,-3,-5,-4)
[1] -4 -9 -25 -28

> c(2,3,5,7) * c(8,9)
[1] 16 27 40 63

> c(2,3,5,7) * c(1,2,3)
[1] 2 6 15 7      Warning Message
 In c(2,3,7) * c(1,2,3)
      longer object length is not a multiple of
 shorter object length

> c(1,2,3,4,5,6) * c(2,3,4)
[1] 2 6 12 8 15 24
```


* find the Sum, Product, maximum, minimum of the values 5, 8, 6, 7, 9, 10, 15, 5

```
> e=c(5,8,6,7,9,10,15,5)
> sum(e)
[1] 65
> Prod(e)
[1] 11340000
> length(e)
[1] 8
> max(e)
[1] 15
> min(e)
[1] 5
```

Matrix:

```
> f<-matrix(nrow=4,ncol=2 ,Data=c(1,2,3,4,5,6,7,8))
> f
       [,1]   [,2]
[1,]    1      5
[2,]    2      6
[3,]    3      7
[4,]    4      8
```

* $X = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$, $y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ 10 & 6 & 12 \end{bmatrix}$

find: $x+y$, $x*y$, $2x+3y$

```
> x<-matrix[nrow=3, ncol=3, data=c(1,2,3,4,5,6,7,8,9))
> y<-matrix (nrow=3, ncol=3, data=c(2,-2,10,4,8,6,10,-11,-12))
> x+y
      [,1]  [,2]  [,3]
[1,]   3     8     17
[2,]         13    -3
```

> x * y

	[,1]	[,2]	[,3]
[1,]	2	16	70
[2,]	-4	40	-88
[3,]	30	36	108

> 2*(x)+3*y

	[,1]	[,2]	[,3]
[1,]	8	20	44
[2,]	-2	34	-17
[3,]	36	30	54

Table :-

```
>w = c(2,4,6,1,3,5,7,18,16,14,17,19,3,3,2,5,0,15,9,14,18,10,12)
>length(w)
>[1] 23
>a = table(w)
>transform(a)
```

	w	freq
1	0	1
2	1	1
3	2	2
4	3	3
5	4	1
6	5	2
7	6	0
8	7	0
9	9	0
10	10	0
11	12	0
12	14	1
13	15	0
14	16	0
15	17	0
16	18	0
17	19	1


```
>breaks = Seq (0,25, 5)
> b = cut(*w, breaks, right = FALSE)
> c = table(b)
transform(c)
```

b	freq
[0,5]	8
[5, 10]	5
[10, 15]	4
[15, 20]	6

AM

6/12/19

PRACTICAL NO:- 2

Q.1 Can the following be p.d.f :-

i) $f(x) = \begin{cases} 2-x & ; 1 \le x \le 2 \\ 0 & \text{other wise} \end{cases}$

ii) $f(x) = \begin{cases} 3x^2 & ; 0 < x < 1 \\ 0 & \text{other wise} \end{cases}$

iii) $f(x) = \begin{cases} \frac{3x}{2}\left(1-\frac{x}{2}\right) & ; 0 \le x \le 2 \\ 0 & \text{other wise} \end{cases}$

Solutions:

i) ~~To pro~~ $f(x) = \begin{cases} 2-x & ; 1 \le x \le 2 \\ 0 & \text{other wise} \end{cases}$

$= \int f(x)\,dx$

$= \int_1^2 (2-x)\,dx$

$= \int_1^2 2dx - \int_1^2 x\,dx$

$= 2x\Big|_1^2 - \frac{x^2}{2}\Big|_1^2$

$= (4-2) - (2-0.5)$

$= 2 - 1.5$

$= 0.5$

$\neq 1$

$\therefore$ This is not a p.d.f

ii) $f(x) = \begin{cases} 3x^2 & ; 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

$= \int f(x)\,dx$

$= \int_0^1 3x^2\,dx$

$= 3\int_0^1 x^2\,dx$

$= 3\frac{x^3}{3}\Big|_0^1$

$= 3\left(\frac{1^3}{3} - \frac{0^3}{3}\right)$

$= \frac{3}{3}$

$= 1$

$\therefore$ This is a p.d.f

iii) $f(x) = \begin{cases} \frac{3x}{2}\left(1-\frac{x}{2}\right) & ; 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$

$= \int f(x)\,dx$

$= \int_0^2 \frac{3x}{2}\left(1-\frac{x}{2}\right)dx$

$$= \int_0^2 \left[\frac{3x}{2} - \frac{3x^2}{4}\right] dx$$

$$= \int_0^2 \frac{3x}{2}dx - \int_0^2 \frac{3x^2}{4} dx$$

$$= \frac{3}{2}\int_0^2 x\,dx - \frac{3}{4}\int_0^2 x^2\,dx$$

$$= \frac{3}{2}\left[\frac{x^2}{2}\right]_0^2 - \frac{3}{4}\left[\frac{x^3}{3}\right]_0^2$$

$$\cancel{= \frac{3}{2}\left[\frac{4}{2} - 0\right] - \frac{3}{4}\left[\frac{4}{2}\right]}$$

$$\cancel{= \left(\frac{3}{2}\times 2\right) - \frac{3}{4}\left(\frac{1}{2}\right)}$$

$$= \frac{3}{2}\left[\frac{4}{2} - 0\right] - \frac{3}{4}\left[\frac{8}{3} - 0\right]$$

$$= \left(\frac{3}{2}\times 2\right) - \frac{3}{4}\left(\frac{8}{3}\right)^2$$

$$= 3 - 2$$

$$= 1$$

$\therefore$ This is p.d.f

Q.2 Can the following be P.m.f

i)

x	1	2	3	4	5
$P(x)$	0.2	0.3	-0.1	0.5	0.1

ii)

x	0	1	2	3	4	5
$P(x)$	0.1	0.3	0.2	0.2	0.1	0.1

iii)

x	10	20	30	40	50
$P(x)$	0.2	0.3	0.3	0.2	0.2

Solutions:

1) Since one Probability is negative
Hence it is not a P.m.f

2) Since $P(x) \geq 0 \quad \forall x$
and $\Sigma P(x) = 1$

$\therefore$ It is a P.m.f

```
> x = c(0,1,2,3,4,5)
> Prob = c(0.1,0.3,0.2,0.2,0.1,0.1)
> sum(Prob)
[1] 1
```

3) > a = c(10, 20, 30, 40, 50)
> Prob1 = c(0.2, 0.3, 0.3, 0.2, 0.2)
> Sum(Prob1)
[1] 1.2

Since $\Sigma P(\text{Prob1}) > 1$
$\therefore$ It is not a p.m.f

Q.3

x	0	1	2	3	4	5	6
$P(x)$	0.1	0.1	0.2	0.2	0.1	0.2	0.1

find:- $P(x \leq 2)$
$P(2 \leq x < 4)$
$P(\text{atleast } 4)$
$P(3 < x < 6)$

Solution:

$P(x \leq 2) = P(0) + P(1) + P(2)$
$= 0.1 + 0.1 + 0.2$
$= 0.4$

$\therefore P(x \leq 2) = 0.4$

$P(2 \leq x < 4) = P(2) + P(3)$
$= 0.2 + 0.2$
$= 0.4$

$\therefore P(2 \leq x < 4) = 0.4$

$P(\text{atleast } 4) = P(4) + P(5) + P(6)$
$= 0.1 + 0.2 + 0.1$
$= 0.4$
$\therefore P(\text{atleast } 4) = 0.4$

$P(3 < x < 6) = P(4) + P(5)$
$= 0.1 + 0.2$
$= 0.3$

$\therefore P(3 < x < 6) = 0.3$

Q.4

1)

x	0	1	2	3	4	5	6
$P(x)$	0.1	0.1	0.2	0.2	0.1	0.2	0.1

2)

x	10	12	14	16	18
$P(x)$	0.2	0.35	0.15	0.2	0.1

Solutions:-

1) $F(x) = 0$ if $x < 0$
$= 0.1$ if $0 \leq x < 1$
$= 0.2$ if $1 \leq x < 2$
$= 0.4$ if $2 \leq x < 3$
$= 0.6$ if $3 \leq x < 4$
$= 0.7$ if $4 \leq x < 5$
$= 0.9$ if $5 \leq x < 6$
$= 1$ if $x \geq 6$

2) $F(x) = 0$ if $x < 0$

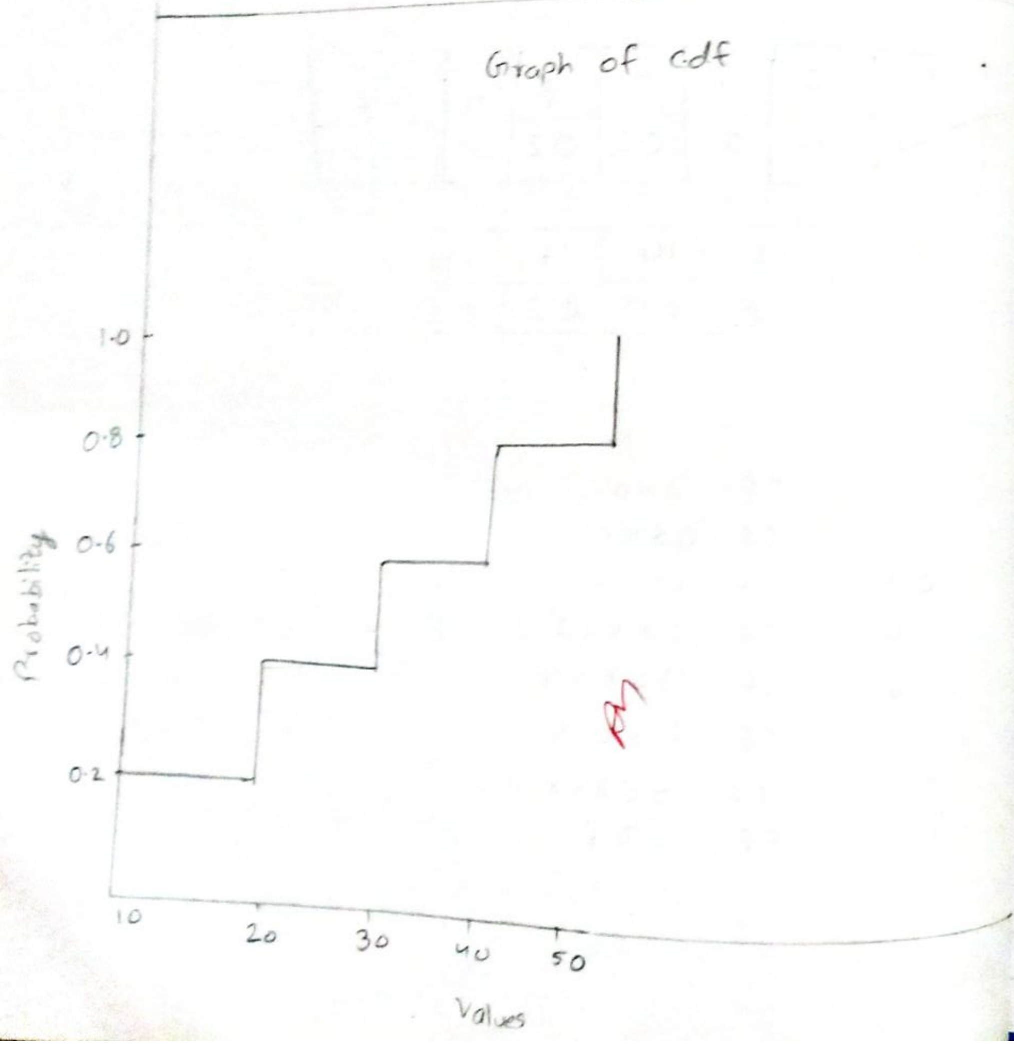
$= \cancel{0.55}\ 0.2$ if $0 \le x < 10$

$= \cancel{0.70}\ 0.55$ if $10 \le x < 12$

$= \cancel{0.90}\ 0.70$ if $12 \le x < 14$

$= 0.9$ if $14 \le x < 16$

$= 1$ if $16 \le x < 18$



11/12/19

Practical : 3

Probability & Binomial Distributions

Q1 find the cdf of following pmf & Draw the graph

X	10	20	30	40	50
P(x)	0.15	0.05	0.3	0.2	0.1

```
> x = c(10, 20,30,40,50)
> Prop = c(0,15,0,25,0.3,0.2,0.1)
> Cumsum = (Prop)
> Cumsum
[1] 0.15 0.25 0.30 0.20 0.10
> Cumsum (Prop)
[1] 0.15 0.25 0.30 0.20 0.10
> Plot(x, cumsum(Prop), xlab="values", ylab='Probability", main=
  "graph of cdf", "s")
```

Q2 Suppose there the 12 mcq in a test, Each question has 5 option & only one of them is correct find the Probability of having

i) 5 correct answer

ii) Atmost 4 correct answer

Solution:

it is given that

$n = 12$, $p = 1/5$, $q = 4/5$

X = Total number of correct answers

$x \sim B(n, p)$

>n=12
>p=1/5
>q=4/5
>x=5
[constant] i) > a = dbinom(5,12,1/5)
>a
[1] 0.05315022
(Atmost) ii) > b= pbinom(4,12,1/5)
[1] 0.9274445

Q3 There are 10 members in a commitee. Probability of any member attending a meeting is 0.9. find the Probability of
1) 7 member attended
2) Atleast 5 member attended
3) Atmost 6 member attended

Solution:-
n=10
p=0.9
q=0.1

X= Total no of member attended
X ~B(n, p)

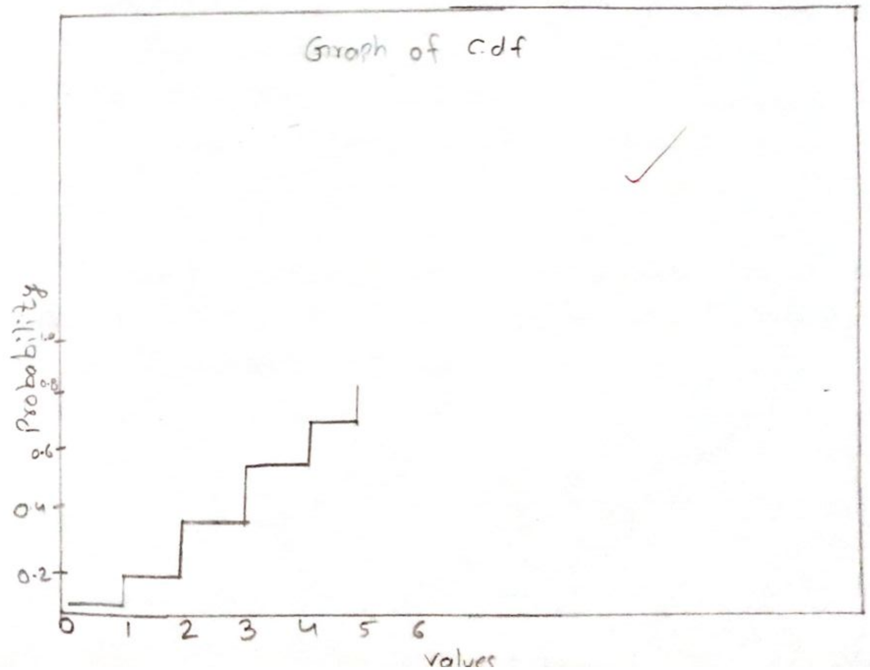
>n=10
>p=0.9
>q=0.1

i) > dbinom (7,10,0.9)
[1] 0.05739563
ii) > 1- pbinom(4,10,0.9)
[1] 0.998~~765~~531
iii) > pbinom (6,10,0.9)
[1] 0.01279520

Q4 find the Cdf & Draw the graph

X	0	1	2	3	4	5	6
P(x)	0.1	0.1	0.2	0.2	0.1	0.2	0.1

```
> x=c(0,1,2,3,4,5,6)
>Prob= c(0.1,0.1,0.2,0.2,0.1,0.2,0.1)
>Cumsum(Prob)
[1] 0.1  0.2  0.4  0.6  0.7  0.9  1.0
>Plot(x, Cumsum(Prob), xlab= "values", ylab=" Probability", main="Graph of cdf",
 "S")
```



AM
11-12-19

Practical: 4

Binomial Distribution

1) find the complete b.D when $n = 5$, $P = 0.1$

2) find Probability of exactly 10 Sucess in 100 trial with $P = 0.1$

Q.3 X follows B.D with $n = 12$, $P = 0.25$ find
 i) $P(x = 5)$
 ii) $P(x \le 5)$
 iii) $P(x > 7)$
 iv) $P(5 < x < 7)$

Q.4 The probability of Salesman make a sale to customer is 0.15. find the Probability:
 i) No Sale for 10 customer
 ii) More than 3 Sale in 20 Customer

Q.5 A Student write 5 mcq Each question has 4 options out of which 1 is correct Calculate the Probability for atleast 3 correct answers

Note:-

i) To find the value of x for which the Probability is P, the commond is qbino (P, n, P)

ii) $P(X = x)$ = dbinom (x, n, p)

iii) $P(X \leq x)$ = pbinom (x, n, p)

iv) $P(X > x)$ = 1 - pbinom(x, n, p)

Solutions:

1) > n = 5, P = 0.1

> dbinom (0:5, 5, 0.1)

[1] 0.59049 0.32805 0.7290 0.00810 0.00045 0.00001

2) > n = 100

> P = 0.1

> x = 10

> dbinom (10, 100, 0.1)

[1] 0.1318653

3) > n = 12

P = 0.25

i] > x = 5

> dbinom (5, 12, 0.25)

[1] 0.1032414

ii) > n = 12

> P = 0.25

> x = 5

> pbinom (5, 12, 0.025)

[1] 0.9455978

iii) > n = 12

> P = 0.25

> x > 7

> 1 - pbinom (7, 12, 0.25)

[1] 0.00278151

iv) > n = 12

> P = 0.25

> ~~x~~ - 5 < x < 7

> dbinom (6, 12, 0.25)

[1] 0.04014945

4) i) > n = 10

> P = 0.15

> x = 0

> dbinom (0, 10, 0.15)

[1] 0.1968744

ii) > n = 20

> P = 0.15

$P(x > 3) = 1 - P(x \leq 3)$

> 1 - pbinom (3, 20, 15)

[1] 0.774844

5) > n = 5

> x = 3

> P = 1/4 = 0.25

> 1 - $P(x \leq 2)$

> 1 - pbinom (2, 5, 0.25)

[1] 0.1035156

Note:-

i) $P(x > 7) = 1 - P(x \leq 7)$
= 1 - pbinom (7, 12, 0.25)

ii) $P(x \geq 7) = 1 - P(x \leq 6)$

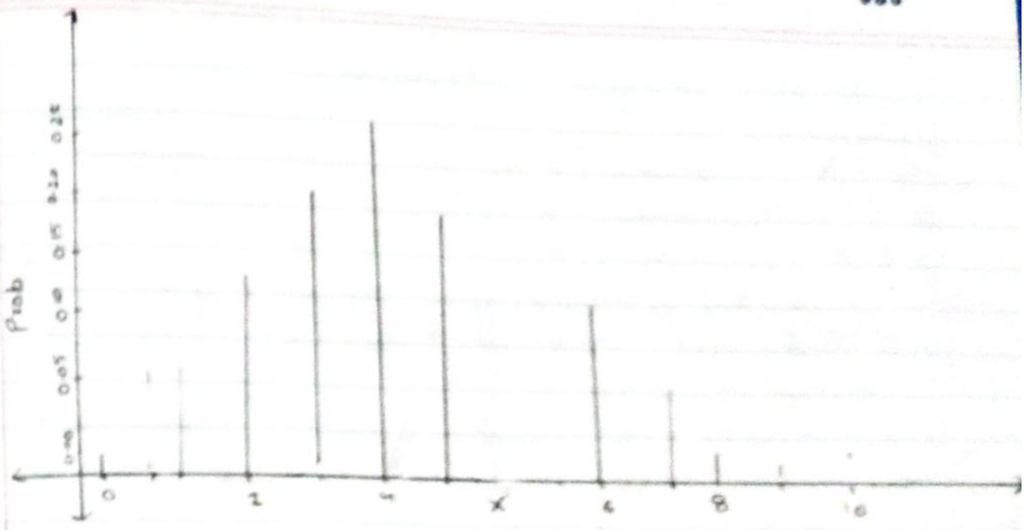
x follows binomial distribution with n=10, P=0.4
Plot the graph of Pmf & C.D.f

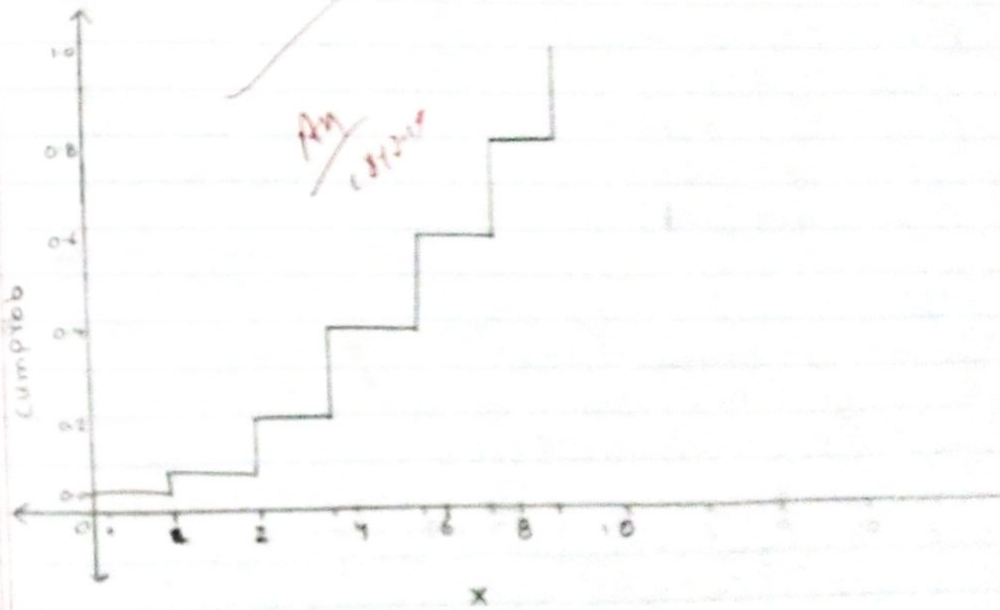
olution:-

```
> n=10
> P=0.4
> x=0:n
> Prob=dbinom(x,n,p)
> Cumprob= pbinom(x,n,p)
> d=dataframe("x values"=x, "Probability"=Prob)
> Print (d)
```

x.values	Probability
0	0.0060466176
1	0.0403107840
2	0.1209323520
3	
4	0.2508226560
5	0.2006581248
6	0.1114767360
7	0.0424673280
8	0.0106168320
9	0.0015728640
10	0.0001048576

```
> Plot (x, Prob, "h")
> Plot (x, cumprob, "s")
```



An

033

Practical: 5

Aim:- Normal Distribution

① $P[X = x] = dnorm(x, \mu, \sigma)$
② $P[X \leq x] = pnorm(x, \mu, \sigma)$
③ $P[X > x] = 1 - pnorm(x, \mu, \sigma)$
④ $P[x_1 < x < x_2] = pnorm(x_2, \mu, \sigma) - pnorm(x, \mu, \sigma)$
⑤ To find the value of K so that $P[x \leq k] = P_1$;
qnorm (P_1, μ, σ)
⑥ To generate 'n' random numbers rnorm$[n, \mu, \sigma]$

Q

1) $X \sim N(\mu = 50, \sigma^2 = 100)$
find. i) $P(x \leq 40)$
ii) $P(x > 55)$
iii) $P(42 \leq x \leq 60)$
iv) $P(x \leq K) = 0.7$; $K = ?$

Solⁿ:-

① >a = pnorm (40, 50, 10)
>cat("P(x<=40)=", a)
P(x<=40) = 0.1586553

② >b = 1- pnorm (55, 50, 10)
>cat ("P(x>55) = ", b)
P(x>55) = 0.3085375

③ >c = pnorm(60, 50, 10) - pnorm(42, 50, 10)
>cat("P(42≤x≤60) = ", c)
P(42≤x≤60) = 0.6294893

④ >d = qnorm(0.7, 50, 10)
>cat("P(x<=K) = 0.7, K = ", d)
P(x<=K) = 0.7. K = 55.24401

2) $X \sim N(\mu = 100, \sigma^2 = 36)$
find : i) $P(x \leq 110)$
ii) $P(x \leq 95)$
iii) $P(x > 115)$
iv) $P(95 \leq x \leq 105)$

Solⁿ:
1) $\mu = 100$
$\sigma^2 = 36$
$\therefore \sigma = 6$
>a = pnorm (110, 100, 6)
>cat ("P(x<=110) = ", a)
P(x≤110) = 0.9522096

2) >b = pnorm (95, 100, 6)
>cat("P(x≤95) = ", b)
P(x≤95) = 0.2023284

3) >c = 1- pnorm(115, 100, 6)
>cat ("P(x>115) = ", c)
P(x>115) = 0.006209665

4) >d = pnorm(105, 100, 6) - pnorm(95, 100, 6)
>cat ("P(95≤x≤105) = ", d)
P(95≤x≤105) = 0.953432

5) >e = qnorm (0.4, 100, 6)
>cat("P(x≤K) = 0.4, K = ", e)
P(x≤K) = 0.4, K = 98.47992

3) Generate 10 random number from normal Distribution with mean = 60 & S.d = 5 also calculate Sample mean, Variance, median & standard Deviation

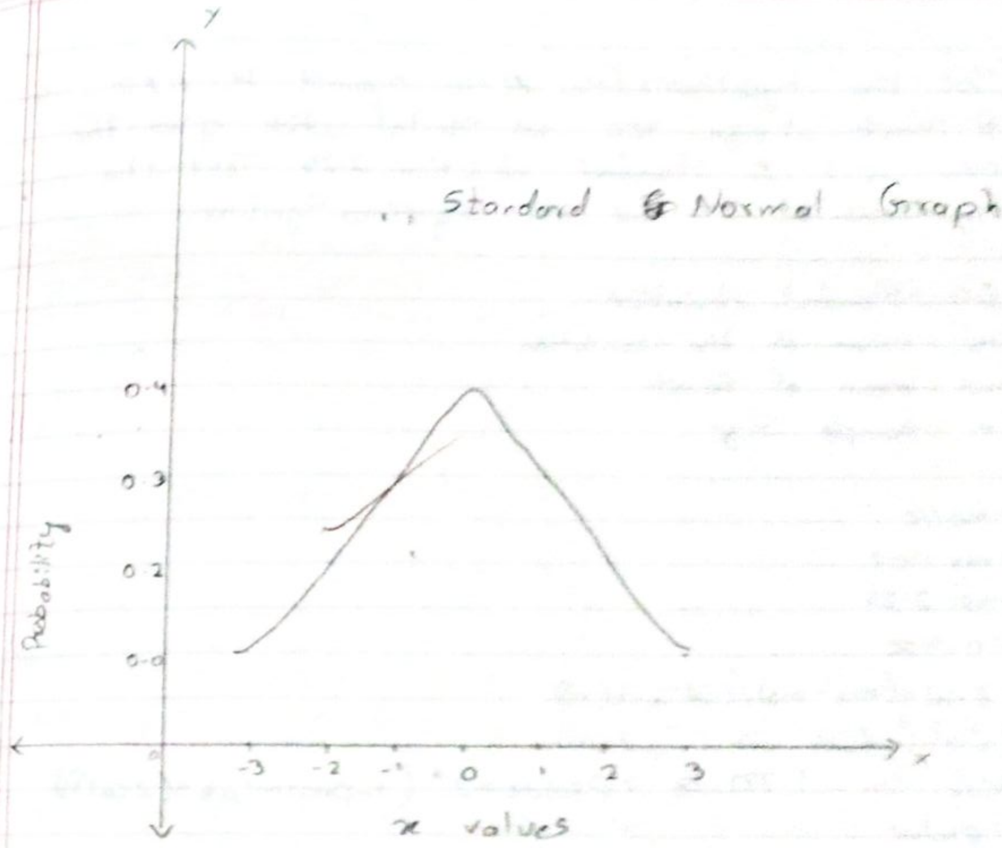
Solution:

```
> x = rnorm(10,60,5)
> am = mean(x)
> am
[1] 60.01123
> me = median(x)
> me
[1] 59.86623
> n=10
> variance = (n-1) * var(x)/n
> variance
[1] 44.45974
> sd = sqrt(variance)
> sd
[1] 6.667814
```

Q.4 Draw the graph of Standard normal Distribution

Solution:

```
> x = seq(-3, 3, by = 0.1)
> y = dnorm(x)
> plot (x,y, xlab = "x values", ylab = "Probability", main = "standard Normal Graph")
```

KM

Practical: 6 :: Z Distribution

1) Test the hypothesis (H_0): $\mu = 10$ against H_1: $\mu \neq 10$
A Sample of size 400 we selected which gives the mean 10.2 & standard deviation 2.25 Test the hypothesis at 5% level of ~~Signifince~~ Significans

S.D = Standard deviation
mo = mean of the population
mx = mean of Sample
n = Sample Size

```
> mo = 10
> mx = 10.2
> sd = 2.25
> n = 400
> zcal = (mx-mo)/(sd/sqrt(n))
> cat("zcal is = ", zcal)
zcal is = 1.777778 > pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.07544036
>
```

Since 0.07544036 is more than 0.05 we will accept the H_0

2) Test the hypothesis (H_0): $\mu = 75$ against (H_1): $\mu \neq 75$
A Sample of Size 100 is selected & sample mean is 80 with S.D of 3 test the hypothesis at 5% level of sigifincans

```
> mo = 75
> mx = 80
> s.d = 3
> n = 100
> zcal = (mx-mo)/(sd/sqrt(n))
> cat("zcal is = ", zcal)
zcal is = -66.66667 > pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
 [1] 0
>
```

Value accepted

3) Test the hypothesis(H_0): $\mu = 25$ against (H_1) $\mu \neq 25$ at 5% level of significans the following Sample is 30 Selected

20	24	27	30	46	35	26	~~45~~	27	10	20
22	30	37	35	21	~~22~~	23	24	25	26	27
19	28	29	30	39	27	15	22	20	18	

```
> x = c(20,24,27,30,46,35,26,27,10,20,22,30,37,35,21,23,24,25,26,27,
   19,28,29,30,39,27,15,22,20,18)
> mx = mean(x)
> mx
 [1] 26.06667
> n = length(x)
> n
  30
> Variance = (n-1)* var(x)/n
```

```
> variance
[1] 52.99556
> sd = sqrt(variance)
> sd
[1] 7.279805
> mo = 25
> zcal = (mx-mo)/(sd/sqrt(n))
> cat("zcal is = ", zcal)
zcal is= 0.8025454>
> Pvalue = 2*(1-pnorm(abs(zcal)))
> Pvalue
[1] 0.4222375
```

Value accepted

4) Experience has shown that 20% students of the college smokes a sample of 400 students reveal that out of 400 only 50 smokes test the hypothesis that experience give the correct proportion or not

```
> P = 0.2
> Q = 1-P
> p = 50/400
> n = 400
> zcal = (p-P)/(sqrt(P*Q/n))
> cat("zcal is ", zcal)
zcal is -3.75 > pvalue = 2*(1-pnorm(abs(zcal)))
> Pvalue
[1] 0.0001768346
> p
[1] 0.125
```

Value rejected ~~accepted~~

5) Test the hypothesis (H_0): $P = 0.5$ against $H_1: P \neq 0.5$ a sample of 200 is selected & the sample proportion is calculated as $p = 0.56$ Test the hypothesis 1% level of significans

```
> n = 200
> P = 0.5
> p = 0.56
> n = 200
> zcal = (p-P)/(sqrt(P*Q/n))
> cat("zcal is ", zcal)
 zcal is 1.697056>
> Pvalue = 2*(1-pnorm(abs(zcal)))
> Pvalue
[1] 0.08968602
p
[1] 0.56
```

H_0 ~~Value~~ accepted

A.K.
22.4.20

Practical 7 Large Sample Tests

1) A Study of noise level in two hospital is calculated below Test the hypothesis that the noise level in two hospital are Same or not

	Hos A	Hos B
No of Sample obs	84	34
mean	61	5
SD	7	8

Ho the noise levels are Same

```
> n1 = 84
> n2 = 34
> mx = 61
> my = 59
> sdx = 7
> sdy = 8
> z = (mx-my)/sqrt((sdx^2/n1) + (sdy^2/n2))
> z
[1] 1.273682
> cat("x calculated is = ", z)
x calculated is = 1.273682 >
> pvalue = 2*(1- pnorm (abs(z)))
> pvalue
[1] ~~0.0455026~~ 0.2027761
```

Since pvalue > 0.05, we ~~reject~~ accept Ho at 5% level of Significance

2) Two random Sample of Size 1000 & 2000 are drawn from two population with the means 67.5 & 68 respectively & with the Same s.d of 2.5, Test the hypothesis ~~of~~ that the mean of two population are equal.

Ho of the Population are Same

```
> n1 = 1000
> n2 = 2000
> mx = 67.5
> my = 68
> sdx = 2.5
> sdy = 2.5
> z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))
> z
[1] -5.163978
> cat("x calculated is = ", z)
x calculated is =
> pvalue = 2*(1- pnorm (abs (z)))
> pvalue
[1] 2.417564e-07
```

Since ~~p~~value > 0.5, we ~~accept~~ accept Ho at 5% level of Significance

3) In fy BSC 20% of a random sample of 400 student had deffective eye Sight In SY Class 15.5% of 500 Students have the Same deffect; is the difference of the Proportion is Same?

Ho - The Proportion of the Population are equal

```
> n1 = 400
> n2 = 500
> P1 = 0.2
> P2 = 0.155
> P = (n1*P1 + n2*P2)/(n1+n2)
> P
[1] 0.175
> z = (P1-P2)/sqrt(P*q*(1/n2 + 1/n2))
> z
[1] 1.76547
```

```
> cat('z calculated is =', z)
z calculated is = 1.765477
> pvalue = 2*(1-pnorm(abs(z)))
> pvalue
[1] 0.07748487
```

Since pvalue > 0.05 we accept Ho at 5% level of Significance

4) From each of the box of the apples a Sample Size of 200 is collected. It is found that there are 44 bad apples in the first Sample & 30 bad apples in second sample. Test the hypothesis that the two boxes are equivalent in term of Bad apples.

Ho The two box are equivalent

```
> n1 = 200
> n2 = 200
> p1 = 44/200
> p2 = 30/200
> p = (n1 * p1 + n2 * p2) / (n1+n2)
> p
[1] 0.185
> q = 1-p
> q
[1] 0.815
> z = (p1-p2) / sqrt(p * q * (1/n1 + 1/n2))
> z
[1] 1.802741
> pvalue = 2* (1-pnorm(abs(z)))
> pvalue
[1] 0.07142888
```

Since pvalue > 0.05, we accept Ho at 5% level of Significance

5) In M.A class out of a Sample of 60 mean height is 63.5 inches with a S.D = 2.5. In a M.com class out of 50 students mean height is 69.5 inches with a S.D = 2.5 Test the hypothesis that the mean of M.A & M.com class are Same.

Ho Heights of two classes are Same

```
> n1 = 60
> n2 = 50
> mx = 63.5
> my = 69.5
> sdx = 2.5
> sdy = 2.5
> z = (mx-my) / sqrt((sdx^2/n1) + (sdy^2/n2))
> z
[1] -12.53359
> pvalue = 2* (1-pnorm(abs(z)))
> pvalue
[1] 0
```

Since pvalue < 0.05, we reject Ho at 5% level of Significance.

AM 5.2.20

.../2/20

Practical-8 Small Sample Test

Q.I The tens are selected & height are found to be 63, 63, 68, 69, 71, 71, 72 cms Test hypothesis that mean height is are 66cm or not at 1

H_0: Mean = 66cms

```
> mean=66
> x=c(63,63,68,69,71,71,72)
> t.test(x)
```

One Sample t-test

data: x

t=47.94, df=6, P-value = 5.22e-09

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

64.66479 71.62092

Sample estimates:

mean of x

68.14286

∴ Pvalue <0.01 is ~~rejecte~~ rejected on H_0 in 1% level of Significance

Q.2 Two random sample was drawn from two different population

Sample 1 = 8, 10, 12, 11, 16, 15, 18, 7

Sample 2 = 20, 15, 18, 9, 8, 10, 11, 12

Test the hypothesis that there is no difference between the population mean at 5% los

H_0: there is no difference in the population Mean

```
> y=c(20,15,18,9,8,10,11,12)
> t.test(x,y)
```

Welch Two Sample t-test

data: x & y

t=-0.36247, df=13.837, P-value = 0.7225

alternative hypothesis: true difference in mean is not equal

-5.192719 3.692719

Sample estimates

mean of x mean of y

12.125 12.875

P-value <0.01 is accepted in H_0 on 1% level of Significance

Q.3 following are the weights of 10 people

Before=(100, 125, 95, 96, 98, 102, 115, 104, 109, 110)

After=(95, 80, 95, 98, 90, 100, 110, 85, 100, 101)

H_0 - The Diet Program is not effective

```
> x=c(100,125,95,96,98,102,115,104,109,110)
> y=c(95,80,95,98,90,100,110,85,100,101)
> t.test(x,y,Paired=T,alternative="less")
```

Paired t-test

data: x and y

t=2.3215, df=9, P-value=0.9773

alternative hypothesis: true difference in means is less than 0

95 Percent Confidence interval:

-Inf 17.89635

Sample estimates:

mean of the differences

10

P-value <0.01 is accepted in H_0 on 1% level of Significance

Q.4 Marks before & after a training program is given below
before = 20, 25, 32, 28, 27, 36, 35, 25
after = 30, 35, 32, 37, 37, 40, 40, 23
Test the hypothesis that training program is effactive or not

Ho - The training program is not effective

```
> x=c(20,25,32,28,27,36,35,25)
> y=c(30,35,32,37,37,40,40,23)
> ttest(x,y,Paired = T, alternative = "greater")
```

Paired t-test

data: x and y
t = -3.3859, df = 7, P-value = 0.9942
alternative hypothesis: true difference in means is greater th
95 Percent confidence interval:
-8.967399   Inf
Sample estimates:
mean of the difference
-5.75

P-value < 0.01 is accepted in Ho at 1% los

Q.5 Two random Sample were drawn from two normals Population & the values are
A = 66, 67, 75, 76, 82, 84, 88, 90, 92
B = 64, 66, 74, 78, 82, 85, 87, 92, 93, 95, 97
Test whether the population have Some Variance at 5% level of Sig

Ho variance of the population are Equal

```
> x=c(66,67,75,76,82,84,88,90,92)
> y=c(64,66,74,78,82,85,87,92,93,95,97)
> var test(x,y)
```

F test to compare two variances

data: x and y
F = 0.70686, num df = 8, denom df = 10, P-value = 0.6359
alternative hypothesis true ratio of variances is not equal to 1
95 percent confidence interval:
0.1833662   3.0360393
Sample estimates:
ratio of variances
0.7068567

Pvalue < 0.7068567 is accepted in Ho at 1% level of Significance

Q.6 The A.P of Sample 100 observation is 52 & S.D is 7 test the hypothesis that the population mean 55 or not at 5% of level of Significance.

Ho: Population mean = 55

```
> n = 100
> mx = 52
> mo = 55
> sd = 7
> zcal = (mx-mo)/(sd/sqrt(n))
> Pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 1.82153e-05
```

Pvalue < 0.01 is accepted in Ho at 1% level of Significance

AM
12.2.20

Practical: 9

Aim: Chi Square Distribution & ANOVA

use the following data to test whether the cleanliness of the home depends on the child

		Condition of Home	
		Clean	Dirty
Condition of Child	Clean	70	50
	fairly clean	80	20
	Dirty	35	45

H_0 condition of the home & child are independent

```
> x=c(70, 80, 35, 50, 20, 45)
> m=3
> n=2
> y=matrix(x,nrow=m, col=n)
> y
      [ ,1]   [ ,2]
[1, ]   70     50
[2, ]   80     20
[3, ]   35     45

> pv = chisq.test(y)
> pv
        Pearson's Chi-Squared test
data: y
X-squared = 25.646, df = 2, p-value = 2.698e-06
```

∴ p.value is less than 0.05, we ~~accept~~ reject H_0 at 5% LOS

Q.2 Table below shows the relation between performance of mathematic & Computer of CS students

		Maths		
		HG	MG	LG
Computer	HG	56	71	12
	MG	47	163	38
	LG	14	42	85

H_0 Performance of Mathematic & Computer are independent

```
> a=c(56,47,14,71,163,42,12,38,85)
> b=3
> c=3
> z=matrix(a,b,c)
> z
        [,1]   [,2]   [,3]
[,1]     56     71     12
[,2]     47    163     38
[,3]     14     42     85
> pv=chisq.test(z)
> pv1
        Pearson's chi-squared test
data: z
X-squared=145.78, df=4, p-value<2.2e-16
```

∴ p-value < 2.2e-16 is ~~rejected~~ less than 0.05, we reject H_0 at 5% LOS

3 Perform ANOVA for the following Data:-

Varieties	observations
A	50, 52
B	53,55,53
C	60,68,57,56
D	52, 54, 54, 55 .

H_0 the means of Varieties ABCD are equal

```
> x1 = c(50,52)
> x2 = c(53,55,53)
> x3 = c(60,68,57,56)
> x4 = c(52, 54, 54,55)
> d= stack(list(b1=x1, b2=x2, b3=x3, b4=x4))
> names(d)
[1] "values"  "ind"
> oneway.test(values~ind, data=d, var.equal=T)
        oneway analysis of means
data: values and ind
F=11.735, num df=3, denom df=9, P-value=0.00183
> anova = aov(values~ind, data=d)
> summary(anova)
```

	Df	Sum Sq	Mean Sq	Fvalues	Pr(>F)
ind	3	71.06	26.688	11.73	0.00183**
Residuals	9	18.17	2.019		

∵ p-value is less than 0.05, we ~~accept~~ reject the H_0 value at 5% LOS

Q:4 perform ANOVA for the following Data

Types	observations
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

H_0 the means of varieties ABCD are equal

```
> x1 = c(6,7,8)
> x2 = c(4, 6,5)
> x3 = c(8, 6,10)
> x4 = c(6,9,9)
> d= stack(list(b1=x1, b2=x2, b3=x3, b4=x4))
> names(d)
[1] "values"  "ind"

> oneway.test(values~ind, data=d var.equal=T)
        One-way analysis of means
data: Values and ind
F=2.6667, num df=3, denom df=8, P-value=0.1189
> anova= aov(values~ind, data=d)
> Summary(anova)
```

	DF	Sum Sq	mean Sq	F values	Pr(>F)
ind	3	18	6.00	2.667	0.119
Residuals	8	18	2.25		

∵ P-values is more than 0.05 we accept the H_0 value at

Q.5 write a R command to open ~~and fil~~ an Excel file in R Software

```
>x= read.csv('S:/users/admin/Destop/marks1.csv')
>x
     Stats   Cal
1      40    60
2      45    48
3      42    47
4      15    20
5      37    25
6      36    27
7      49    57
8      59    58
9      20    25
10     27    27
```

Practical: 10

11/3/20

Non-Parametric test

Q.1 following are amount of sulphide oxide emitated by factory.

17, 15, 20, 29, 90, 80, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26

Apply Sign test; to test the hypothesis that the population media 21.5 against the alternative it is less than 21.5

H_0: Population median = 21.5

H_1: It is less than 21.5

```
> x = c(17, 15, 20, 29, 90, 80, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26)
> m = 21.5
> sp = length (x [x > m])
> Sn = length (x [x < m])
> n = sp + Sn
> pv = pbinom (sp, n, 0.5)
> [1] 0.7482777
```

Since, p-value > 0.05, we accept the H_0 at 5% level of significance.

Q.2 for the observation 12, 19, 31, 28, 43, 40, 55, 49, 70, 63.

Apply sign test, to test population median 25 against the alternative more than 25.

→ H_0: median is 25

H_1: It is more than 25

```
> a = c(12, 19, 31, 28, 43, 40, 55, 49, 70, 63)
> m = 25
> sp = length (a[a > m])
> Sn = length (a[a < m])
> n = sp + Sn
```

[1] > n

[1] 10

> pv = pbinom (Sn, n, 0.5)

> pv

0.0546845

Since p-value > 0.05, we accept the H_0 at 5% los

3 for the following data

60, 65, 63, 89, 61, 71, 58, 51, 48, 66

Test the hypothesis using wilcoxon's sign of rank test for testing the alternative it is greater than 60

H_0: median is 60

H_1: It is greater than 60

```
> x = c(60, 65, 63, 89, 61, 71, 58, 51, 48, 66)
> mu = 60
wilcox.test(x, alter = 'greater', mu = 60)
```

wilcoxn signed rank test with continuity correction

data: x

v = 29, p-value = 0.2386

alternative hypothesis: true location is greater than 60

Since p-value > 0.05, we accept the H_0 at 5% los

Note: If the alternative is less, we have to write wilcox.test(x, alter = "less", mu = 60.

If the alternative is not equal to we have to write ~~wilcoxon signed~~ wilcox.test(x, area = "two-sided", mu = 60))

Q.4 Test the hypothesis, median is 12 against the alternative that it is less than 12 using wilcoxon's value: 12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 10

→ H_0: Median is 12

H_1: It is less than 12

```
> x = c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 10)
> mu = 12
wilcox.test(x, alter = "less", mu = 12)
```

wilcoxon signed rank test with continuity correction

data: x

v = 25, p-value = 0.2521

alternative hypothesis: true location is less than 12

Since, p-value > 0.05, we accept the H_0 at 5% los

M
4.7.20