Practical No:- 1

* Basics of R-Software

R is a software for data analysis of statistics computing. Thus software is used for effective data handling & output storage is possible.
It is capable of graphical display.
It is a free software.

1) $2^2 + \sqrt{25} + 35$
→ 2^2 + sqrt(25) + 35
= 44

2) $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$
→ 2 * 5 * 3 + 62 ÷ 5 + sqrt(49)
= 49.4

3) $\sqrt{75 + 4 \times 2 + 9 \div 5}$
→ sqrt (75 + 4 * 2 + 9/5)
= 9.262829

4) $42 + |-10| + 7^2 + 3 \times 9$
→ 42 + abs (-10) + 7^2 + 3*9
= 128

Q.] $x = 20$; $y = 30$
Find, $x+y$; x^2+y^2 ; $\sqrt{y^3-x^3}$; abs(x-y)

i) $x+y$
$= 50$

ii) x^2+y^2
→ x^2+y^2
$= 1300$

iii) $\sqrt{y^3-x^3}$
→ y^3-x^3
$= 137.8405$

iv) abs(x-y)
$= 10$

Q] Calculate the following.

1) c(2,3,4,5)^2
→ 4 9 16 25

2) c(4,5,6,8) * 3
→ 12 15 18 24

3) c(2,3,5,7) * c(-2,-3,-5,-4)
→ -4 -9 -25 -28

4) c(2,3,5,7) * c(8,9)
→ 16 27 40 63

5) c(1,2,3,4,5,6)^c(2,3)
→ 1 8 9 64 25 216

Scanned with CamScanner

2) Find the sum, prod, min, max of the given value
5, 8, 6, 7, 9, 10, 15, 5
x = c(5, 8, 6, 7, 9, 10, 15, 5)
length(x)
= 8

sum(x)
= 55

max(x)
= 15

min(x)
= 5

prod(x)
= 11340000

Matrix calculation

1) $\begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$

→ x ⟵ matrix(nrow = 4, ncol = 2, data = c(1, 2, 3, 4, 5, 6, 7, 8))

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

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))

i) x+y

$$\begin{bmatrix} 3 & 8 & 17 \\ 0 & 13 & -3 \\ 13 & 12 & 21 \end{bmatrix}$$

ii) x * y

$$\begin{bmatrix} 2 & 16 & 70 \\ -4 & 40 & -88 \\ 30 & 36 & 108 \end{bmatrix}$$

iii) 2 * x + 3 ** y

$$\begin{bmatrix} 8 & 20 & 44 \\ -2 & 34 & -17 \\ 36 & 30 & 54 \end{bmatrix}$$

Q] Perform the following

x = c(2, 4, 6, 1, 3, 5, 7, 18, 16, 14, 17, 19, 3, 3, 2, 5, 10, 15,
14, 18, 10, 12)

→ length (x)
23
a = table (x)
transform (a)

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

```
breaks = seq(0,20,5)
b = cut(x, breaks, light=FALSE)
c = table(b)
transform(c)
```

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

780

Practical No:- 2

iii)

Title :- Problems ~~of~~ on P.D.F and C.D.F

Q.1] Can the following be P.D.F?

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

To prove :- $\int f(x)\,dx = 1$

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

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

$= [2x]_1^2 \left[\frac{-x^2}{2}\right]_1^2$

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

$\neq 1$

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

To prove :- $\int f(x)\,dx = 1$

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

$= \left[\frac{3x^3}{3}\right]_0^1$

$= \frac{\cancel{3}}{\cancel{3}} - 0$

$a^n = \frac{a^{n+1}}{n+1}$

$= 1$ Hence it is P.D.F.

Scanned with CamScanner

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

To prove :- $\int f(x)\,dx = 1$

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

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

$$= \left[\frac{3x^2}{4}\right]_0^2 - \left[\frac{3x^3}{12}\right]_0^2$$

$$= 3 - 2$$

$$= 1$$

Hence, it 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

$P(3) = -0.1$

Since, one probability is negative
Hence, it is not P.M.F.

ii)

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

Since $P(x) \geq 0 \; \forall x$
and $\Sigma P(x) = 1$
Hence, it is a P.M.F.

iii>

x	10	20	30	40	50
$P(x)$	0·2	0·3	0·3	0·2	0·2

$\Sigma P(x) = 0\cdot2 + 0\cdot3 + 0\cdot3 + 0\cdot2 + 0\cdot2$
$= 1\cdot2$

Since $\Sigma P(x) \neq 1$

$P(x)$ is not a P.M.F.

Q3] Find $P(x \leq 2)$, $P(2 \leq x < 4)$, $P(\text{atleast } 4)$, $P(3 < x < 6)$

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

i> $P(x \leq 2) = P(0) + P(1) + P(2)$
$= 0\cdot1 + 0\cdot1 + 0\cdot2$
$= 0\cdot4$

ii> $P(2 \leq x < 4) = P(2) + P(3)$
$= 0\cdot2 + 0\cdot2$
$= 0\cdot4$

iii> $P(\text{atleast } 4) = P(4) + P(5) + P(6)$
$= 0\cdot1 + 0\cdot2 + 0\cdot1$
$= 0\cdot4$

iv> $P(3 < x < 6) = P(4) + P(5)$
$= 0\cdot1 + 0\cdot2$
$= 0\cdot3$

Q.4. Find c.d.f.

i)

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

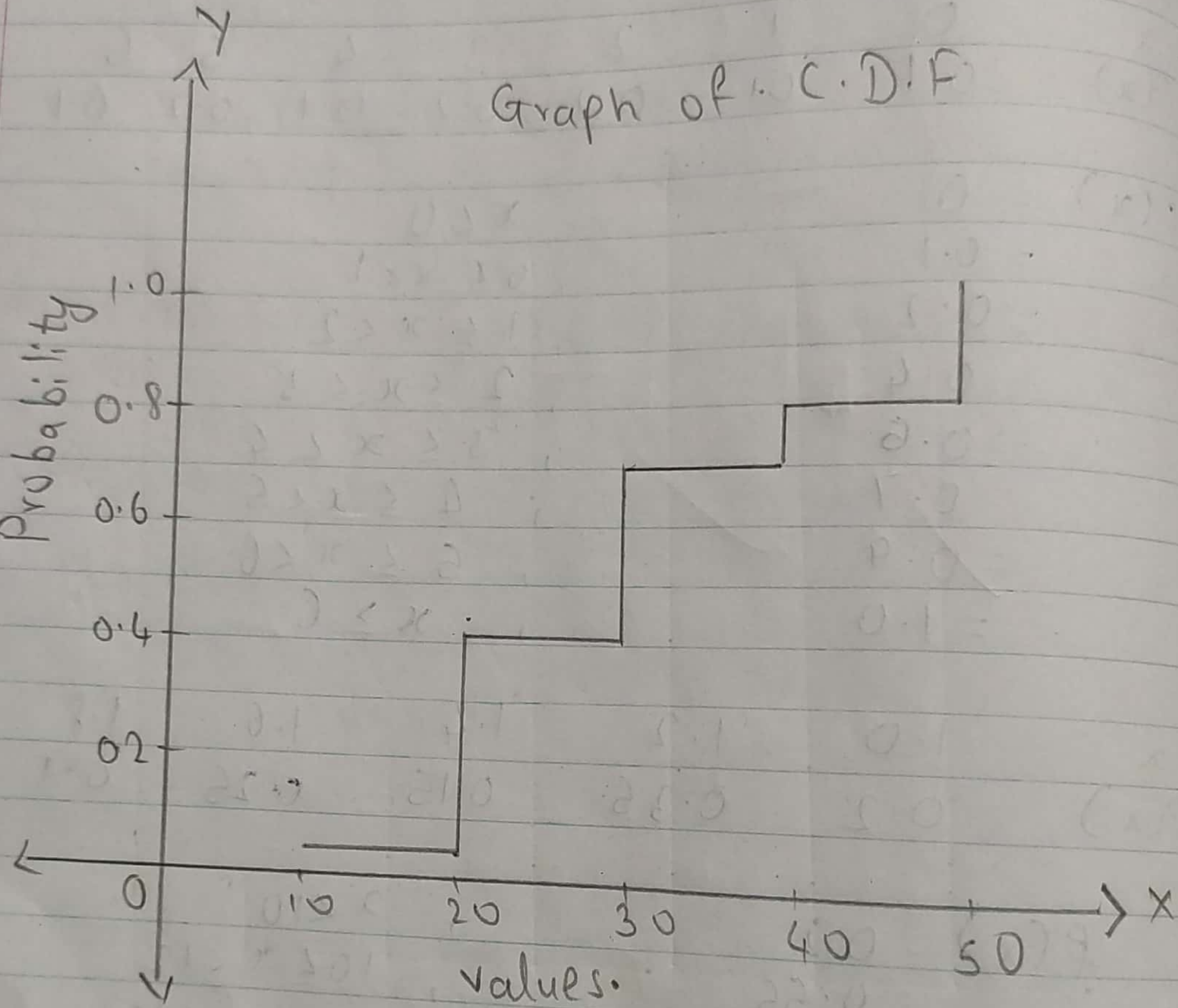
$$
f(x) = \begin{cases}
0 & , x < 0 \\
0.1 & , 0 \le x < 1 \\
0.2 & , 1 \le x < 2 \\
0.4 & , 2 \le x < 3 \\
0.6 & , 3 \le x < 4 \\
0.7 & , 4 \le x < 5 \\
0.9 & , 5 \le x < 6 \\
1.0 & , x \ge 6
\end{cases}
$$

ii)

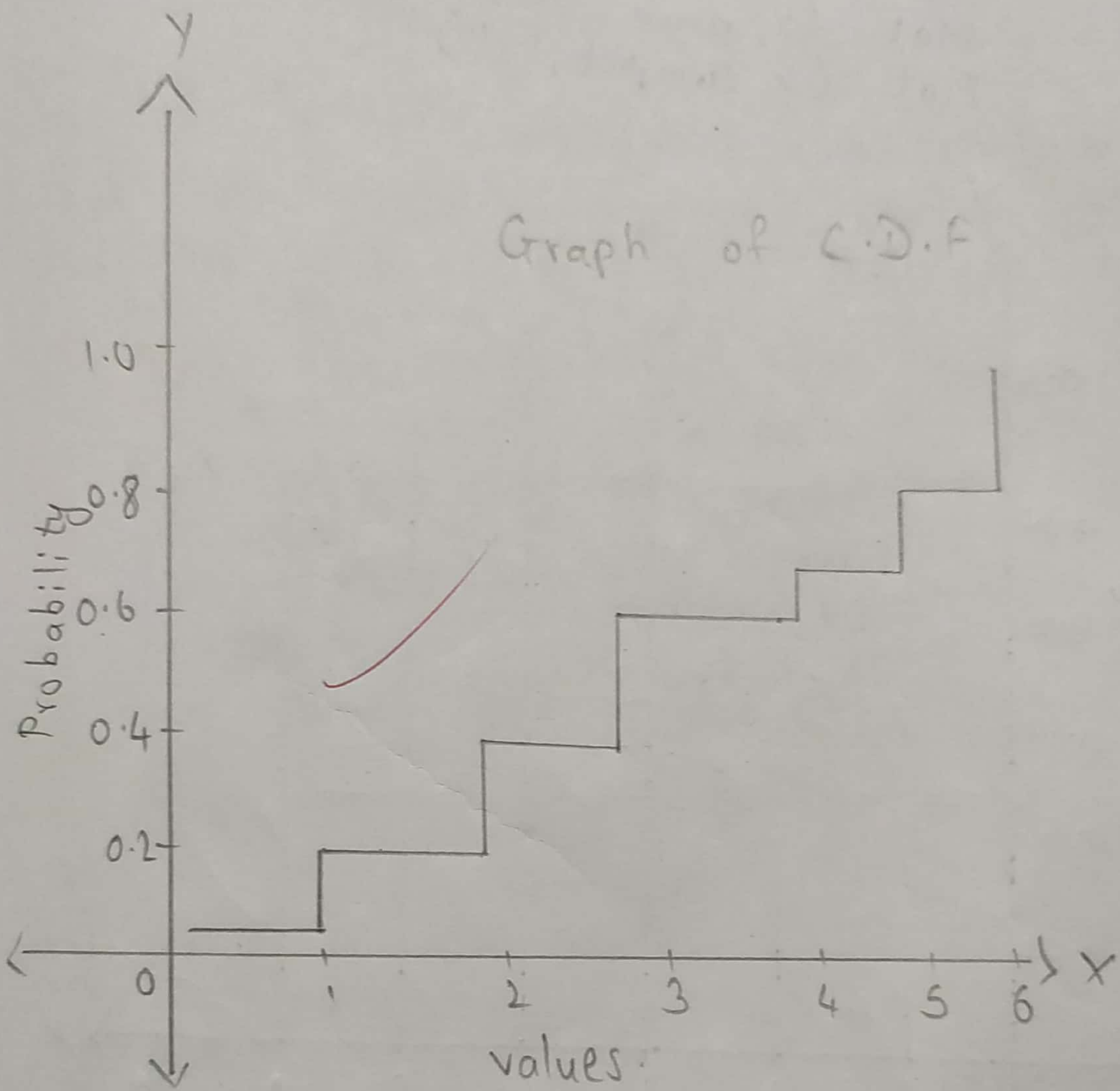
x	1.0	1.2	1.4	1.6	1.8
$P(x)$	0.2	0.35	0.15	0.25	0.1

$$
f(x) = \begin{cases}
0.2 & , x < 10 \\
0.55 & , 10 \le x < 12 \\
0.70 & , 12 \le x < 14 \\
0.90 & , 14 \le x < 16 \\
\cancel{0}1.0 & , x \ge 18
\end{cases}
$$

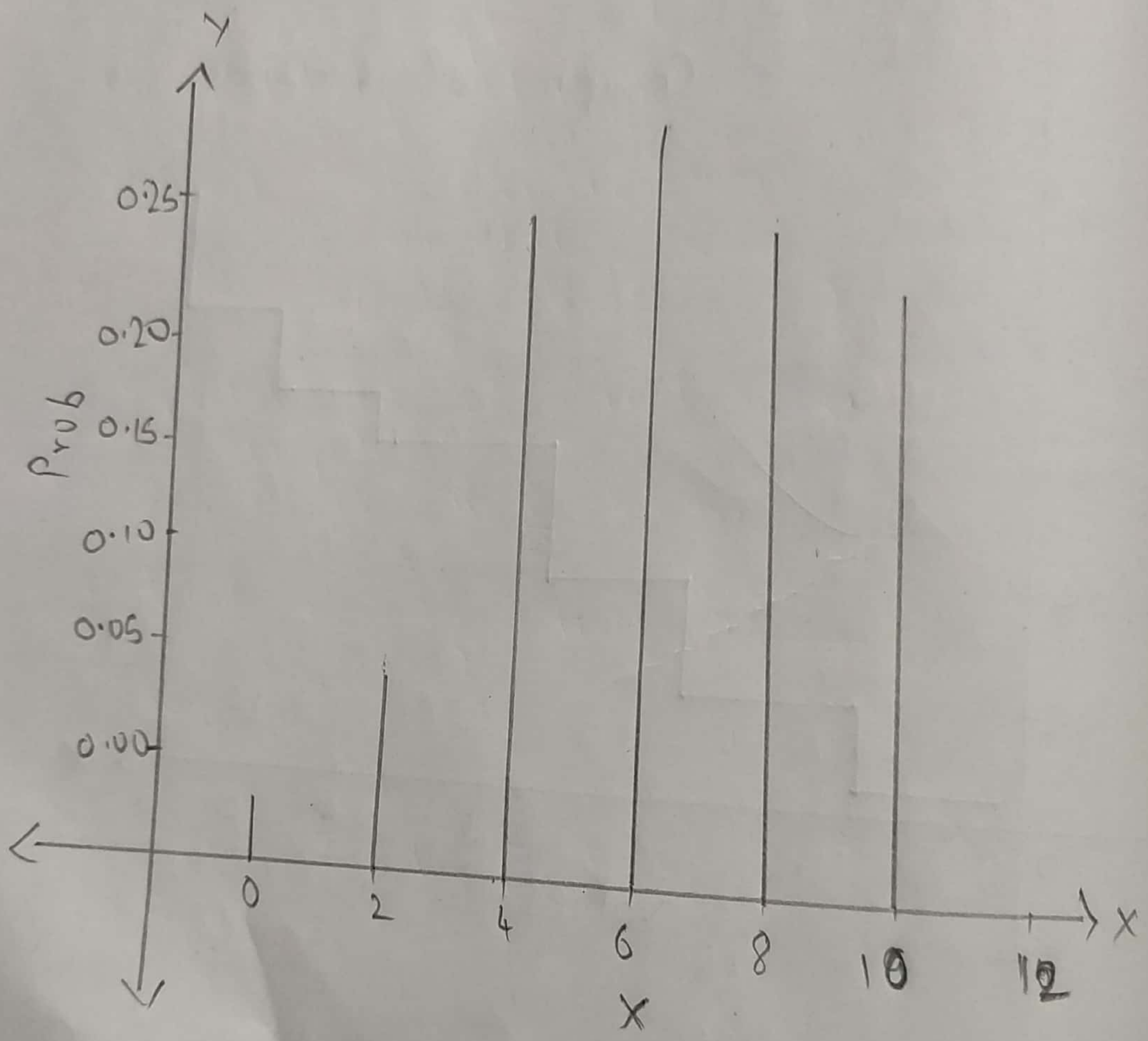
Graph of C.D.F



Graph of C.D.F



```
plot (x, prob, "h")
plot (x, cumprob, "s")
```



Practical No:- 3

Title:- Binomial Distribution and Probability

Q.] Find the c.d.f of the following P.D.F. and draw the graph.

x	10	20	30	40	50
$P(x)$	0.15	0.25	0.3	0.2	0.1

$P(x) = 0$ if $x < 10$
$= 0.15$ $10 \le x < 20$
$= 0.40$ $20 \le x < 30$
$= 0.70$ $30 \le x < 40$
$= 0.90$ $40 \le x < 50$
$= 1.00$ $x \ge 50$

```
-> x = c(10,20,30,40,50)
x
[1]   10  20  30  40  50
P(x) = c(0.15, 0.25, 0.3, 0.2, 0.1)
cumsum(P(x))
[1]  0.15   0.40   0.70   0.90   1.00
P(c)f (x, cumsum(Px), xlab = "values", ylab = "probability")
       Main= "Graph of C.D.f ", "s")
```

033

Q.] Suppose there are 12 MCQ's in a ~~test~~ each question has 5 options & only one is correct. Find the probability of having :-

1> Five correct answers.

2> Almost four correct answers.

→ It is given that

$n = 12$, $P = \frac{1}{5}$, $Q = \frac{4}{5}$

x = Total no. of correct answers.

$x \sim B(n, p)$

1> $n = 12$; $p = 1/5$, $Q = \frac{4}{5}$; $x = 5$

~~x = Total no. of correct answers~~

d binom (5, 12, 1/5)

[1] 0.05315

2> $n = 12$; $P = 1/5$; $q = 4/5$; $x = 4$

p binom (4, 12, 1/5)

[1] 0.9274

Q.] There are 10 members in a committee. The probability of any member attending a meeting is 0.9. Find the probability.

1> 7 members attended.

2> Atleast ~~8~~ 5 members attended

3> Atmost 6 members attended

→ It is given that

$n = 10$, $P = 0.9$, $Q = 0.1$

Q

1) $n=10$; $P=0.9$, $Q=0.1$, $x=7$
dbinom(7,10,0.9)
[1] 0.05739

2) $n=10$; $P=0.9$, $q=0.1$, $x=5$
1- Pbinom (5,10,0.9)
[1] 0.99836

3) $n=10$; $P=0.9$, $q=0.1$ $x=6$
Pbinom(6,10,0.9)
[1] 0.01279

Q]. Find the C.D.F ~~of~~ & 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

$$
f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.1 & 0 \le x < 1 \\ 0.2 & 1 \le x < 2 \\ 0.4 & 2 \le x < 3 \\ 0.6 & 3 \le x < 4 \\ 0.7 & 4 \le x < 5 \\ 0.9 & 5 \le x < 6 \\ 1.0 & x \ge 6 \end{cases}
$$

Scanned with CamScanner

O.P.O

→ x = c(0,1,2,3,4,5,6)

x

[1] 0 1 2 3 4 5 6

Px = c(0.1,0.1,0.2,0.2,0.1,0.2,0.1)

cumsum(Px)

[1] 0.1 0.2 0.4 0.6 0.7 0.9 1.0

PDf:(x, cumsum(Px), xlab="values", ylab="probab

main:"Graph of C.D.f:", "S")

Practical No:- ~~5~~ 4

Q.1] Find the complete binomial distribution
$n=5$, & $P=0.1$

2]. Find the probability of exactly 10 success in 100 trials ~~trails~~.
with $P=0.1$

3]. X follows binomial distribution with $n=12$, $P=0.25$
Find i) $P(X=5)$
ii) $P(X\le 5)$
iii) $P(X>7)$
iv) $P(5<x<7)$

4]. Probability of salesman makes a sales to be customer is 0.15. Find the probability.
i) No sale for 10 customers.
ii) More than 3 sale in 20 customer.

5]. A student writes 5 Mcq each questions has 4 options out of which 1 is correct. Calculate the probability for atleast 3 correst answers.

6] X follows binomial distribution $n=0$, $P=0.4$, Plot the graph PMF & CDF.

1> n=5, P=0.1
→ dbinom (0:5, 5, 0.1)
[1] 0.59049 0.32805 0.07290 0.00810 0.00045
0.

2. n=100, x=10, P=0.1
→ dbinom (10, 100, 0.1)
[1] 0.1318653

3. n=12, P=0.25
→ (i) P(x=5)
dbinom (5, 12, 0.25)
[1] 0.1032414

ii) P(x ≤ 5)
pbinom (5, 12, 0.25)
[1] 0.9455978

iii) P(X>7)
P(X>7) = 1-P(X ≤ 7) = 1-pbinom (7, 12, 0.5)
[1] 0.00278151

iv) P(5 < x < 7)
dbinom (6, 12, 0.25)
[1] 0.00278151

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 ≤ 3) = 1-Pbinom(3, 20, 0.15)

[1] 0.3522748

5) n=5, P=1/4 , x ≥ 3 or x ≤ 2

→ 1-Pbinom (2,5,1/4)

[1] 0.1035156

6) n=10 ; P=0.4

x = 0: n

prob = dbino(x,n,p)

cumprob = pbinom (x,n,p)

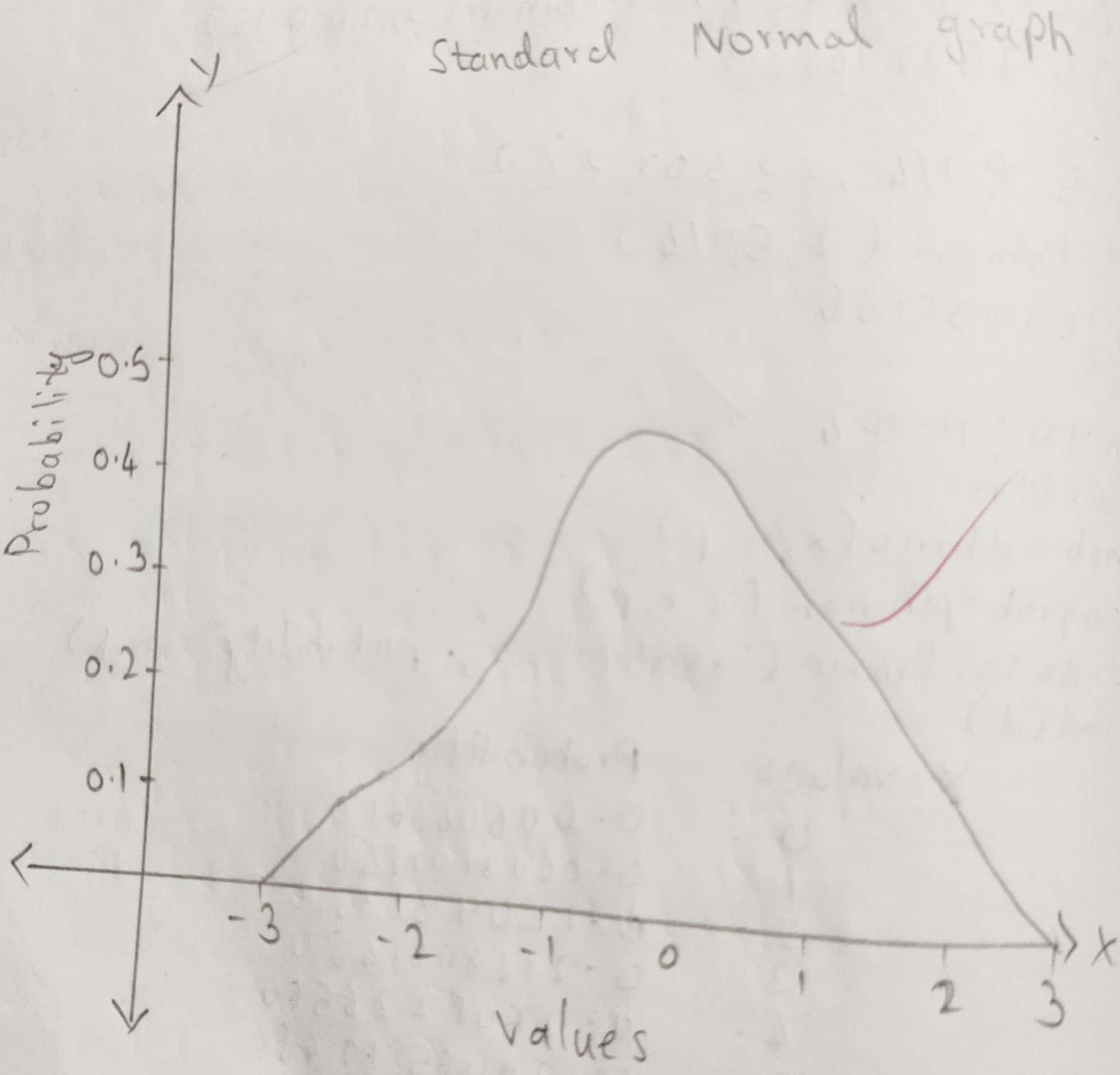
d=data. frame ("xvalues" = x, "probability" = prob)

Print(d)

	X.values	Probability
1	0	0.006046617
2	1	0.0403107840
3	2	0.1209 323520
4	3	0.2149908480
5	4	0.2508226560
6	5	0.2006581248
7	6	0.1114767360
8	7	0.0424673280
9	8	0.0106183200
10	9	0.0015728640
11	10	0.0001048576

Draw the graph of standard normal distribution.

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



Practical No:-5

Normal Distribution

1) $P[X=x] = \text{dnorm}(x, \mu, \sigma)$

2) $P[X \leq x] = \text{pnorm}(x, \mu, \sigma)$

3) $P[X > x] = 1 - \text{pnorm}(x, \mu, \sigma)$

4) $P[x_1 < x < x_2] = \text{pnorm}(x_2, \mu, \sigma)$

5) To find the value of k so that :-
$P(X \leq k) = P_1$; $\text{qnorm}(P_1, \mu, \sigma)$

6) To generate 'n' random numbers :-
$\text{rnorm}(n, \mu, \sigma)$

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

Q.2]. $X \sim N(\mu = 100, \sigma^2 = 36)$
i) $P(X \leq 110)$
ii) $P(X \leq 95)$
iii) $P(X > 115)$
iv) $P(95 \leq X \leq 105)$
v) $P(X \leq k) = 0.4$; $k = ?$

Q.3). Generate 10 random numbers from normal distribution with mean(μ) = 60, S.D = 5 Calculate the sample, mean, median, variance, S.D.

Q.4). Draw the graph of standard normal distribution.

Q.1).

→ i> a=pnorm (40,50,10)

cat("P (x<=40) is =",a)

P(x<=40) is =0.1586553

ii> b=1-pnorm (55,50,10)

cat("P (x>55) is=",b)

P(x>55) is= 0.8085375

Scanned with CamScanner

Q2.]

i> a=pnorm (110,100,6)
cat("P(x<=110) is=",a)
P(x<=110) is = 0.9522096

ii> b = pnorm(95,100,6)
cat("P(x<=95) is =",b)
P(x<=95) is = 0.2023284

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

iv> d=pnorm (95,100,6) - pnorm(105,100,6)
cat("P(95<=x<=105) is =",d)
P(95<=x<=105) is = -0.593432

v> e = qnorm (0.4,100,6)
cat("P(x<=k)=0.4 ,k is =",e)
P(x<=k) = 0.4 ,k is = 98.47992

LAD

Q.3.]

→ n=10, r=10, σ=5

x = rnorm (10,60,5)

[1] 66.05727 51.97402 59.97334 60.87027

58.27801 56.90084 63.89487 56.54952

[9] 58.59008 65.11264

am=mean(x)

>am

[1] 59.81959

me=median(x)

>me

[1] 59.815959

variance=(10-1) * var(x)/10

>variance

[1] 16.93294

sd=sqrt(variance)

>sd

[1] 4.114966

Practical No:- 6

Z distribution

1) $H_0 : \mu = 10$ against $H_1 : \mu \neq 10$. A sample of size 400 are selected which gives a mean 10.2 at a standard deviation 2.25. Test the hypothesis at 5% level of significance.

mo = (mean of population) = 10
mx (mean of sample) = 10.2
sd (standard deviation) = 2.25
n = (sample size) = 400

```
> zcal (mx-mo) / (sd/sqrt(n)))}
> cat ("zcal is :", zcal )
> zcal is : 1.79
> pvalue = 2 * (1-pnorm (abs(zcal)))
> pvalue
> 0.075
```

Thus, hypothesis accepted.

2) Test the hypothesis
$H_0 : \mu = 75$ $H_1 : \mu \neq 75$
A sample of size 100 is selected and the sample mean is 80 with the standard deviation of 3. Test the hypothesis at 5% level of significance.

mo = 75
mx = 80
sd = 3
n = 100

zcal = (mx - mo) / (sd / sqrt(n))
cat ("zcal is:", zcal)
zcal is : 16.6667
pvalue = 2 * (1-pnorm (abs(zcal)))
pvalue
0
Thus, hypothesis rejected.

3) Test the hypothesis
H_0 $\mu = 25$ against H_1 $\mu \neq 25$
A sample of 30 is selected. Test the hypothesis
5% level of ~~signifa~~ significance. A sample of following
30 is selected 20, 24, 27, 35, 30, 46, ~~26~~, 27, 10, 2
30, 37, 35, 21, 22, 23, 24, 25, 26, 27, 28, ~~29~~, 30, 39, 27,
19, 22, 20, 18.
x = c(20, 24, 7, 35, ... , 19, 22, 20, 18)
mx = mean(x)
mx = 26.066
n = length (x)
variance = (n-1) * var (x) / n
sd = sqrt (variance)
sd = 7.2798.
mo = 25
zcal = (mx - mo) / (sd / sqrt (n))
zcal = 0.8025
pvalue = 2 * (1-pnorm (abs (zcal))
pvalue = 0.4223
Thus, hypothesis rejected.

4) Experience has shown that 20% students of a college smokes. A samples of 400 students reveal that out of 400 only 50 smokes. Test the hypothesis that the experience keeps the correct proportion or not.

046

p = 0.2
Q = 1-P
p = 50/400
n = 400
zcal = (p-P)/(sqrt(P* Q/n))
zcal
-3.75
pvalue = 2* (1-~~pnorm~~ (abs(zcal)))
pvalue
0.0001768346

Thus, hypothesis is rejected.

5) Test the hypothesis : $H_0 : P = 0.5$ against $P \neq 0.5$
A sample of 200 is selected where sp = 0.56 and pp = 0.5 level of significance is 1%.

P = 0.5
Q = 1-P
sd = 0.56
n = 200
zcal = (p-P) / (sqrt(P* Q/n))
zcal = 1.6970
pvalue = 2* (1-pnorm (abs(zcal)))
pvalue = 0.08968

Thus, hypothesis is accepted.

SAD

Practical No:- 7

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	59
S.D	7	8

H_0 : 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 ("z calculated is =",z)
z calculated is =1.273682
pvalue = 2*(1-pnorm(abs(z)))
pvalue
[1] ~~0.0455026~~ 0.2027761
```

since pvalue > 0.05, we accept ~~reject~~ H_0 at 5%

Two random sample of size 1000 and 2000 are drawn from two population with the means 67.5 and 68 respectively and with the same S.D of 2.5 Test the hypothesis that the mean of two populat are equal.

H_0 :- Two population ~~mea~~ are equal

$n_1 = 1000$

$n_2 = 2000$

mx = 67.5

my = ~~58~~ 68

sdx = 2.5

sdy = 2.5

```
> z=(mx-my)/sqrt((sdx^2/n1)+(sdy^2/n2))
> z
[1] -5.163978
> pvalue 2 * $ (1-pnorm(abs(z)))
> pvalue
> 2.417564e-07
```

Since pvalue > 0.05, we accept H_0 at 5%.

In a F.Y.BSc 20% of a random sample of 400 students had defective eye sight. In SY class 15.5% of 500 students had the same defect. Is the difference of proportion is same?

H_0 :- The proportion of the population are equal

n = 400

n2 = 500

p1 = 0.2

p2 = 0.155

```
> p = (n1*p1+n2*p2)/(n1+n2)
> p
[1] 0.175
> q = 1-p
> q
[1] 0.825
> z = (p1-p2)/sqrt(p*q*(1/n1+1/n2))
> z
[1] 1.76547
> pvalue = 2*(1-pnorm(abs(z)))
[1] 0.07748487
```

Since pvalue > 0.05, we accept H_0 at 5%.

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 and 30 bad apples in the second sample. Test the hypothesis that the two boxes are equivalent in terms of numbers of bad apples.

Ho:- Two boxes are equal

```
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%.

In a MA class out of a sample of 60 mean is 63.5 inch with a S.D 2.5. In a MCom class out of 50 students mean height 69.5 inches with a S.D of 2.5 Test the hypothesis that the mean of MA and MCom class are same.

Ho:- Heights of two classes are same

```
>n1 = 60
 n2 = 60
 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%

Small sample Test

1) The flower stem are selected and the heights are found to be 63, 63, 68, 69, 71, 71, 72. Test the hypothesis that the mean height is ~~60~~ 66 cm or not at 1% loss level of significance.

H_0:- mean = ~~60~~ 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 = 5.220-09

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:-

64.66479   91.62092

sample estimates:

mean of x

68.14286

∴ p-value is less than 0.01 we reject H_0 at % level of significance.

2) Following are the weights of 10 people before and after diet. Program. Test hypothesis that the diet program is effective or not:

Before (kg): 100, 125, 95, 96, 98, 112, 115, 104, 109, 110

After (kg):- 95, 80, 95, 98, 90, 100, 110, 85, 100, 101

H_0:- The diet program is not effective.

~~Pair~~ Paired t-test

data : x and y
t = 2.6089 , df = 9 , p-value = 0.9858
alternative hypothesis:- true difference in means is less than 0.
95 percent confidence interval:-
-Inf , 18.72908
sample estimates:
mean of the differences
11

∴ p-value is greater than 0.05 we accept H_0 at 5% level of significance.

3) Two random sample are drawn 2 different popul
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 the two population mean at 5% level of significance.
H_0:- There is no difference in the population mean
> x = c(8, 10, 12, 11, 16, 15, 8, 7)
> y = c(20, 15, 18, 9, 8, 10, 11, 12)
which welch two sample t-test
data: x and y
t = -0.36247 , df = 13.837 , p-value = 0.7225
alternative hypothesis:- true difference in means is not equal to 0.
95 percent confidence interval:
-5.192719 3.692719

mean of x   mean of y
12.25   12.875

∴ p-value is greater than 0.05 we 052 accept Ho at 5% level of significance.

4) The marks before and after training program are given below

Before :- 20, 25, 32, 28, 27, 36, 35, 25

After :- 30, 35, 32, 37, 37, 40, 40, 25

Test the hypothesis training Program is effective 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)
>t.test(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 mean is greater than 0.

95 percent confidence interval:

-3.967399 Inf

sample estimates

mean of the differences

-5.75

∴ p-value > 0.05 we accept the Ho at 5% level of significance.

5> Two random sample that drawn from two normal population and 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 same variance at 5% level of significance

H_0:- The variances of two 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)
```

F test to compare two variances

data:- x and y

F = 0.70686 , num df = 8, denom df = 10, p-value = 0.63

alternative hypothesis:- true ratio of variances is not equal to 1

95 percent confidence interval:-

0.1833662 3.0360393

ratio of variances:-

0.7068567

∴ P-value > 0.05 we accept the H_0 at 5% level of significance.

6) The arithmetic mean of a sample of 100 observation is 52 if the S.D is 7. Test the hypothesis that the population mean is 55 or not at 5% level of significance.

H0:- Population mean = 55
n = 10
mx = 52
~~SD~~ mo = 55
SD = 7

```
>zcal = (mx-mo) / (sd/sqrt(n))
>zcal
[1]  4.285714
>p-value = 2 * (1-pnorm(abs(zcal)))
>p-value
[1]  1.82153e-05
```

∴ p-value < 0.05 we reject H0 at 5% level of significance.

Practical No :- 9

Chi square distribution and ANOVA Bar

1) Use the following data to test whether the cleanliness of home depends upon the child.

		Codn of Home	
		Clean	Dirty
Codn of child	Clean	70	50
	Fairly	80	20
	Dirty	35	45

H_0 :- codn of Home and child are independent

```
> x = c (70,80,35,50,20,45)
> m = 3
> n = 2
> y = matrix (x, nrow = m, ncol = n)
> y
>        [1]    [2]
  [1]    70     50
  [2]    80     20
  [3]    35     45
> pv = chisq.test (y)
> pv
        Pearson's chi-square test
data :- y
x-squared = 25.646, df = 2, p-value = 2.698e-06
```

∴ P-value is less than 0.05 we reject H_0 at 5% level of significance.

2) Table below shows a relation between the performance of mathematic and computer of CS student. 051

		Maths		
		H.G	MG	LG
Comp	HG	56	71	12
	MG	47	163	38
	LG	14	42	85

H_0:- Performance between maths and computer are independent

```
> x=c(56,47,14,71,163,42,12,38,85)
> m=3
> n=3
> y=matrix(x,nrow=n, ncol=n)
> y
     [1]   [2]   [3]
[1]  56    71    12
[2]  47    163   38
[3]  14    42    85
> p-v=chisq.test(y)
> pv
        Pearson's chi-squared test
data: y
x-squared=145.78, df=4, p-value=2.2e-16
```

∵ Pvalue < 0.05 we reject H_0 at 5% level of significance.

3> Perform ANOVA for the following Table data

variety	observations
A	50, 52
B	53, 55
C	60, 58, 57, 56
D	52, 54, 54, 55

4>

H_0:- The mean of variety A, B, C, D are equal

```
x1 = c(50,52)
x2 = c(53,55)
x3 = c(60,58,57,56)
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	F value	Pr(>F)
ind	3	71.06	23.688	11.73	0.00183
Residuals	9	13.17	2.019		

∵ P-value < 0.05 we reject H_0 at 5% level of significance.

4) Perform ANOVA for the following data:

types	observation
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

H_0:- The mean of variety A, B, C, D 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))
> name(d)
[1] "values"   "ind"
> oneway test (values~ind, data=d, var equal =T)
          one-way analysis of means
data: values and ind
F = 2.667, num df = 3, denom df = 8, p-value = 0.8839
> anova = aov (values~ind, data=d)
> summary (annova)
            Df  sum sq  mean sq  F value  Pr(>F)
ind          3    18     6.00     2.667    0.119
Residuals    8    18     2.25
```

∵ p-value > 0.05 we accept H_0 at 5% level of significance.

> x = read.csv("c:/users/administrator/deskt
Mar

> x

	STATS	cal
1	40	60
2	45	48
3	42	47
4	15	20
5	37	25
6	36	27
7	49	~~27~~ 57
8	59	58
9	20	25
10	27	27

Practical No:-10

Topic: Non-parametric Test

Q.1) Following are the amounts of sulphur oxide emmited by a factory

17 15 20 29 19 18 22 25 27 9 24
20 17 6 24 14 15 23 24 26

Apply sign test to test the hypothesis that the population median is 21.5 against the alternative that it is less than 21.5.

H_0:- Population median = 21.5
H_1 :- Population median < 21.5.

```
>x = c(17,15,20,29,19,18,22,25,27,9,24,20,17,6
       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)
> pv = 0.4119015
```

Note:- If the alternative is greater than median then pv = pbinom(sn, n, 0.5).

Q.2]. For the observations :- 12 19 31 28 43 40 55 49 70 63.

058

Apply sign test to test ~~Si~~ the population median is 25 against the alternative that it is greater than 25.

Ho :- Population median = 25

H₁ :- Population median > 25

```
>x = c(12,19,31,28,43,40,55,49,70,63)
> m = 25
>sp = length (x[x>m])
>sn = length (x[x<m])
>n = sp+sn
>pv = pbinom(sn,n,0.5)
> pv = 0.0546875
```

Q3]. For the following data:- 60 65 63 89 61 71 58 51 48 66

Test the hypothesis using ~~willkoxen~~ wilcoxon signedrank test. For test the hypothesis median is 60 against the alternative it is greater than 60.

Ho :- The median = 60

H₁ :- Median > 60.

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

Wilcoxon signed rank test with continuously correction

data: x

V = 29 , p-value = 0.2386

alternative hypothesis: true location is greater than 60.

Note:- If the alternative is less we have to wri
wilcox.test (x, alter = "less", mu = _)

If the alternative is not equal to
wilcox.test (x, alter = " ~~t.si~~ "two.sided", mu = _)

Q.4]. Using wilcoxon test, test the hypothesis that the median is 12 against the hypothesis <12

12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20

H_0:- The median = 12
H_1:- Median < 12

```
> x = c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20)
> wilcox.test(x, alter = "less", mu = 12)
        Wilcoxon signed rank test continuity correction
data: x
V = 25, p-value = 0.2521
alternative hypothesis: true location is less than 12
```

AY
11.7.20