PERFORMANCE

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

Exam Seat No. ______

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 fybsc.cs Roll No. 1757 Year 2019-20

This is to certify that the work entered in this journal is the work of Mst. / Ms. Priyanka Kumari Saw who has worked for the year 2019-20 in the Computer Laboratory.

Teacher In-Charge

Head of Department

Date : ______

Examiner

★★ INDEX ★★

PRACTICAL - 01

AIM : Basics of R software

1) R is a software for stastical analysis and data computing.

2) It is an effective data handling software and outcome storage is possible.

3) It is capable of graphical display
4) It is a free software.

Q1 Solve the followings

1) $4+6+8\div2-5$

```
> 4+6+8/2-5
[1] 9
```

2) $2^2+|-3|+\sqrt{45}$

```
> 2^2+abs(-3)+sqrt(45)
[1] 13.7082
```

3) $5^3+7\times5\times8+46/5$

```
> 5^3+7*5*8+46/5
[1] 414.2
```

4) $\sqrt{4^2+5\times3+7/6}$

```
> sqrt(4^2+5*3+7/6)
[1] 5.671567
```

Scanned by TapScanner

5) round off

46÷7+9*8

>round(46/7+9*8)

[1] 79

Q2 >c(2,3,5,7)*2

[1] 4 6 10 14

>c(2,3,5,7)*c(2,3,6,2)

[1] 4 9 30 14

>c(2,3,5,7)^2

[1] 4 9 25 49

>c(6,2,7,5)/c(4,5)

[1] 1.50 0.40 1.75 1.00

>c(2,3,5,7)*c(2,3)

[1] 4 9 10 21

>c(1,6,2,3)*c(-2,-3,-4,-1)

[1] -2 -18 -8 -3

>c(4,6,8,9,4,5)^c(1,2,3)

[1] 4 36 512 9 16 2125

Q3 >x=20 >y=30 >z=2

>x^2+y^3+z

[1] 27402

>sqrt(x^2+y)

[1] 20.73644

>x^z+y^z

[1] 1300

Q4 >x<-matrix(nrow=4,ncol=2,data=c(1,2,3,4,5,6,7,8))

>x

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

Q5 Find x+y and 2x+3y where $x = \begin{bmatrix} 4 & -2 & 6 \\ 7 & 0 & 7 \\ 9 & -5 & 3 \end{bmatrix}$

$y = \begin{bmatrix} 1 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 5 \end{bmatrix}$

```
>x<-matrix(nrow=3, ncol3, data=c(4,7,9,-2,0,-5,6,7,3))
```

> x

	[,1]	[,2]	[,3]
[1,]	4	-2	6
[2,]	7	0	7
[3,]	9	-5	3

```
>y<-matrix(nrow=3, ncol=3, data=c(10,12,15,-5,-4,
 -6,7,9,5))
```

> y

	[,1]	[,2]	[,3]
[1,]	10	-5	7
[2,]	12	-4	9
[3,]	15	6	5

> x+y

	[,1]	[,2]	[,3]
[1,]	14	-7	13
[2,]	19	-4	16
[3,]	24	-11	8

> 2*x+3*y

	[,1]	[,2]	[,3]
[1,]	38	-19	33
[2,]	50	-12	41
[3,]	63	-28	21

Q6 Marks of stastics of CS Batch A

x = c(60, 20, 35, 24, 46, 56, 55, 45, 27, 22,
47, 58, 54, 40, 50, 32, 36, 29, 35, 39)

```
> x = c(data)
> breaks = seq(20, 60, 5)
> a = (x, breaks, right = FALSE)
> b = table(a)
> c = transform(b)
> c
        a        Freq
1    [20,25)      3
2    [25,30)      2
3    [30,35)      1
4    [35,40)      4
5    [40,45)      1
6    [45,50]      3
7    [50,55)      2
8    [55,60)      4
```

PRACTICAL-02

TOPIC: Probability distribution

1) Check whether the followings are p.m.f or not.

x	$P(x)$
0	0.1
1	0.2
2	-0.5
3	0.4
4	0.3
5	0.5

If the given data is p.m.f then

$\sum P(x) = 1$

$\therefore P(0) + P(1) + P(2) + P(3) + P(4) + P(5) = P(x)$

$= 0.1 + 0.2 - 0.5 + 0.4 + 0.3 + 0.5$

$= 1.0$

$\because P(2) = -0.5$, it can't be a

$\therefore$ probability mass function

$\because P(x) \geq 0 \ \forall, x$

2)

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

The condition for p.m.f is $\sum P(x) = 1$

So, 42

$\sum P(x) = P(1) + P(2) + P(3) + P(4) + P(5)$

$= 0.2 + 0.2 + 0.3 + 0.2 + 0.2$

$= 1.1$

$\therefore$ The given data is not a pmf because the $P(x) \neq 1$

3)

x	$P(x)$
10	0.2
20	0.2
30	0.35
40	0.15
50	0.1

The condition for p.m.f is

1) $P(x) \geq 0 \ \forall x$ satisfy

$\sum P(x) = P(10) + P(20) + P(30) + P(40) + P(50)$

$= 0.2 + 0.2 + 0.35 + 0.15 + 0.1$

$= 1$

$\therefore$ The given data is p.m.f

Code:

```
> Prob = c(0.2, 0.2, 0.35, 0.15, 0.1)
> sum(Prob)
[1] 1
```

Q2 Find the c.d.f for the following p.m.f and sketch the graph.

x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1

$F(x) = 0$ $x < 10$

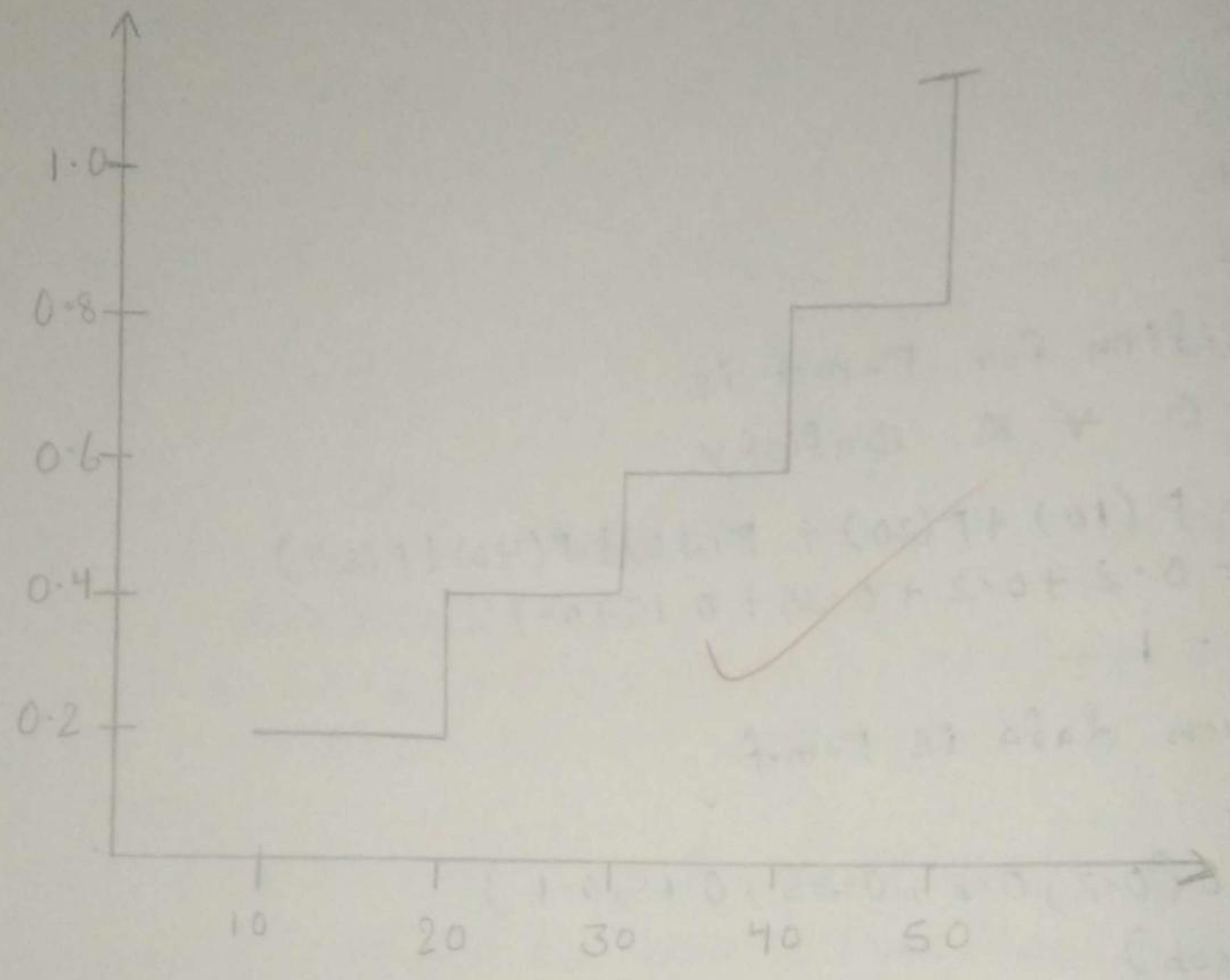
0·2 $10 \le x < 20$

0·4 $20 \le x < 30$

0·75 $30 \le x < 40$

0·90 $40 \le x < 50$

1·0 $x \ge 50$



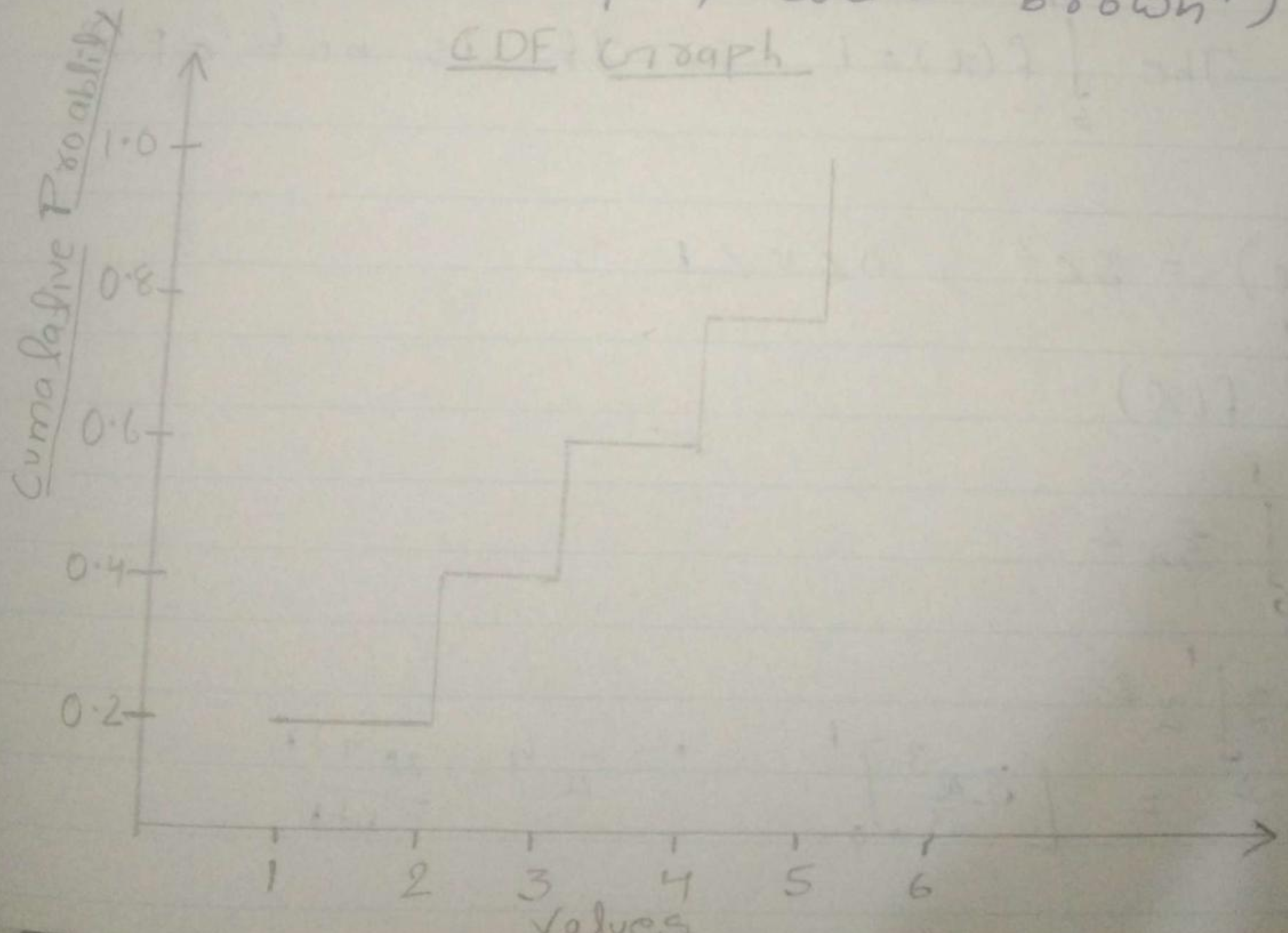

```
> X = c(10,20,30,40,50)
> plot(x, cumsum(prob), "s")
```

Q2 Find

x	1	2	3	4	5	6
$P(x)$	0·15	0·25	0·1	0·2	0·2	0·1

$$
F(x) = \begin{cases} 0 & x < 1 \\ 0.15 & 1 \le x < 2 \\ 0.40 & 2 \le x < 3 \\ 0.50 & 3 \le x < 4 \\ 0.70 & 4 \le x < 5 \\ 0.90 & 5 \le x < 6 \\ 1.00 & x \ge 6 \end{cases}
$$

```
> Prob = c(0.15, 0.25, 0.1, 0.2, 0.2, 0.1)
> sum(prob)
[1] 1
> cumsum(prob)
[1]  0.15, 0.40, 0.50, 0.10, 0.90, 1.00
> x = c(1, 2, 3, 4, 5, 6)
> Plot(x, cumsum(prob), "s", xlab=("value",
ylab="Cumulative Probability",
main="CDF Graph, "col = "brown")
```



Q3 Check that whether the following is p.d.f or not

(i) $f(x) = 3 - 2x \; ; \; 0 \leq x \leq 1$

(ii) $f(x) = 3x^2 \; ; \; 0 < x < 1$

(i) $f(x) = 3 - 2x$

$\rightarrow \int_0^1 f(x)$

$= \int_0^1 (3-2x)\,dx$

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

$= \left[3x - x^2\right]_0^1 = 2$

$\because$ The $\int_0^1 f(x) = 1 \quad \therefore$ It is not a Pdf

2) $f(x) = 3x^2 \; ; \; 0 < x < 1$

$\int_0^1 f(x)$

$= \int_0^1 3x^2$

$= 3\int_0^1 x^2$

$= \left[3\frac{x^3}{3}\right]_0^1 \quad \because x^n = \frac{x^{n+1}}{n+1}$

$= x^3$

$= 1$

The $\int_0^1 f(x) = 1$ $\therefore$ It is a pdf.

PRACTICAL-03

TOPIC: Binomial distribution

\# $P(X = x) = dbinom(x, n, P)$
\# $P(X \le x) = Pbinom(x, n, P)$
\# $P(X > x) = 1 - pbinom(x, n, P)$
\# If x is Unknown
$P_1 = P(X \le x) = qbinom(P_1, n, P)$

1) Find the Probability of exactly 10 success in hundred trials with $P = 0.1$.

2) Suppose there are 12 mcq, Each question has 5 options out of which 1 is correct. Find the Probability of having exactly 4 correct answers.
ii) Atmost 4 correct answers
iii) More than 5 correct answers

3) Find the Complete distribution when $n=5$ and $P = 0.1$.

4) $n=12$, $P = 0.25$, find the following Probabilities.

i) $P(X=5)$ iii) $P(X>7)$
ii) $P(X \le 5)$ iv) $P(5 < X < 7)$

1 > x = dbinom (10, 100, 0.1)
> x
[1] 0.1318653

Q: 2/ i) > dbinom (4, 12, 0.2)
[1] 0.1328756
ii) > pbinom (4, 12, 0.2)
[1] 0.927445
iii) > 1-pbinom (5, 15, 0.2)
[1] 0.01940528

3/ dbinom (0:5, 5, 0.1)
0 - 0.59049
1 - 0.32805
2 - 0.07290
3 - 0.00810
4 - 0.00045
5 - 0.00001

4 1) > dbinom (5, 12, 0.25)
[1] 0.1032414
2) > pbinom (5, 12, 0.25)
[1] 0.9455978
3) > 1 - pbinom (7, 12, 0.25)
[1] 0.00278151
4) > dbinom (6, 12, 0.25)
[1] 0.04014943

5

Q: 5) The Probability of a Salesman making a sale to customer 0.15 find the Probability of
1) No sales out of 10 customer
11) More than 3 sales out of 20 customer

6) A sales man has 20% Probability of making a sale to customer out of 30 customers what minimum number of sales he can make with 88% of probability.

7) X follows binomial distribution with $n=10$, $p=0.3$ plot the graph of p.m.f and c.d.f.

Scanned by TapScanner

5// >dbinom(0, 10, 0.15)
[1] 0.1968744
>1-pbinom(3,20,0.15)
[1] 0.3522748

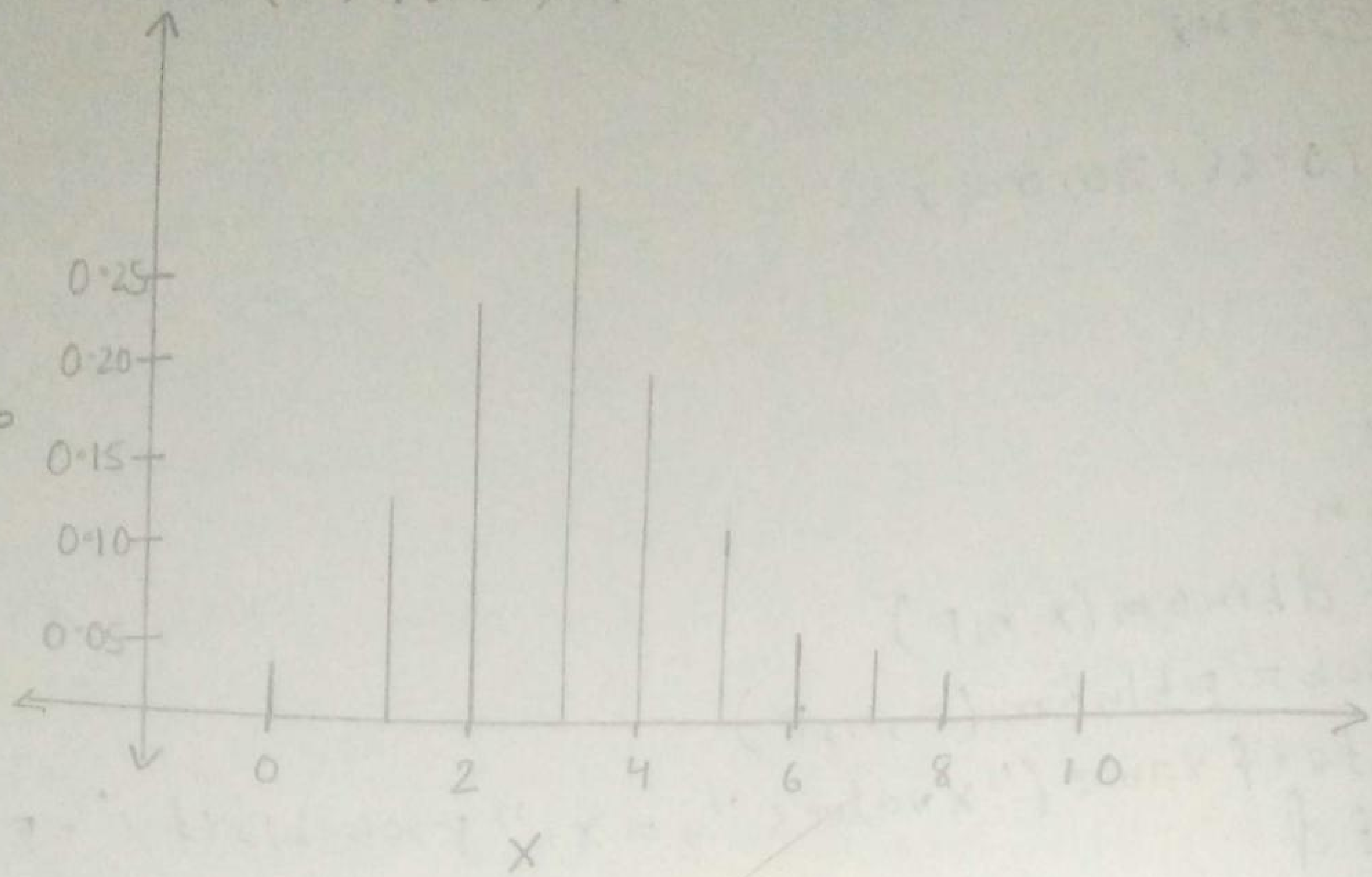
6// qbinom(0.88, 30, 0.2)
[1] 9

7// >n = 10
>p = 0.3
>x = 0:n
>prob = dbinom(x,n,p)
>cumprob = pbinom(x,n,p)
>d=data.frame("xvalues", = x, " probability" = prob)
>print d

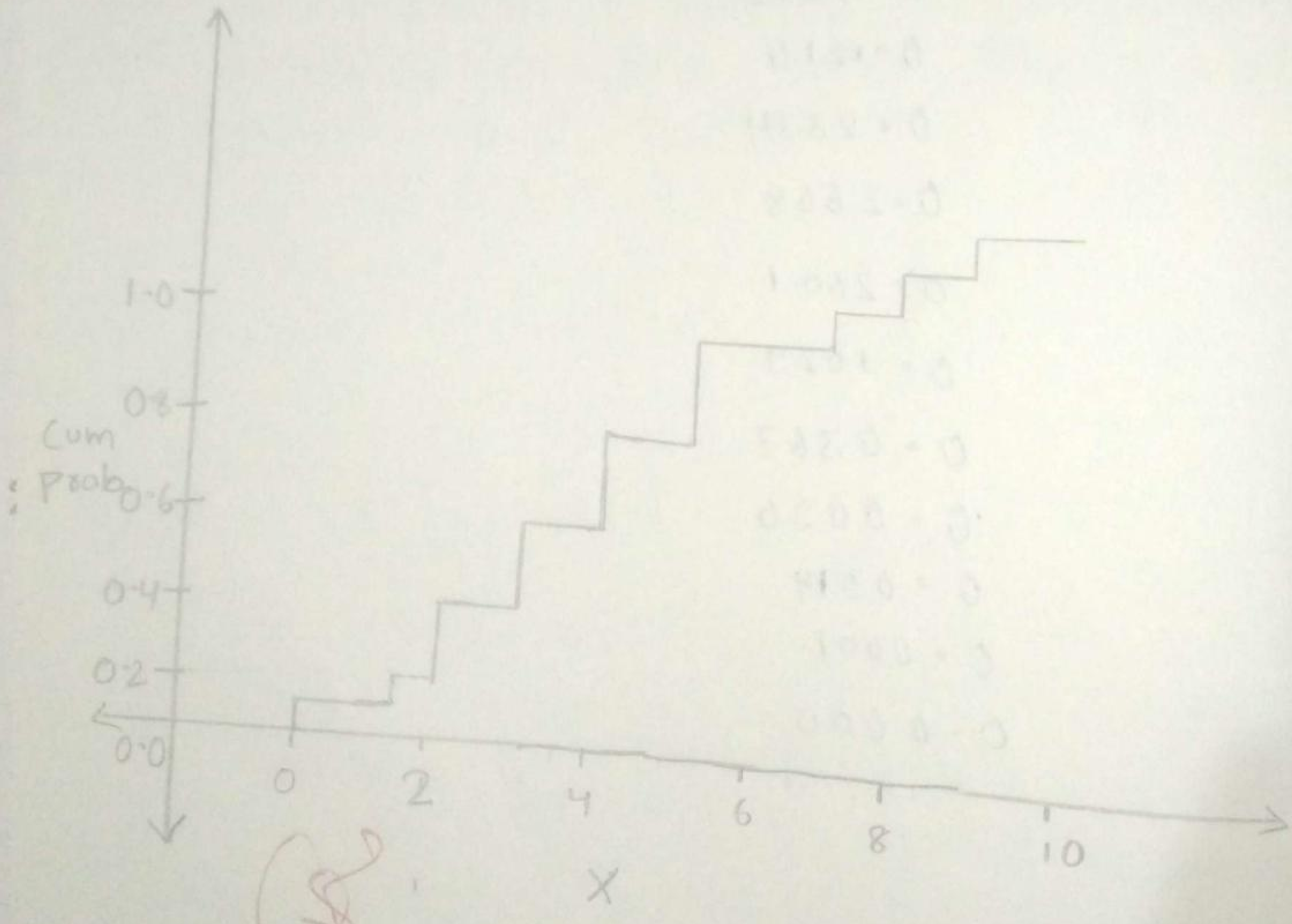
,	xvalues	probability
1	0	0.0282
2	1	0.1210
3	2	0.2334
4	3	0.2668
5	4	0.2001
6	5	0.1029
7	6	0.0367
8	7	0.0090
9	8	0.0014
10	9	0.0001
11	10	0.0000

Scanned by TapScanner

> plot (x, Prob, "h")



> plot (x, cumprob, "s")



PRACTICAL-04

AIM: Normal Distribution

(i) $P(X = x) = \text{dnorm}(x, \mu, \sigma)$
(ii) $P(X \leq x) = \text{pnorm}(x, \mu, \sigma)$
(iii) $P(X > x) = 1 - \text{pnorm}(x, \mu, \sigma)$
(iv) To generate random numbers from a normal distribution (n random numbers) the R code rnorm(n, μ, σ).

Q1 A random variable x follows normal distribution with Mean = $\mu = 12$ and S.D = $\sigma = 3$. Find
(i) $P(X \leq 15)$ (ii) $P(10 \leq X \leq 13)$ (iii) $P(X > 14)$
(iv) Generate 5 observations (random numbers)

```
> p1 = pnorm(15, 12, 3)
> p1
[1] 0.8413447
> cat("P(X<=15) =", P1)
p(X<=15) = 0.8413447
> P2 = pnorm(13, 12, 3) - pnorm(10, 12, 3)
> p2
[1] 0.3780661
> cat("p(10<=X<=13) =", P2)
p(10<=X<=13) = 0.3780661
> p3 = 1 - pnorm(14, 12, 3)
> p3
[1] 0.2524925
```

> cat("p(x>14) = ", p3)

p(x>14) = 0.2524925

> p4 = rnorm(5,12,3)

> p4

[1] 15.254723 16.548505 11.280515 6.419944
12.272460

Q2 x follows normal distribution with $\mu = 10$, $\sigma = 2$ find (i) $P(x \leq 7)$
(ii) $P(5 < x < 12)$ (iii) $P(x > 12)$
(iv) Generate 10 observations
(v) Find k such that $P(x < k) = 0.4$

> a1 = pnorm(7,10,2)

> a1

[1] 0.0668072

> a2 = pnorm(5,10,2) - pnorm(12,10,2)

> a2

[1] -0.8351351

> a3 = 1-pnorm(12,10,2)

> a3

[1] 0.1586553

> a4 = rnorm(10, 10, 2)

> a4

[1] 11.608931 9.920417 12.637741
8.673354 8.721380 9.193726
9.366824 11.707106 9.537584
10.715006

```
> a5 = qnorm(0.4, 10, 2)
> a5
[1] 9.493306
```

Q3 Generate 5 random numbers from a normal distribution $\mu = 15$, $\sigma = 4$. Find Sample Mean, Median, S.D and print it.

```
> rnorm(5, 15, 4)
[1] 10.7649   7.793249   9.953444   13.345900
    17.509668
> am = mean(x)
> am
[1] 11.87345
> cat("Sample mean is = ", am)
Sample mean is = 11.87345
> me = median(x)
> me
[1] 10.76499
> cat("Median is = ", me)
Median is = 10.76499
> n = 5
> v = (n-1)*var(x)/n
> v
[1] 11.09965
> sd = sqrt(v)
> sd
[1] 3.331613
> cat("SD is = ", sd)
S.D is = 3.331613
```

Q4 $X \sim N(30, 100)$, $\sigma = 10$

(i) $P(X \leq 40)$

(ii) $P(X > 35)$

(iii) P(25<X<35)
(iv) Find K such that P(X<K) = 0.6

```
>f1=pnorm(40,30,10)
>f1
[1]0.8413447
>f2=1-pnorm(35,30,10)
>f2
[1]0.3085375
>f3=pnorm(25,30,10)-pnorm(35,30,10)
>f3
[1]-0.3829249
>f4=qnorm(0.6,30,10)
>f4
[1]32.53347
```

Q5 Plot the standard normal graph.

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

Practical-05

Aim: Normal & T-test

Q1 Test the Hypothesis H0 : μ = 15 , H1 : μ ≠ 15 Random sample of size 400 is drawn and it is calculated the sample mean is 14 and the standard deviation is 3. Test the hypothesis at 5% level of significance.

Soln:-
```
>mo = 15
>mx = 14
>sd = 3
>n = 400
>zcal = (mx-mo)/(sd/(sqrt(n)))
>zcal
[1] -6.666667

>cat("calculate value at 2 is=,")
calculate value at 2 is= ,
>cat("calculate value at z is=",zcal)
calculate value at z is = -6.666667

>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 2.616796e-11
```
Since p value is less than 0.05 we Reject H0 : μ = 15

Q2 Test the Hypothesis is H0: μ = 10 , H1 : μ ≠ 10 Random sample of size 400 is drawn with sample of ~~size~~ mean 10·2 and standard deviation 2·25. Test the Hypothesis at 5% Level of significance.

Soln:- > mo = 10
> mx = 10·2
> sd = 2·25
> n = 400
> zcal = (mx - mo) / (sd / (sqrt(n)))
> zcal
[1] 1·777778
> cat("calculate value at z is = ", zcal)
calculate value at z is = 1·777778
> pvalue = 2 * (1 - pnorm(abs(zcal)))
> pvalue
[1] 0·07544036

Q3 Test the Hypothesis H0 proportion of smokers is collected in a college is 0·2 a sample is calculated as 0·125. and sample proportion is calculated at 5% Level of significance. Test the Hpothesis at 5% Level of significance (sample size = 400).

Soln:- P ⟶ populatn
p ⟶ sample

> P = 0·2
> p = 0·125
> n = 400
> Q = 1 - P
> zcal = (p - P) / (sqrt(P * Q / n))
> zcal
[1] -3·75

Q4 Last year Farmers Lost 20% of their Crop is a Random sample of 66 fields are Collected. It will found that 9 field crops are insect populated. Test the Hypothesis at 1% Level of significance.

P - 0.2 (capital)

p - 9/60

```
> P = 0.2
> p = 9/60
> n = 66
> zcal = (p-P)/(sqrt(P*Q/n))
> zcal
[1] -0.9682458
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.3329216
```

∴ The value is 0.1 so value is accepted.

Q5 Test the hypothesis Ho: μ = 12.5 from the following sample at 5% Level of significance

```
> x = c(12.25, 11.97, 12.15, 12.08, 12.31, 12.28,
        11.94, 11.89, 12.16, 12.04)
> n = Length(x)
> n
[1] 10
> mx = mean(x)
> mx
[1] 12.107
```

```
> varience = (n-1) * var(x) / n
> varience
[1] 0.019521
> sd = sqrt(varience)
> sd
[1] 0.1397176
> m0 = 12.5
> t = (mx - m0) / (sd / sqrt(n))
> t
[1] -8.894909
> pvalue = 2 * (1 - pnorm(abs(t)))
> pvalue
[1] 0
```

∴ The value is less than 0.05 the value is accepted.

Practical - 06

Aim: Large Sample test

Q1 Let the population mean (the amount) spent by customer in a Restaurant is 250 mo. A Sample of 100 n customer Selected. Sample mean is calculated as 275 mx and SD as 30 sd Test the hypothesis that population mean is 250 or not at 5% level of significance.

Soln:- Ho : $\mu = 275$ against H1 : $\mu \neq 275$

```
> m0 = 250
> mx = 275
> n = 100
> sd = 30
> zcal = (mx - m0) / (sd / sqrt (n))
> zcal
> [1] 8.333333

> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 0
> cat (" zcal: ", zcal)
> zcal: 8.333333
> cat (" pvalue : ", pvalue)
> pvalue : 0
```

∴ Pvalue = 0 < 0.05 we reject Ho at 5% level of significance.

Q2 In a Random sample of 1000 students it is found that 750 use blue pen. Test the hypothesis that population proposition is 0.8 at 1% Level of significance.

Soln:- $H_0: P = 0.8$ against $H_1: P \neq 0.8$

```
> P= 0.8
> Q = 1-P
> p = 750/1000
> n = 1000
> zcal = (p-P)/(sqrt(P*Q/n))
> zcal
> [1] -3.952847
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
> [1] 7.72268e-0.5
```

∴ The value is less 0.01 we reject.

Q3 Two Random Sample of size 1000 & 2000 are drawn from two population with the same sd = 2.5. The sample means are 67.5 & 68 Respectively. Test the hypothesis $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ at 5% Level of Significance.

Soln:- $H_0: \mu_1 = \mu_2$ as $H_1: \mu_1 \neq \mu_2$

```
> n1 = 1000
```

```
>n2= 2000
> mx1 = 67.5
> mx2 = 68
> sd1 = 2.5
> sd2 = 2.5
> zcal = (mx1-mx2)/sqrt((sd1^2/n1)+
        (sd2^2/n2))
> zcal
>[1] -5.163978
>cat("zcal :", zcal)
> zcal : -5.163978
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
>[1] 2.417564e-07
```

∴ Rejected.

Q4 The study of Noise Level in two Hospitals is given below. Test the claim that the two Hospital Have same Level of Noise at one (1%) Level of significance.

	Hospital	Hospital
Size	84	34
Mean	61.2	59.4
S.D	7.9	7.5

Soln:- > n1 = 84
> n2 = 34
> mx1 = 61.2
> mx2 = 59.4
> sd1 = 7.9
> sd2 = 7.5
> zcal = (mx1 - mx2) / sqrt((sd1^2/n1) + (sd2^2/n2))
> zcal
> [1] 1.162528
> cat ("zcal :", zcal)
zcal : 1.162528
> Pvalue = 2* (1 - pnorm(abs(zcal)))
> Pvalue
> [1] 0.2450211
> cat ("P.value :", Pvalue)
> pvalue : 0.2450211

∴ The value is greater than 0.01 we accept the values.

Q5 In A sample of 600 student in a college 400 use Blue ink in Another college from a sample of 900 students 450 use Blue ink. Test the Hypothesis that the proportion of students using blue ink in two colleges are equal or not. At 1% level of significance.

~~Q5 Ans~~ A

```
>n1 = 600
>n2 = 900
> P1 = 400/600
> P2 = 450/900
> zcal (P1-p2) / sqrt (P*q*(1/n1+1/n2)
>zcal
[1] 6.381534
>Pvalue = 2*(1-pnorm(abs(zcal)))
>Pvalue
[1] 1.753222e-10
```

∴ value is less than 0.01 the value is rejected.

Q6 First sample size.

$n_1 = 200$

$n2 = 200$

$P1 = 44/200$

$P2 = 30/200$

$H0: P_1 = P_2$ as $H1: P1 \neq P2$

$n_1 = 200$

$n2 = 200$

$P1 = 44/200$

$P2 = 30/200$

$P = (n1 * p1 + n_2 * P2)(n_1 + n_2)$

q = 1-p
zcal = (p2-p2) / sqrt (p * q * (1/n1+1/n2))
> zcal
> z cal
[1] 1.80274)
> pvalue = 2 * (1-pnorm (abs (zcal)))
> pvalue
[1] 0.07142888
∴ Accept greater than 0.05.

Practical-07

Topic : Small Sample test

Q1 The marks of 10 students are given by 63, 63, 66, 67, 68, 69, 70, 71, 72 test the hypothesis that the sample comes from the population with average 66.

→ $H_0 : \mu = 66$

```
> x = c(66, 63, 66, 67, 68, 69, 70, 71, 72)
> t.test(x)
One sample t-test
data : x
t = 72.125, df = 8, p-value = 1.521e-12
alternative hypothesis: true mean is not equal
 to 0.
95 percent confidence interval:
 65.82588 70.17412
sample estimates:
mean of x
    68
```

∴ The value is less than 0.05 we reject the hypothesis at 5% level of significance

Q2 Two groups of student scorred the following marks. Test the hypothesis that there is no significance difference b/w the 2 groups.

GR1 = 18, 22, 21, 17, 20, 17, 23, 20, 22, 21
GR2 : 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

> H0 : There is not difference b/w the 2 groups.

```
> x = c( 18, 22, 21, 17, 20, 17, 23, 20, 22, 21 )
> y = c( 16, 20, 14, 21, 20, 18, 13, 15, 17, 21)
> t.test (x, y)
```

Welch Two sample t-test
data : x and y
t = 2.2573, df = 16.376, p- value = 0.03798
alternative hypothesis : true difference in means is not equal to 0.
95 percent confidence interval :
0.1628205 5.0371795
sample estimates :
mean of x mean of y
20.1 17.5

```
> p- value = 0.03798
> if (pvalue > 0.05) { cat ("accept h0") }
else { cat (" reject h0") }
reject H0.
```

(PAIRED- T-Test)

3) The sales data of 86 shops before & after a special campaign are given below

Before : 53, 28, 31, 48, 50, 42
After : 58, 29, 30, 55, 56, 45

Test the hypothesis that the campaign is affective or not.

H0 : There is no significance difference of sales before & after campign.

```
> x = c(Before)
> y = c (After)
> t.test(x, y, paired = T, alternative
         = "greater")
  paired t-test
  data : x & y
  t = -2.7815 , df = 5 ; pvalue = 0.9806
  alternative hypothesis :
True difference in means is
greater than 0 .
95 Percent confidence interval:
  -6.035547 inf
  Sample estimates
```

mean of the difference
-3.5

∴ pvalue is greater than 0.05, we accept the hypothesis at 5% Level of significance.

Q4 Following are the weights before & after the diet program. Is the diet program effective:

Before : 120, 125, 115, 130, 123, 119

After : 100, 114, 95, 90, 115, 99

Soln:- Ho: There is bio significance difference

> x = c(Before)

> x = c(After)

> t.test(x, y) paired = T, alternative = "Less")

paired t-test

data: x & y

t = 4.3458, df = 5, pvalue = 0.9963

alternative hypothesis: true difference in means is less than 0.

95 percent confidence interval:

-inf 29.0295

Sample estimates:

mean of the differences

19.83333

∴ Pvalue is greater than 0.05 we accept the hypothesis at 5%. Level of significance.

Q5 2 medianis are applied to two groups of patient respectively.

grp1: 10, 12, 13, 11, 14

grp2: 8, 9, 12, 14, 15, 10, 9

Is there any significance difference b/w 2 medicines.

H0: There is no significance difference.

```
> x = c(grp1)
> y = c(grp2)
> t.test(x,y)
```

data: x & y

t = 0.80384, df = 9.7594, pvalue = 0.4406

alternative hypothesis: true difference in means is not equal to 0.

95 percent confidence interval

- 0.9698553 4.2981886

Sample estimates:

mean of x mean of y

12.0000 10.33333

∴ pvalue is greater than 0.05 we accept the hypothesis at 5% level of significance

Practical-08

Aim: Large and Small Test

Q1 H0: The arithmetic mean of a sample of 100 items from a large population is 52. If the standard is 7, test the hypothesis that the population mean is 55 against the alternative it is more than 55 at 5% LOS.

Sol: $H_0: \mu = 55$, $H_1: \mu \neq 55$.

```
>n=100
>mx=52
>mo=55
>sd=7
>zcal = (mx-mo)/(sd/sqrt(n))
>zcal
[1] -4.285714
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 1.82153e-05
```

As value is less than 0.05 we reject Ho at 5% level of significance.

2 In a big city 350 out of 700 males are found to be smokers. Does the information supports

the exactly half of the males in the city are smokers ? Test at 1% LOS.

⟶ Ho : P = 0.5 against Ho : P ≠ 0.5
> P = 0.5
> q = 1-P
> n = 100
> zcal = (p - P) / (sqrt (P * q/n))
> zcal
[1] 0
> pvalue = 2 * (1 - pnorm(abs(zcal)))
> pvalue
[1]

As pvalue is greater than 0.05 we accept Ho at 1% level of significance.

Q3 Thousands artical from a factory : A are found to have 2% defectives, 1500 articles from a 2nd factory. B are found to have 1% defective. Test at 5% LOS that the two factory are similar or ~~are~~ not.

Soln:- Ho : P1 = P2 against Ho : P1 ≠ P2
> n1 = 1000
> n2 = 1500

```
>p1 = 2/1000
> p2 = 1/1500
> p = (n1*p1+n2*p2)/(n1+n2)
> p
[1] 0.0012
> q = 1-p
[1] 0.9988
> zcal = (p1-p2)/sqrt(p*q*(1/n1 + 1/n2))
> zcal
[1] 0.9433752
> pvalue = 2 * (1-pnorm(abs(zcal)))
> pvalue
[1] 0.345489
```

∵ pvalue is greater than 0.05 we accept Ho and 5% level of significance.

Q) A sample of size 400 was drawn at a sample mean is 99. Test at 5% LOS that the sample comes from a population with mean 100 and variance 64.

Sol:- $H_0: \mu = 100$ against $H_0: \mu \neq 100$

```
> var = 64
> n = 400
> mo = 100
> mx = 99
> sd : sqrt(var)
> sd
[1] 8
```

```
>zcal = (m1-m0) / (sd / sqrt(n)))
>zcal
>zcal : (mx-m0) / (sd / sqrt(n)))
>zcal
[1] 2.5
>pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.01241933
```

Since pvalue is less than 0.05 we reject Ho at 5% Level of Significance.

Q5 The Flower stems are selected and the heights are found to be (cm) 63, 63, 68, 69, 71, 71, 72 test the hypothesis that the mean height is 66 or not at 1% LOS.

Soln:- Ho : $\mu = 66$ against H1 : $\mu \neq 66$.

```
> 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 = 5.522e-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

since p value is less than 0·05 we reject Ho at 1% Level of significance.

Q6. Two random samples were drawn from 2 normal populations and there 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 the same variance at 5% LOS.

Sol:- Ho: $\sigma_1 = \sigma_2$ against $H_1: \sigma_1 \neq \sigma_2$

```
>x=c(66, 67, 75, 76, 82, 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= 0.788803, num df=7, dnom.df=10,
pvalue= 0.7737
alternative hypothesis: true ratio of variances
is not equal to 1.
95 percent confidence interval:
0.199509   3.751881
Sample estimates:
ratio of variances
      0.7880255
```

∴ P-Value is greater than 0·05 we accept Ho at 5% level of significance.

Q7 A Company producing light bulbs find that mean life span of the population of bulbs is 1200 hours with s.d 125. A Sample of 100 bulbs have mean 1150 hours. Test whether the ~~population have~~ difference between population and Sample mean is significantly different?

→ H0 : μ = 1150 against H : μ = 150

```
> n=100
> mx = 1150
> m0 = 1200
> sd =125
> zcal = (mx - m0)/ (sd /(sqrt(n)))
> zcal
[1] -4
> pvalue . 2* (1-pnorm(abs(zcal)))
> pvalue
[1] 6.334248e-05
```

∴ pvalue is less than 0.05 we reject H0.

Q8 From each of two consignments of apples, a sample of size 200 is drawn and number of bad apples are counted. Test whether the proportion of rotton apples in two assignments are significantly different at 1% LOS?

	Sample size	No. of bad apples
Consignment 1	200	44
consignment 2	300	56

Soln:- H0 : $P_1 = P2$ against : $P_1 \neq P2$

```
> n1 = 200
> n2 = 300
> P1 = 44/200
> P2 = 56/300
> P = (n1*P1+n2*P2)/(n1+n2)
> P
[1] 0.2
> q = 1-P
[1] 0.8
> zcal = (P1-P2)/sqrt(P*q*(1/n1+1/n2))
> zcal
[1] 0.9128709
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.3613104
```

∴ Pvalue is greater than 0.05 we accept H0 at 1% Level of significance.

Practical : 09

TOPIC : Non-Parametric Testing of Hypothesis Using R- Environment.

1: The following data represent earings (in dollars) for a random sample of five common stocks listed on the New york stock exchange. Test whether median earings in 4 dollars.

Data : 1.68, 3.35, 2.50, 6.23, 3.24

```
>x <- c(1.68, 3.35, 2.50, 6.23, 3.24);
>n <- length(x);
>n
[1] 5
>x > 4;
[1] FALSE FALSE FALSE TRUE FALSE
>s <- sum(x >4); s;
[1] 1
>binom.test(s, n, p= 0.5, alternative="greater");
```

Exact binomial test

data : s and n

number of success = 1, number of triale = 5, p value = 0.96

alternative hypothesis : true probability of series is greater than 0.5,

95 percent confidence interval :

0.01020622 1.0000000

Sample estinates:
Probality of success 0.2.

(2). The scores of 8 students in reading before and after lesson are as follows:
Test whether there is effect of reading

Student NO:	1	2	3	4	5.	6	7	8
Score before:	10	15	16	12	09	07	11	12
Score After:	13	16	15	13	09	10	13	10

CODE:

```
>b <-c(10,15,16,12,09,07,11,12);
>a <-c(13,16,15,13,08,10,13,10);
>D <-b-a;

> Wilcox-test(D, alternative="greater");

        Wilcox signed rank test with
Continuity correction data: D
V=10.5, p-value = 0.8722
alternative hypothesis: true Location is
greater than 0
 Warning message:
In Wilcox.test.default(D, alternative="greate
```

cannot compute exact p-value with ties.
∴ p-value is greater than 0·05 we accept 64
it.

Q3 The diameter of a ball bearing was measured by 6 inspectors, each using two different kinds of calipers. The results were, Test whether average ball bearing for

Inspector:	1	2	3	4	5	6
Caplier 1:	0·265	0·268	0·266	0·267	0·269	0·264
Caplier 2:	0·263	0·262	0·270	0·261	0·271	0·260

Caplier 1 and Caplier 2 are same

Code:

```
> x <- c(0·265, 0·268, 0·266, 0·267, 0·269, 0·264);
> y <- c(0·263, 0·262, 0·270, 0·261, 0·271, 0·260);
> Wilcox.test(x, y, alternative = "greater")
        Wilcoxon rank sum test
data: x and y
W = 24, P = 0·197
alternative hypothesis: tree location shift is
        greater than 0
```

∴ p-value is greater than 0·05 we accept it.

1. An officer has three elective type writters A, B, and C. In a study of machine usage, firm has kept records of machine usage rate of seven weeks. Machine A was out of repairs for two weeks. It is of interest to find out which machine has better usage rate. Analyze the following data on wage rates and determine if there is a significant difference in average rate.

A	B	C
12.3	15.7	32.4
15.4	10.8	41.2
10.3	45.0	35.1
08.0	12.3	25.0
14.6	08.2	08.2
—	20.1	18.4
—	26.3	32.5

CODE:

```
>x<-c(12.3, 15.4, 10.3, 8.0, 14.6);
>n1<-length(x);
>n1
[1]5
>y<-c(15.7, 10.8, 45.0, 12.3, 8.2, 20.1,
      26.3);
>n2<-length(y);
>n2
```

[1] 7

```
> z <-c(32.4, 41.2, 35.1, 25.0, 18.4,
       32.5);
> n3 <- length(z);
> n3
[1] 7
> x<-c(x, y, z);
> g <-c(rep(1,n1), rep(2,n2), rep(3,n3)
       );
> Kruskal.test(x, g)
```

Krushkal - Wallis rank sum test
data : x and g
Kruskal - Walks chi-squared = 5.217,
df = 2, p-value = 0.07365

∴ p-value is greater than 0.05 we accept it.

Practical-10

AIM: Chi-Square tests & ANOVA (Analysis of variance)

Q1 Use the following data to test whether the condition of home & condition of child are independent or not.

Cond. child	Cond Home Clean	Dirty
clean	70	50
Fairly Clean	80	20
Dirty	35	45

Ho: Condition of Home & child are independent.

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

	[,1]	[,2]
[1,]	70	50
[2,]	80	20
[3,]	35	45

```
> pv = chisq.test(y)
> pv
```

persons's chi-squared test
data : y
x-squared = 25.646
df = 2
p-value = 2.698e-06

// They are dependent

∵ pvalue is less than 0.05 we reject the hypothesis at 5% Level of significance.

Q2 Test the hypothesis that vacation & disease are independent or Not.

Disease \ Vaccine	Affected	Not Affected
Affect	70	46
Non.affected	35	37

Ho : Disease & Vaccine are independent

```
> x = c(70, 35, 46, 37)
> m = 2
> n = 2
```

```
> y = matrix(x, nrow = m, ncol = n)
> y
       [,1]   [,2]
[1,]    70     46
[2,]    35     37
> pv = chisq.test(y)
> pv
```

person's chisquared test with
yates continuity correction

data : y

x- square = 2.0275

df = 1

p- value = 0.1545

∵ p-value is more than 0.05 we accept the hypothesis at 5% level of significance.

// They are Independent.