Scanned with CamScanner

Pratical - 1

Basics of R. software

→ R is software for data analysics and statictical comparing
→ This software is used for effective data handling and output storage is available.
→ Its capable of graphical diplayed.
→ Its 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] 9.262826
> 42 + abs(-10) + 7^2 + 3*9
[1] 128
```

$(2^2 + \sqrt{25} + 35)$

$(2 \times 5 \times 3 + 62 \div 5 + \sqrt{49})$

$(\sqrt{76 + 4 \times 2 + 9 \div 5}$

$(42 + |-10| + 7^2 + 3 \times 9)$

```
> x = 20
> y = 30
> x+y
[1] 50
> x^2 + y^2
[1] 1300
> abs (x-y)
[1] 10
> sqrt (y^3 - x^3)
[1] 137.8405
```

$(x^2 + y^2)$

$(|x-y|)$

$(\sqrt{y^3 - x^3}$

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

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

> c(2,3,5,9) + c(-2,-3,-5,-4)
[1]  0  0  0 3
> c (2,3 5,9) *  c(8,9)
[1]  16  27  40  63
```

Q2 Find the sum product Maximum, Minimum value
5, 8, 6, 7, 9, 10, 15, 5

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

> x - Matric Cnrow = 4, ncal = 2, data = c(1,2,3,4,5,6,7

> x

	[,1]	[,2]
[1,]	1	5
[2,]	2	6
[3,]	3	7
[4,]	4	8

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

> x <- Matrix Cnrow = 3, ncal = 3, dato = c(1,2,3,4,6,5,7,8,9))

> x

	[,1]	[,2]	[,3]
[1,]	1	4	7
[2,]	2	5	8
[3,]	3	6	9

> y <- Matrix (nrow = 3, ncol = 3, dato = c(2,-2,10,4,8,6,10,-11,12)

> y

	[,1]	[,2]	[,3]
[1,]	2	4	10
[2,]	-2	8	-11
[3,]	10	6	12

> x+y

3	8	17
0	13	-3
13	12	21

```
> x * y
        2     16     70
       -4     40    -88
       30     36    103
> 2 * x + 3 * y
        8     20     44
       -3     54    -17
       36     30     54

> x = c(2,4,6,1,3,5,7,18,16,14,17,19,3,3,2,5,0,15,7,10,
> length (x)
  [1] 25
> a = table (x)
> transform (a)
       x        freq
       0
       1
       2
       3
       4
       5
       6
       7
       8
       9
       10
       12
       14
       15
       16
       17
```

Imarks = seq (0, 20, 5)
G = cut [x, breaks, right = FALSE]
c = table (b)
transform (c)

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

Pratical no. 2

1) Can the following be P.D.F

$$f(x) = \begin{cases} 2-m & 1 \le x \le 2 \\ 0 & \end{cases}$$

Sol^n :- $\int f(x)\,dx = 1$

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

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

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

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

$= 0.5$

$\neq 1$ $\quad\quad$ $\therefore$ It is not a P.D.f

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

Sol^n :- $= \int f(x)\,dx = 1$

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

$= 3\left[\frac{x^3}{3}\right]_0^1 \quad = [1-0] \quad = 1$

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

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

sol^n :-

$$\int f(x)\,dx = 1$$

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

$$= \int_0^2 \left(3x/2 - \frac{3x^2}{4}\right)dx$$

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

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

$$= 3/2\,(2-0) - 3/12\,(8-0)$$

$$= 6/2 - 24/12$$

$$= 3/1 - 3/1 \quad = 1/1$$

$$= 1$$

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

2] Can the following be p.d.f

①

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

$\because$ One of the value is negative $\therefore$ It is not p.d.f

②

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

It is P.M.F

$\Sigma P(x) = 0.1 + 0.3 + 0.2 + 0.2 + 0.1 + 0.1 = 1$

Scanned with CamScanner

3) find $P(x \le 2)$ $P(2 \le x < 4)$ $P(\text{at least } 4)$ $P(3 < x < 6)$

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

$P(x \le 2) = P(x=0) + P(x=1) + P(x=2)$

$= 0.1 + 0.2 + 0.1$

$= 0.4$

$P(2 \le x < 4) = P(x=2) + P(x=3)$

$= 0.2 + 0.2$

$= 0.4$

$P(\text{at least } 4) = P(x=5) + P(x=6)$

$= 0.2 + 0.1$

$= 0.3$

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

$= 0.1 + 0.2$

$= 0.3$

4) find c.d.f

(1)

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

$P(x) = 0 \qquad x < 0$

$= 0.1 \qquad 0 \le x < 1$

$= 0.2 \qquad 1 \le x < 2$

$= 0.4 \qquad 2 \le x < 3$

$= 0.6 \qquad 3 \le x < 4$

$= 0.7$

$= 0.9 \qquad 5 \leq x < 6$

$= 1 \qquad x \geq 6$

X	10	12	14	16	1.8
P(x)	0.2	0.35	0.15	0.2	0.1

$P(x) = 0 \qquad x < 10$

$= 0.2 \qquad 10 \leq x < 12$

$= 0.55 \qquad 12 \leq x < 14$

$= 0.70 \qquad 14 \leq x < 16$

$= 0.9 \qquad 16 \leq x < 18$

$= 1 \qquad 18 \leq x$

Pratical - 3

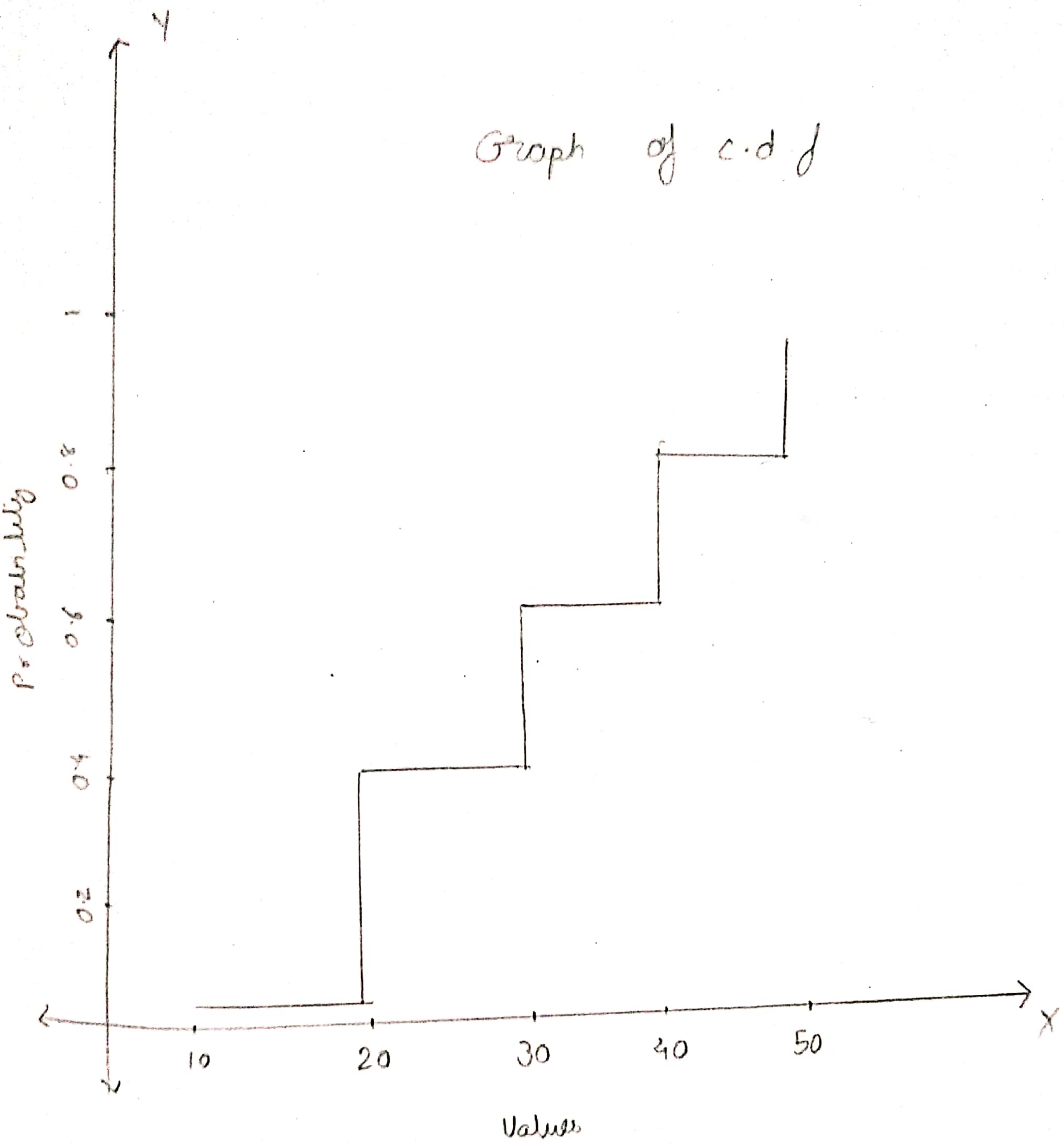
Probability of Binomial Distribution

Find the c.d.f of following P.M.F and draw graph

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

solⁿ :-

```
x = c (10, 20, 30 ,40, 50)
prob = c(0.15, 0.25 , 0.3 , 0.2, 0.1)
 Cumsum (prob)
[1]  0.15 , 0.40  0.70  0.90  1.00
plot (x , cumsum (prob) , x lab = "values" , ylab = "probability"
      main = " graph of c.d.f ", "s" )
```


Y
Graph of c.d.f
1
0.8
0.6
0.4
0.2
Probability
10
20
30
40
50
X
Values

(ii) > c = pbinom (4, 12, 1/5)
> c
[1] 0.927445

3) There are 10 member in a committee the probability of any member attending a meeting is 0.9. Find the probability
(i) 7 member attend (ii) atleast 5 member
(iii) atmost 6 member attends

solⁿ :-

$n = 10$ $p = 0.9$ $q = 0.1$

(i) x = Total no. of member attendance

$X \sim B(n, p)$

```
> x=7
> c= dbinom (7,10,0.9)
> c
[1] 0.05739563
```

(ii)

```
> x=5
> c= 1- pbinom (4,10,09)
> c
[1] 0.99 ~~83651~~ 9853
```

```
(iii) > x=6
> c = pbinom (6,10,0.9
> c
[1] 0.0127952
```

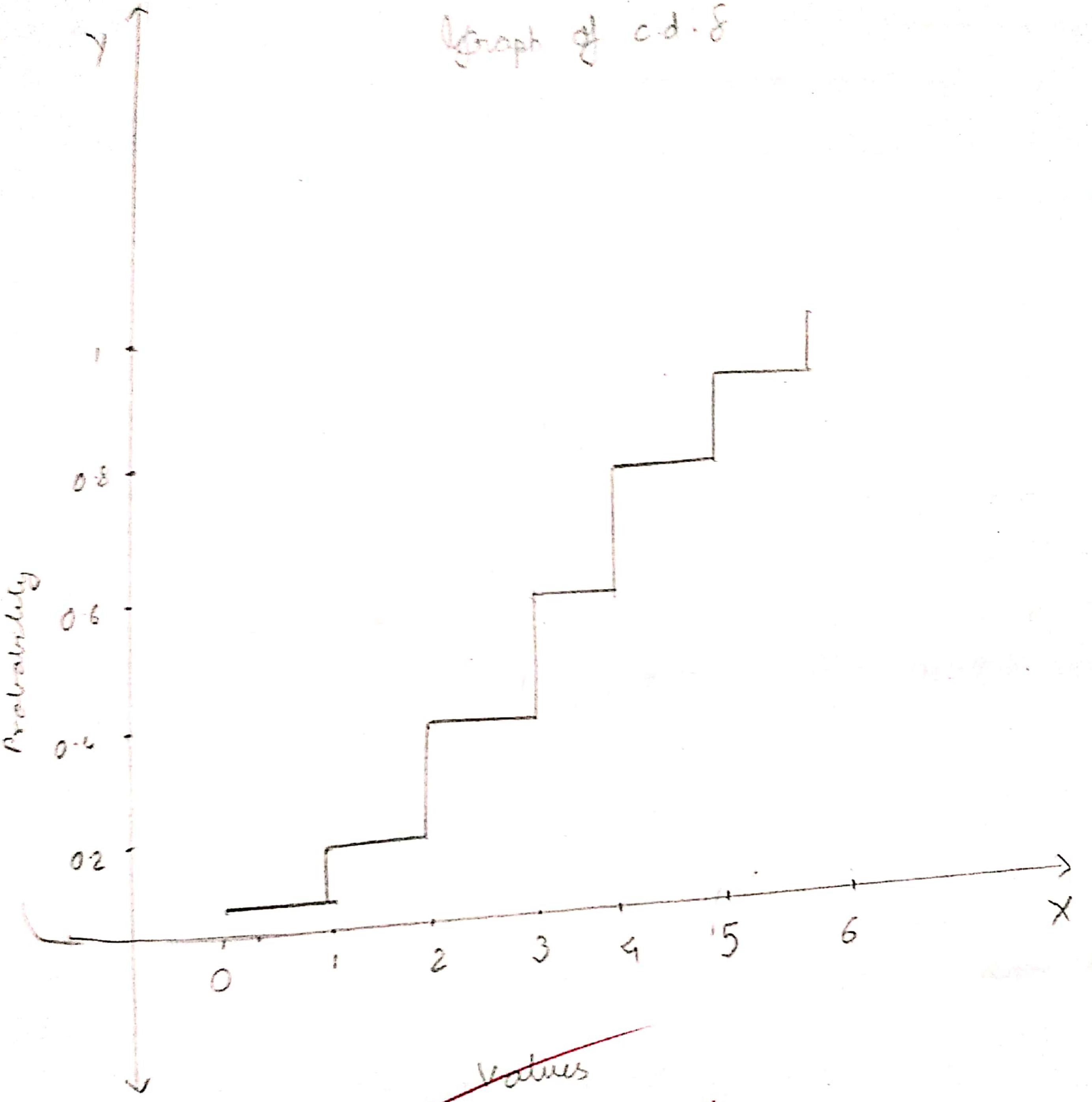
4. Find the c.d.f and draw the graph.

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

Sol^n :-

```
> 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   1.00
> plot (x, cumsum (prob), x lab = "values", y lab = "probability"
        main = " graph of c.d.f", "s")
```

Graph of c.d.f



~~Values~~

AM
11.12.16

Pratical - 21

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 $P=0.9$

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] $P(5 < x < 7)$

4] The probability of sales man make a sale to customer is 0.15. find the probability
i] No sale for 10 customers
ii] More than 3 sale in 20 customer.

5] The student write 5 m.c.q. Each question has 4 option out of which only one is correct. find the probability for atleast 3 correct ans.

6] X follows B.D. with $n=10$, $p=0.4$. plot the graph of p.m.f and c.d.f.

Note :-

i] $P(X=x) = dbinom(x, n, p)$
$P(X \le x) = pbinom(x, n, p)$
$P(X > x) = 1 - pbinom(x, n, p)$
To find the value of x for which the probability is P, the command is $qbinom(P, n, p)$

solution :-

1) $n = 5$, $p = 0.1$
```
> dbinom(0:5, 5, 0.1)
[1] 0.59049   0.32805   0.07290   0.00810   0.00045 0.00001
```

2] $n = 100$
$p = 0.1$
$x = 10$
```
> dbinom(10, 100, 0.1)
[1] 0.1318653
```

3] $n = 12$
$p = 0.25$
$x = 5$

①
```
> dbinom(5, 12, 0.25)
[1] 0.1032414
```

②
```
> pbinom(5, 12, 0.25)
[1] 0.9455978
```

Scanned with CamScanner

(4) > dbinom (6, 12, 0.25)
[1] 0.04014945

(3) > 1-pbinom (7, 12, 0.25)
[1] 0.00278151

$P(X > 7) = 1 -$ p~~binom~~$(X \leq$
$P(X \geq 7) = 1 -$ p~~binom~~ $(X \geq 6)$

4] n = 10 , p = 0.15 , x = 0

> dbinom (0, 10, 0.15)
[1] 0.1968744

n = 20 , p = 0.15
P (X > 3) = 1 - P(X ≤ 3) = 1 - pbinom (3 , 20, 0.15)

> 1- pbinom(3 , 20, 0.15)
[1] 0.3522748

5] n = 5 x = 2 P = 0.25

P(X ≥ 3) = 1 - P (X ≤ 2) = 1 - pbinom (2, 5, 0.25)

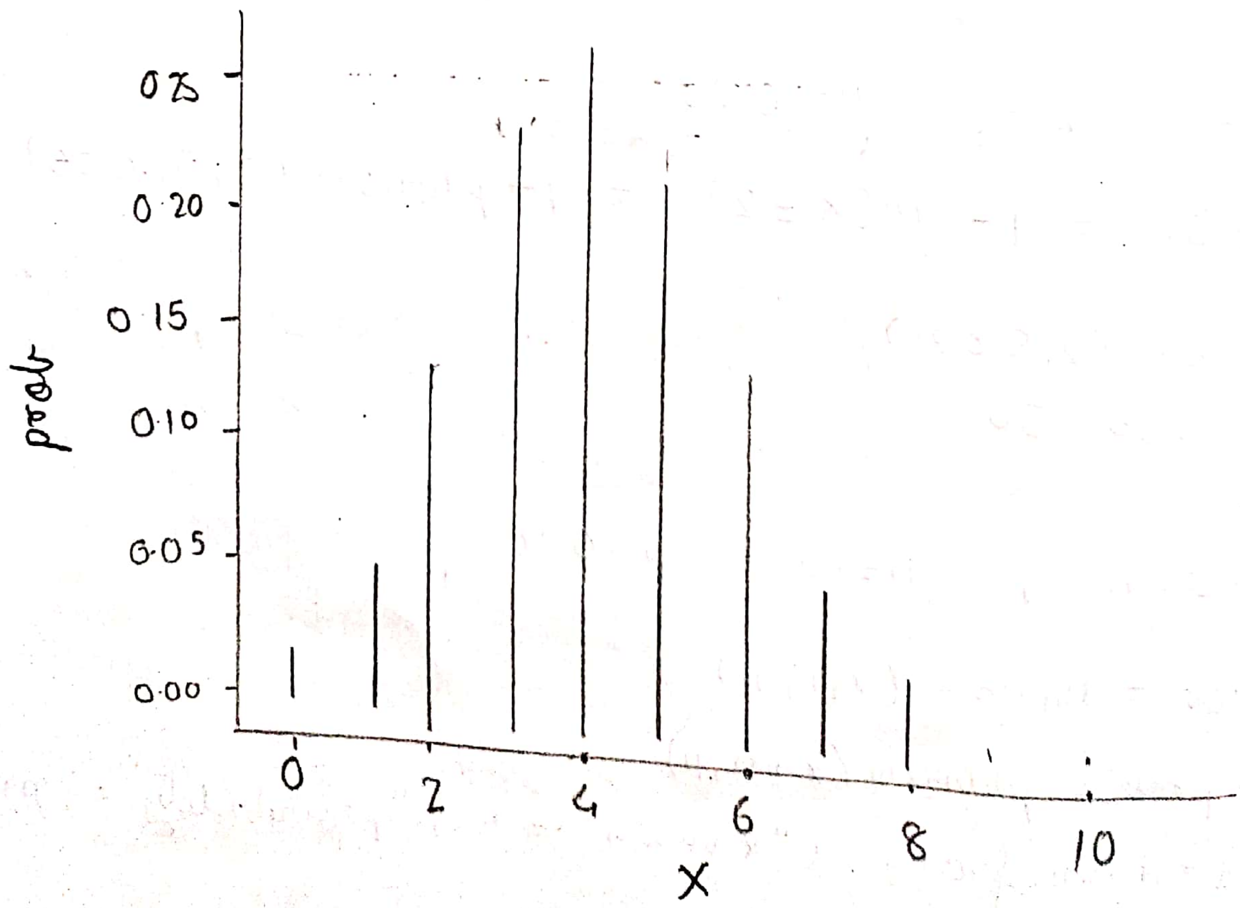
> 1 - pbinom (2, 5, 0.25)
[1] 0.1035156.

6] x = 0 : n , n = 10 p = 0.4

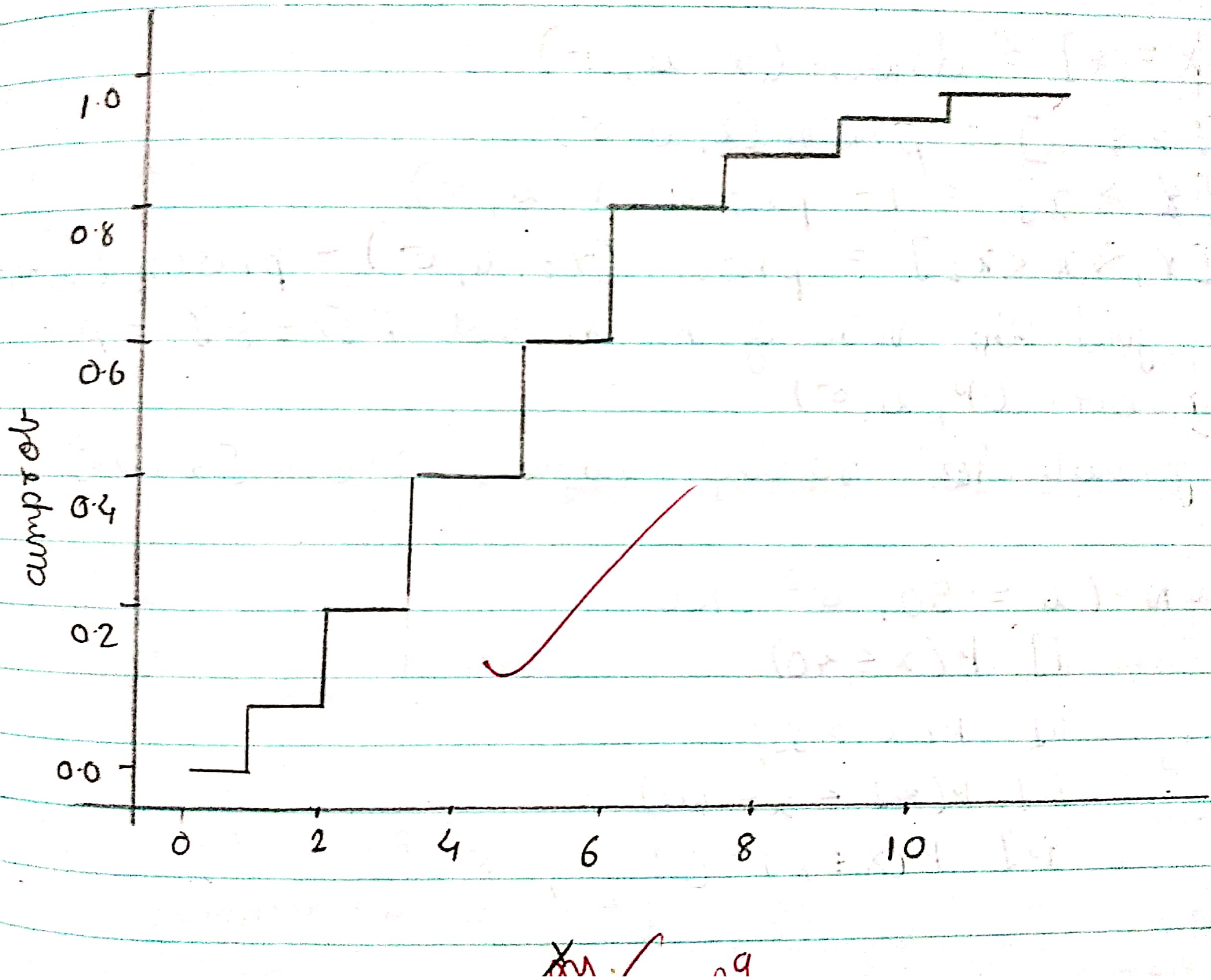
> prob = dbinom (x, n, p)
> cumprob = pbinom (x , n, p)
> d = data.frame ("x values" = x , "probability" = prob)
> print (d)

> plot (x, prob, "h")

	x values	probability
1	0	0.0060466176
2	1	0.0403107840
3	2	0.1209323520
4	3	0.2149908480
5	4	0.2508226560
6	5	0.2006581248
7	6	0.1114767360
8	7	0.0424673280
9	8	0.0106168320
10	9	0.0015728640
11	10	0.0001048576



> plot (x, cumprob, "s")



Pratical - 5

Aim : Normal Distribution.

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

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

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 ,0 ,10)

> cat (" p(x <=k) =0.7, k=" ,d)

p (x<=k) = 0.7 k = 5.2440

2] $X \sim N(\mu = 100, \sigma^2 = 36)$

find :- i) $P(X \le 110)$

ii] $P(X \le 95)$

iii] $P(X > 115)$

iv] $P(95 \le x \le 105)$

sol :.

i] $\mu = 100$

$\sigma^2 = 36$

$\therefore \sigma = 6$

> 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 (105 ,100, 6) - pnorm (95 ,100, 6)

> cat ("P (95 ≤ x ≤ 105) =", d)

P (95 ≤ x ≤ 105) = 0.953432

5] $x1e$ = 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 numbers from normal distribution with mean = 60 & sd = 5 also calculate sample mean, variance, median & standard deviation

sol^n :-

```
> x = rnorm( 10, 60, 5)
> a.m = mean (x)
> a.m
  [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
```

Drau

solut

>

>

>

Scanned with CamScanner

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")
```

Pratical - 6 : z Distribution

1] Test the hypothesis (Ho) : $\mu = 10$ 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 significance.

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
> z cal = (mx - mo) / (sd / sqrt(n))
> cat ("z cal is =", z cal)
z cal is = 1.77778 > pvalue = 2*(1-pnorm(abs(zcal
> pvalue
[1] 0.07544036
```

since 0.07544036 is more than 0.05 we will accept the Ho

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

42

```
> mo = 75
> mx = 80
> sd = 3
> n = 100
> zcal = (mx - mo) / (sd / sqrt(n))
> cat (" zcal is = ", zcal)
zcal is = -66.6667
> pvalue = 2* (1-pnorm(abs(zcal)))
> pvalue
[1] 0
>
```

Value accepted

3) Test the hypothesis (H_0) : $= 25$ against (H_1) : $\mu \neq 25$ at 5% level of significance the following sample is 30 selected.

20	24	27	30	46	35	26	~~46~~	27	10	20
22	30	37	35	21	~~22~~	23	24	25	26	27
14	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, 14, 28, 29, 30, 39, 27, 15, 22, 20, 18)
> mx = mean (x)
> mx
[1] 26.06667
```

```
> n = length (x)
> n
  [1] 30
> variance = (n-1) * var(x)/n
> variance
  [1] 52.99556
> sd = sqrt (variance)
> sd
  [1] 7.279805
> m0 = 25
> zcal = (mx - m0) / (sd / sqrt (n))
> cat ("zcal is = ", zcal)
zcal is 0.80254547
> pvalue = 2 * (1- pnorm (abs (zcal)))
> pvalue
  [1] 0.422375
```

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.

Q. 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 significance

```
> 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 accepted.

Pratical - 7 Large Sample Test.

A study of noise level in two hospital is calculated below. Test the hypothesis that the noise level in two hospital are same ~~as~~ is or not.

	Hos A	Hos B
No. of sample	84	34
abs	8	
meas	61	S
sd	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 ("x calculated is =", z)
x calculated is = 1.273682
> pvalue = 2 * (1- pnorm (abs (z)))
> pvalue
[1] 0.2027761
```

∵ pvalue > 0.05, we accept H_0 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 that the mean of two population are equal

H_0 of the population are same.

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

∵ pvalue > 0.5 we accept H_0 at 5% l.o.s.

3] In fy B.S.C 20% of a random sample of 400 students had defective eye sight. In S% class 15.5% of 500 student have the same effect; is the difference of the proportion is same?

H_0 - The proportion of 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/n1))

> z

[1] 1.76547

> cat ("z calculated is =", z)

z calculated is =1.76547

> pvalue = 2* (1-pnorm (abs (z)))

> pvalue

[1] 0.07748487

∵ pvalue > 0.05 we accept H_0 at 5% LOS

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

H_0: 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

∵ pvalue > 0.05, we accept H_0 at 5% LOS

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 of ~~at~~ 50 students mean height is 69.5 incles with a s.d = 2.5 Test the hypothesis that the mean of M.A & M.com class are same.

Ho : Height 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
```

∴ p value < 0.05 , we reject Ho at lo.f.

Praticall - 8. Small sample Test.

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

H_0 : mean = 66 cms

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

One sample t-test

data : x

t = 47.94 , df = 6 , pvalue = 5.22e -09

Alternative hypothesis : true mean is not equal to 0

95 percent confidence interval :

64.66479 71.62092

sample estimates :

mean at x

68.14286

∴ pvalue < 0.01 is rejected on H_0 at 1% lof.

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

03

> x = c (8, 10, 12, 11, 16, 15, 18, 7)
> y = c (20, 15, 18, 9, 8, 10, 11, 12)
> t. test (x, y)

. Two sample t-test

data x & y
t = 0.36 247, df = 13.837 . p-value = 0.7225
alter native hypothesis true difference in mean is not equal ~ 5.192719 8.69 2719
sample estimates
mean of x mean of y
12.175 12.875
p-value < 0.01 is accepted in H_0 on 1% los

Q3 :: Following are the weights of 10 peaple
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 per cent confidence interval
~ Inf 17.89635
sample estimats :: mean of the differences 10
p-value < 0.01 is accepted in H_0 on 1% level of significance

Q3 :- 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 effective or not

H_0 - 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 = -3.3869, df = 7 . p.value = 0.9942

alternative hypothesis true difference in means is greater than 0

95 percent confidence interval.

- 8.967399   Inf

sample estimates

mean of the difference

-6.75

p-value < 0.01 is accepted in H_0 at % los

5. Two random sample were drawn from two normals populations & 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% los

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 = 0.6

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

p value < 0.7068567 is accepted in H_0 at 1% los

Q6 The A.P of sample of 100 observation is 52 if S.D is 7 test the hypothesis that the population mean 55 or not at 5% los.

H_0 population mean = 55

```
> n = 100
> mx = 52
> m0 = 55
> sd = 7
> zcal = (mx - m0) / (sd / sqrt (n))
> pvalue = 2 * (1- pnorm (abs (zcal)))
> pvalue
[1] 1.82153e-05
```

P value < 0.01 is accepted ~~at~~ in H_0 at 1% level of significance.

Pratical - 9   chi square distribution & ANOVA

1. 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 and 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 . pvalue = 2.698e-0
```

∵ p value is less than 0.05 we reject H_0 at 5% los

Q2: Table below shows the relation between performance of mathematics & computer of C.S students.

		Maths		
		H.G	M.G	L.G
Computer	H.G	56	71	12
	M.G	47	163	38
	L.G	14	42	85

H_0 performance of mathematics & C.S are independent.

```
> a = c (56, 47, 14, 17, 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)
> pv
        Person's chi-squared test
data - z
x - squared = 145.78, df = 4, p value < 2.2e -16
```

∵ p-value < 2.2e -16 is less than 0.05, we reject H_0 at 5% los.

Q3 :- 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 A B C D 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"
> one way. test (values ~ ind, data - d, var. equal =T)

        one way analysis of means
data    values and ind
F = 11.735, num df = 3, denom df = 0, p value = 0.00183
> anova = aov (values ~ind, data =d)
> summary (anova)
```

	DF	sum Sq	Mean Sq	F values	Pr(>F)
ind	3	71.06	26.688	11.73	0.00183 ''
Residuals	9	18.17	2.019		

∵ p-values is less than 0.05, we reject the H_0 value at 5% los.

04 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 value is more than 0.05 we accept the H_0 value at level of significance

Write a R command to open ~~an~~ an excel file in R software.

```
> x = read.csv ("s:/users/admin/Desktop/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

Non - Parametric Test.

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, 28
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 media 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
```

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

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 (o [o > m])
> sn = length (o [o < m])
> n = sp + sn
> n
[1] 16
> pv = pbinom (sn, n, 0.5)
> pv
0.0546845
```

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

Q3 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
> mv = 60
```

wilcoxn signed rank test with continuity correction.

data : x

v = 29, p-value = 0.2386

alternative hypothesis true location is greater than 60

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

Note : If the alternative is less, we have To write wilcoxin test.

(x, alter = " less , mu=60)

If the alternative is not equal to we have to write wilcoxin Test (x, area = " two sided ", mu = cal)

Q. Test the hypothesis median is 12 against the alternative that it is less than 12 using wilcoxin 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

wil cox. test (x, alter = "less", mu = 12)

wil coxon signed rank test with continuity correction.

data : x

v = 25 , p value = 0.2521

alternative hypothesis location is less than 12

∵ p value > 0.05 . we accept the H_0 at 5% los.

Scanned with CamScanner