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 FY BSC CS 'A' **Roll No.** 1730 **Year** 2019 - 20

This is to certify that the work entered in this journal is the work of Mst. / ~~Ms.~~ Dinesh K. Lohar

____

who has worked for the year 2019 - 20 in the Computer Laboratory.

Teacher In-Charge

Head of Department

Date : ________________

Examiner

Scanned with CamScanner

★★ INDEX ★★

SEM II.

No.	Title	Page No.	Date	Staff Member's Signature
1.	Practical 1 → Basic of R software	45-46	28.11.19	AV
2.	Practical 2 - Probability Distribution	47-49	5/12/19	AV
3.	Practical 3 - Binomial Distribution	50-52	19/12/19	AV
4.	Practical 4 - Normal Distribution	53-55	2/1/20	AV
5.	Practical 5 Norm and T-test	56-57	16/01/20	AV
6.	Practical 6 Testing of Hypothesis & Large Sample Test	58-60	23/01/20	AV
7.	Practical 7 Small Sample Test	60 – 62	06/02/20	AV
8.	Practical 8 Large and Small Sample Test	63-65	13/02/20	AV

Scanned with CamScanner

★★ INDEX ★★

Scanned with CamScanner

PRACTICAL - 1

Topic :- Basic of R software.

1. R is a software for statistical analysis and data computing.
2. It is any effective data handling software and outcome storage is possible.
3. It is capable of graphical display.
4. It is a free software.

Q. Solve the following :-

1. $4 + 6 + 8 \div 2 - 5$

⇒ 4 + 6 + 8/2 - 5

[1] 9

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

⇒ 2**2 + abs(-3) + sqrt(45)

[1] 13.7082

3. $5^3 + 7 \times 5 \times 8 + 46 \div 5$

⇒ 5**3 + 7*5*8 + 46/5

[1] 414.2

4. $\sqrt{4^2 + 5 \times 3 + 7 \div 6}$

⇒ sqrt(4**2 + 5*3 + 7/6)

[1] 5.671567

Scanned with CamScanner

5. round off 46÷7 +9*8
=> round(46/7 +9*8).
[1] 79.

Q. Vector problems : solve the following :-

1. c(2,3,5,7) * 2
[1] 4 6 10 14

2. c(2,3,5,7) * c(2,3)
[1] 4, 9, 10, 21

3. c(2,3,5,7) * c(2,3,6,2)
[1] 4 9 30 14

4. c(2,3,5,7) ^ 2
[1] 4 9 25 49

5. c(1,6,2,3) * c(-2,-3,-4,-1)
[1] -2 -18 -8 , -3

6. c(4,6,8,9,4,5) ^ c(1,2,3)
[1] 4 36 512 9 16 125

7. c(6,2,7,5) / c(4,5)
[1] 1.50 0.40 1.75 1.00

Scanned with CamScanner

Q. Solve →
$x = 20, y = 30, z = 2$
find,
i) $x^2 + y^3 + z$
ii) $\sqrt{x^2 + y}$
iii) $x^z + y^z$

⇒ i) > x**2 + y**3 + z
⇒[1] 27402

ii) > sqrt (x**2 + y)
[1] 20.73644

iii) > x**z + y**z
[1] 1300

* MATRIX :-
syntax - x <- matrix (nrow = 4, ncol = 2, data = c(1,2, ...))

⇒ 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

Q. Find $x + y$ and $2x + 3y$, where $x = \begin{bmatrix} 4 & -2 & 6 \\ 7 & 0 & 7 \\ 9 & -5 & 3 \end{bmatrix}$ $y = \begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 5 \end{bmatrix}$

⇒ > x <- matrix (nrow = 3, ncol = 3, data = c(4,7,9,-2,0,-5,6,7,3))
> y <- matrix (nrow = 3, ncol = 3, data = c(10,12,15,-5,-4,-6,7,9,5))
> x

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

> y

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

Scanned with CamScanner

```
> 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
```

Q. Marks of statistics of CS students A Batch out of 60

x
59
20
35
24
46
56
55
45
27
22
47
58
54
40
50
32
36
29
35
39

```
> x = c(59, 20, 35, ......, 39)
> length(x)
> [1] 20
> breaks = seq(20, 60, 5)
> a = cut (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
```

A.Vw 28.11

PRACTICAL - 2

Aim :- Probability Distribution.

Q.1 Check whether the following are p.m.f or not :-

i)

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

since, $p(2) = -0.5$, can't be p.m.f
because $p(x) \geq 0 \ \forall x$.

ii)

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

```
> p = c(0.2,0.2,0.3,0.2,0.2)   > sum(p)   [1] 1.1
```

$\therefore \quad \Sigma p(x) = 1.1$

$\therefore$ It can't be p.m.f because $\Sigma p(x) = 1$.

iii)

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

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

$\therefore \quad \Sigma p(x) = 1$

It is a p.m.f

Q.2

i) Find c.d.f for the following p.m.f and sketch the graph :-

a)

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

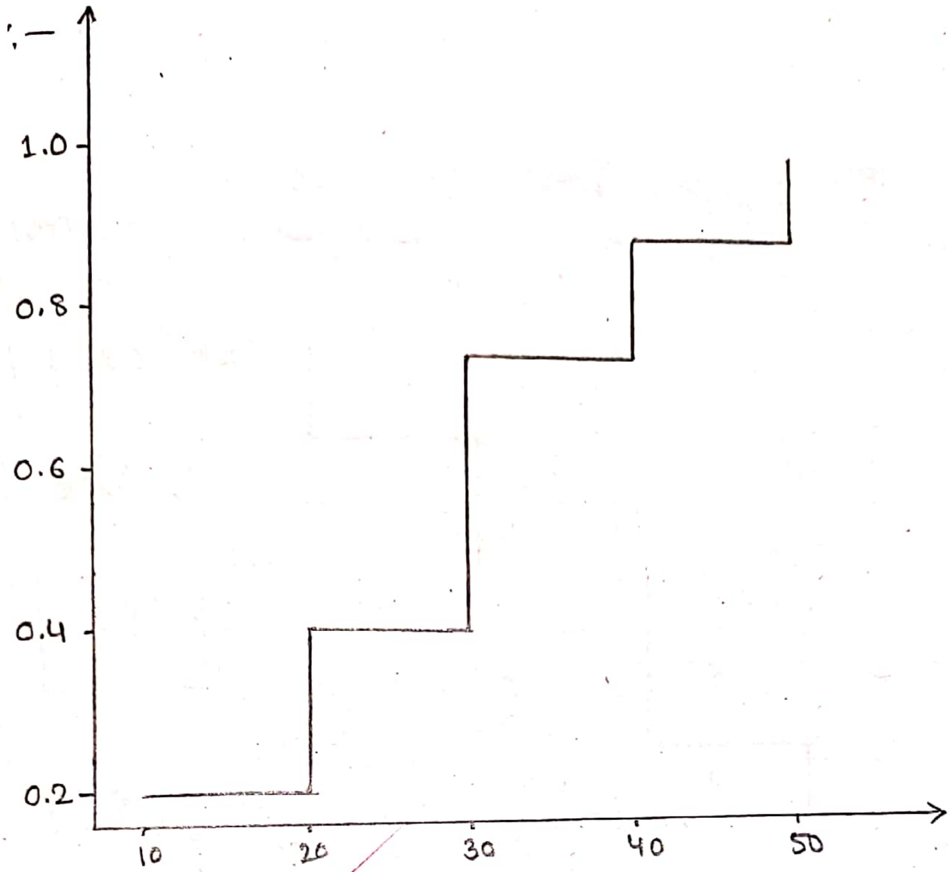
```
> prob = c (0.2, 0.2, 0.35, 0.15, 0.1)
> sum (prob)
[1] 1
```

$F(x) = 0$	$x < 10$
0.2	$10 \le x < 20$
0.4	$20 \le x < 30$
0.75	$30 \le x < 40$
0.9	$40 \le x < 50$
1.0	$x \ge 50$

```
> cumsum (prob)
[1] 0.20  0.40  0.75  0.90  1.00
> x = c(10, 20, 30, 40, 50)
> plot (x, cumsum (prob), "s")
```


Scanned with CamScanner

Graph :-



b)

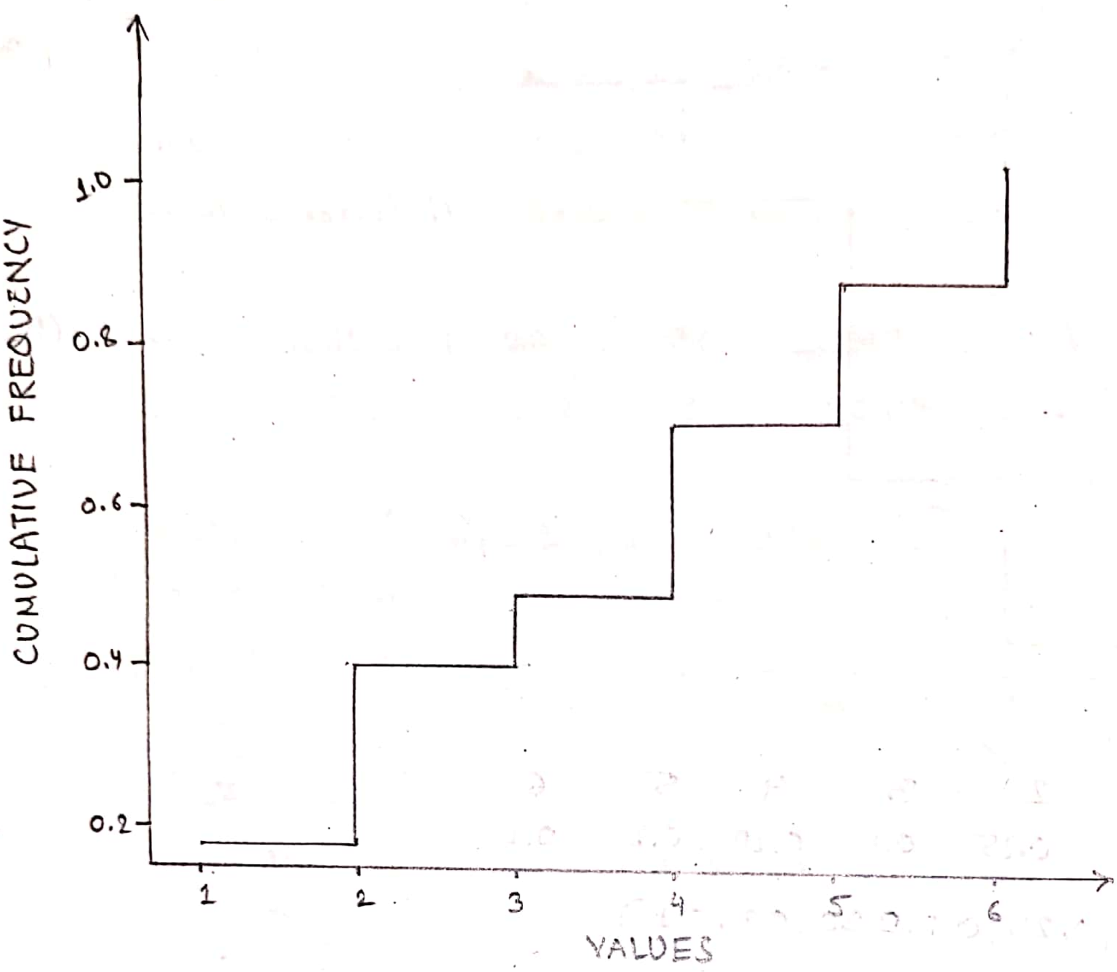
x	1	2	3	4	5	6
$p(x)$	0.15	0.25	0.1	0.20	0.2	0.1

```
>prob=c(0.15,0.25,0.1,0.20,0.2,0.1)
>sum(prob)
[1] 1
>x=cumsum(prob)
>x
[1]  0.15  0.40  0.50  0.70  0.90  1.00
>y=c(1,2,3,4,5,6)
>plot(y,x,"s",xlab="VALUES",ylab="CUMULATIVE
      FREQUENCY",main="C.D.F")
```

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

Scanned with CamScanner

Graph :-



Q.3 Check whether the following is p.d.f or not

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

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

i) Solution :-

$$\int_0^1 f(x)\, dx$$

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

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

$$= [3x - x^2]_0^1$$

$$= 2$$

$\because$ It is not equal to 1 , It is not a p.d.f

Scanned with CamScanner

ii) Solution :-

$$\int_0^1 f(x)\, dx$$

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

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

$$= [x^3]_0^1$$

$$= 1$$

Since, integration is equal to 1

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

AY 5-12-19

Scanned with CamScanner

PRACTICAL - 3

Topic :- Binomial Distribution

command → $P(X = x) =$ dbinom (x, n, p) exactly
→ $P(X \leq x) =$ pbinom (x, n, p) Atmost
→ $P(X > x) = 1 -$ pbinom (x, n, p) More than
→ If x is unknown & $P_1 = P(X \leq x)$
qbionm (P_1, n, p)

Q.1) Find the probability of exactly 10 success in 100 trials with $p = 0.1$?

Q.2) Suppose there are 12 MCQ. Each question has 5 options out of which one is correct. Find the probability having :- $(P = 0.2)$
i) Exactly 4 correct answers.
ii) Atmost 4 correct answers.
iii) More than 5 correct answers.

Q.3) Find the complete distribution when $n = 5$ and $P = 0.1$?

Q.4) $N = 12$, $P = 0.25$, Find the following probability :-
i) $P(X = 5)$
ii) $P(X \leq 5)$
iii) $P(X > 7)$
iv) $P(5 < X < 7)$

Scanned with CamScanner

Q.5) The probability of a salesman making a sale to a customer 0.15. Find the probability of 51)

i) no sales out of 10 customer

ii) more than 3 sales out of 20 customer

iii)

Q.6) A salesman has a 20% probability of making a sale out of 30 customer. What minimum no. of sales he can make with 88% probability.

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

Scanned with CamScanner

Solution :-

1.)

```
>dbinom (10,100,0.1)
[1] 0.13186653
```

2.) i)
```
> dbinom(4,12,0.2)
[1] 0.1328756
```

ii)
```
> pbinom(4,12,0.2)
[1] 0.9274445
```

iii)
```
> 1-pbinom (5,12,0.2)
[1] 0.01940528
```

3.)
```
> dbinom (0:5,5,0.1)
[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001
```

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

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

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

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

5.) i)
```
> dbinom (0,10,0.15)
[1] 0.1968744
```

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

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

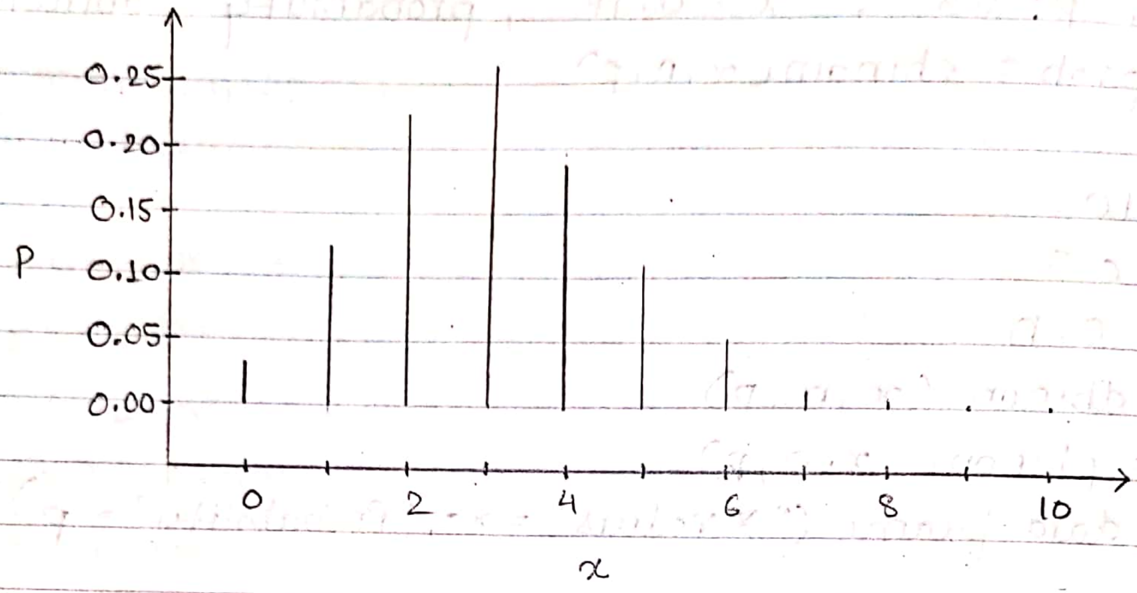
Scanned with CamScanner

7) $n=10$, $p=0.3$, $x=0:n$, probability = dbinom(x,n,p)
cumprob = pbinom(x,n,p)

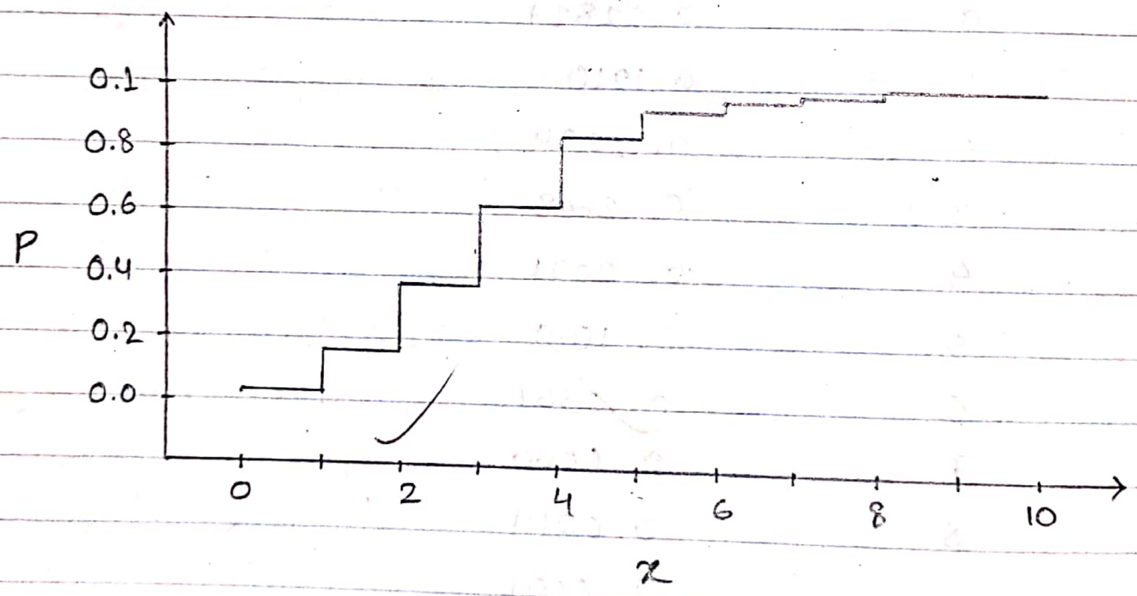
```
> n=10
> p=0.3
> x=0:n
> p=dbinom(x,n,p)
> cp=pbinom(x,n,p).
> d=data.frame("X values"=x, "Probability"=p)
> d
```

	X values	Probability
1	0	0.02824
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

```
>plot(x,p,"h")
```



> plot (x, cp, "s")



19-12-19

PRACTICAL - 4

Topic:- Normal Distribution

Commands :-

i) $P(X=x)$ = dnorm (x, μ, σ)

ii) $P(X \le x)$ = pnorm (x, μ, σ)

iii) $P(X > x)$ = 1 - pnorm (x, μ, σ)

iv) generate random number from normal distribution (n random numbers) the R code is " rnorm (n, μ, σ).

Problems :-

1.) A random variable X follows normal distribution with mean equals to $\mu = 12$ & standard Deviation $= \sigma = 3$. Find :-

i) $P(X \le 15)$

ii) $P(10 \le X \le 13)$

iii) $P(X > 14)$

iv) Generate 5 observations (Random Numbers)

⇒ $\mu = 12$, $\sigma = 3$

i)
```
> pnorm (15, 12, 3)
[1] 0.8413447
> cat ("P(X<=15) = ", pnorm(15,12,3))
P(X<=15) = 0.8413447
```

ii)
```
> 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
```

iii)
```
>p3=1-pnorm(14,12,3)
>p3
[1] 0.2524925
>cat ("P(X>14) = ", p3)
P(X>14) = 0.2524925
```

iv)
```
>rnorm(5,12,3)
[1] 14.447671  14.222767  9.717719  8.081383  13.963679
```

2.) X follows normal distribution with $\mu = 10$, $\sigma = 2$. Find :-

i) $P(X \le 7)$

ii) $P(5 \le X \le 12)$

iii) $P(X > 12)$

iv) Generate 10 obs.

v) Find K such that $P(X < K) = 0.4$

$\Rightarrow$ $\mu = 10$, $\sigma = 2$

i)
```
>p1=pnorm(7,10,2)
>p1
[1] 0.0668072
```

ii)
```
>p2=pnorm(12,10,2)-pnorm(5,10,2)
>p2
[1] 0.8351351
```

Scanned with CamScanner

iii) > p3 = 1 - pnorm(12,10,2)
> p3
[1] 1586553

iv) > rnorm(10,10,2)
[1] 10.669229 7.370285 11.599006 9.448043 6.386619
9.270702 7.147375 12.389127 9.323450 9.215697

v) > qnorm(0.4,10,2)
[1] 9.493306

3.) Generate 5 random numbers, normal distribution with mean = 15, S.D = 4. Find sample mean, median, & S.D and print it.

⇒ > x = rnorm(5,15,4)
> x
[1] 11.48082 12.85051 17.05461 19.64004 18.29075
> n = 5
> am = mean(x)
> am
[1] 15.86334
> md = median(x)
> md
[1] 17.05461
> v = (n-1) * var(x)/n
> v = 9.971698
> sd = sqrt(v)
> sd
[1] 3.157799
> cat("Sample Mean = ", am)
Sample Mean = 15.86334
> cat("Sample Median = ", md)
Sample Median = 17.05461
> cat("Sample S.D = ", sd)
Sample S.D = 3.157799

Scanned with CamScanner

4.) X follows N (30,100)

$X \sim N(\mu, \sigma^2)$

$\therefore \mu = 30, \sigma = 10$

i) $P(X \leq 40)$

```
> p1 = pnorm(40,30,10)
> p1
[1] 0.8413447
```

ii) $P(X > 35)$

```
> p2 = 1- pnorm(35,30,10)
> p2
[1] 0.3085375
```

iii) $P(25 \leq X \leq 35)$

```
> p3 = pnorm(35,30,10) - pnorm(25,30,10)
> p3
[1] 0.3829249
```

iv) Find K such that $P(X < K) = 0.6$

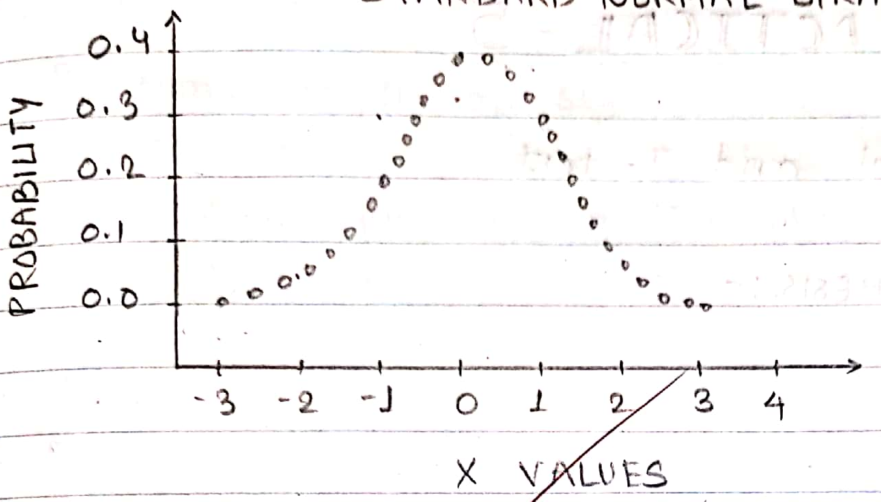
```
> qnorm(0.6,30,10)
[1] 32.53347
```

5.) Plot the standard normal graph.

```
i) > x = seq(-3,3, by = 0.1)
   > y = dnorm(x)
   > plot (
```

Scanned with CamScanner

STANDARD NORMAL GRAPH



PRACTICAL - 5

Topic :- Normal and T- test.

Q.1) TEST THE HYPOTHESIS :-

H_0 : $\mu = 15$

H_1 : $\mu \neq 15$

Random sample of size 400 is drawn and it is, calculate the sample mean is 14, standard deviation is 3, test the hypothesis at 5% level of significance.

```
> mo =15; mx=14; sd=3; n=400
> zcal= (mx-mo)/(sd/(sqrt(n)))
> cat("Calculated value of z : ", zcal)
  Calculated Value of z : -6.666667
> pvalue =2 * (1-pnorm (abs(zcal)))
> pvalue
[1]  2.616796e-11
```

The PValue is less than 0.05 then the value as Null Hypothesis ∴ Rejected.

Scanned with CamScanner

Q.2) TEST THE HYPOTHESIS :-

Ho : $\mu = 10$

H1 : $\mu \neq 10$

Random sample of size 400 is drawn with same mean 10.2 and Standard Deviation 2.25, Test the hypothesis of 5% of level of significance.

```
> mo = 10
> mx = 10.2
> n = 400
> sd = 2.25
> zcal = (mx-mo) / (sd /sqrt(n)))
> cat ("Calculated Value of z : ", zcal)
> Calculated Value of z : 1.77778
> pvalue = 2* (1-pnorm (abs(zcal)))
> pvalue
[1] 0.07544036
```

The Pvalue is greater than 0.05.
The value is Accepted.

Q.3) Proportion of smokers in a college is 0.2. A sample is collected and sample proportion is calculated 0.125, Test the hypothesis at 5% level of significance. (Sample size is 400).

TEST THE HYPOTHESIS ; Ho

```
> p = 0.2
> p = 0.125
> n = 400
> Q = 1-P
> zcal = (p-P) /(sqrt (P*Q)/n))
> cat (" Calculated Value of z : ", zcal)
  Calculated Value of z : -3.75
> Pvalue = 2*(1-pnorm (abs (zcal)))
```

Scanned with CamScanner

```
>pvalue
[1] 0.0001768346
```

The Pvalue is smaller than 0.05
The Pvalue is rejected

Q.3) Last Year farmers lost 20% of their crop. A ran. sample of 60 fields are collected and it is found that 9 fields crops are insect polluted test the hypothesis at 1% level of significance.

```
>P=0.2;p=9/60;n=60; Q=1-P
>zcal = (p-P)/(sqrt(P*Q/n))
>cat("Calculated value of z: ",zcal)
> Calculated Value of z is : -0.9682458
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.3329216
```

The Pvalue is greater than 0.01
The Pvalue is Accepted

Q.4) TEST THE HYPOTHESIS: Ho = μ=12.5
From the following sample at 5% level of significance.

```
>x=c(12.25, 11.97, 12.08, 12.31, 12.28, 11.94, 11.89, 12.16,
    12.04,12.15)
>n=length(x)
>n
[1] 10
>mx=mean(x)
>variance = (n-1)*var(x)/n
>sd= sqrt(variance)
>mo=12.5
>t= (mx-mo)/(sd/(sqrt(n)))
>cat ("Calculated value of t :",t)
Calculated Value of t :
```

Scanned with CamScanner

```
> pvalue = 2 * (1-pnorm(abs(t)))
> pvalue
[1] 0
```

The pvalue is less than 0.05
The Pvalue is rejected.

AN
16.01.

PRACTICAL - 6

Topic :- Large Sample Test

1. Let the population mean (the amount spent per customer in a restaurant) is 250 a sample of 100 customers selected the sample mean is calculated as 275 and S.D 30, test the hypothesis that population mean is 250 or nor at 5% level of significance.

$H_0 : \mu = 275$ against $H_1 : \mu \neq 275$

```
> mx = 275
> mo = 250
> sd = 30
> n = 100
> zcal= (mx-mo) / (sd / sqrt (n))
> cat ("Calculated Value of z : ",zcal)
 Calculated value of z:  8.33333
> pval= 2 * (1-pnorm(abs(zcal)))
> cat ("Calculated Pvalue : ", pval)
Calculated Pvalue : 0
```

∴ Pvalue is less than 0.5 then the value of p is rejected.

Scanned with CamScanner

2. In a random sample of 1000 students. It is found that 750 use blue pen. Test the hypothesis that the population proportion is 0.8 at 1% level of significance.

⇒ Ho : P = 0.8 against H₁ : P ≠ 0.8

```
> P=0.8; q=1-P; p=750/1000; n=1000
> zcal = (p-P)/sqrt (P*(q/n)))
> cat ("Calculated
Calculated Value of z : -3.952847
> pval = 2* (1-pnorm(abs(zcal)))
> cat ("Calculated
Calculated Pvalue : 7.72268e-05
```

∴ Pvalue is less than 0.1, then value is rejected.

3. Two random samples of size 1000 and 2000 are drawn from two population with the sample standard Deviation 2.5. The sample means are 67.5 and 68. Test the Hypothesis Ho : $\mu_1 = \mu_2$ as $H_1 : \mu_1 \neq \mu_2$ at 5% level of significance.

⇒ Ho : $\mu_1 = \mu_2$ 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)))
> cat (" Calculated Value of z : ", zcal)
Calculated Value of z : -5.163978
> Pval = 2* (1-pnorm (abs(zcal)))
> cat (" Calculated P Value : ", pval)
Calculated P Value : 2.417564e-07
```

∴ The Pvalue is less than 0.05 then the value is Rejected.

Scanned with CamScanner

4.) The study of noise level in two hospitals is given below :-
Test the claim that the two hospitals have same noise at 1% level of significance.

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

Ho: $\mu_1 = \mu_2$ & $H_1 : \mu_1 \neq \mu_2$

```
>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)))
>cat("Calculated Value of z : ",zcal)
>Calculated Value of z : 1.162528
>pval = 2*(1-pnorm(abs(zcal)))
>cat(" Calculated P Value : ", pval
Calculated PValue : 0.245021L
```

∴ Pvalue is greater than 0.01 ~~then~~ the value is accepted.

Scanned with CamScanner

5. In a sample of 600 students 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 two colleges are equal or not at 1% level of significance.

$H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

```
> n1 = 600
> n2 = 900
> p1 = 400/600
> p2 = 450/900
> p = (n1*p1 + n2*p2)/(n1+n2)
> q = 1-p
> zcal = (p1-p2)/sqrt(p*q*(1/n1 + 1/n2))
> cat ("Calculated value of z: ", zcal)
Calculated Value of z : 6.381534
> pval = 2 * (1-pnorm (abs (zcal)))
> cat(" Calculated Pvalue : ", pval)
Calculated Pvalue : 1.753222e-10
> p
[1] 0.5666667
```

∴ Pvalue is less than ~~0.01~~, then the value is rejected.

6. $n1 = 200$, $n2 = 200$, sample proportion $(p1) = 41/200$, $p2 = 30/200$, at 5% level of significance.

$H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

```
> n1 = 200
> n2 = 200
> p1 = 44/n1
> p2 = 30/n2
> p = (n1*P1 + n2*P2) /(n1+n2)
> q = 1-p
> zcal = (p1-p2) /sqrt(p*q* (1/n1 + 1/n2))
> pval = 2* (1-pnorm(abs(zcal)))
> p
> cat("Calculated Value of z : ", zcal)
Calculated value of z : 1.802741
> cat ("Calculated Pvalue : ", pval)
Calculated Pvalue : 0.07142888
> p
[1] 0.184
```

∴ Pvalue is greater than 0.05, then the value is accepted.

AM
27.01.20

PRACTICAL-7

Topic:- Small Sample Test.

1) The marks of 10 students are given by 63, 63, 66, 67, 68, 69, 70, 70, 71, 72. Test the hypothesis that the sample comes from opposition that the sample with average 66.

⇒ $H_0 . \mu = 66$

```
> x = c(63,63,66,67,68,69,70,70,71,72)
> t.test(x)
```

One Sample t-test

data: x
t = 68.319, df = 9, p-value = 1.558e-13
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
65.65171 70.14829
sample estimates:
mean of x.
67.9

As pvalue = 1.558e-13 < 0.05, we reject it

```
> pvalue = 1.558e-13
> if (pvalue > 0.05) {cat("Accept Ho")} else {cat ("Reject Ho")}
Reject Ho
```

Scanned with CamScanner

2) Two groups of students score the following marks. Test the hypothesis that there is no significant difference between two groups.

Group 1 : 18, 22, 21, 17, 20, 17, 23, 20, 22, 21

Group 2 : 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

⇒ Ho : There is no difference between two 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 and y
      20.1        17.5

>pvalue = 0.03798
>if (pvalue >0.05){cat ("Accept Ho")} else{cat("Reject Ho")}
Reject Ho
```

3) The sales data of 6 shops before and after a specials campaing are given below. Test the hypothesis that campaign in effective way or not.
Before: 53, 28, 31, 48, 50, 42
After: 58, 29, 30, 55, 56, 45

⇒ Ho - there is no significance difference before and after the campaing.

```
> x = c(53, 28, 31, 48, 50, 42)
> y = c(58, 29, 30, 55, 56, 45)
> t.test(x, y, paired = T, alternative = "greater")

        Paired t-test

data: x and y
t = -2.7815, df = 5, p-value = 0.9806
alternative hypothesis: true difference in means
                        is greater than 0
95 percent confidence intervals:
   -6.035547        Inf
sample estimates:
mean of the differences
            -3.5

> pvalue = 0.9806
> if (pvalue > 0.05) {cat("Accept Ho")} else {cat("Reject Ho")}
Accept Ho
```

Scanned with CamScanner

4) Two medicines are applied to two group of patients respectively
Group1 - 10, 12, 13, 11, 14
Group 2 - 8, 9, 12, 14, 15, 10, 9
Is there any significance difference between two medicines.

⇒ Ho - there is no significance difference

```
> x=c(10,12,13,11,14)
> y=c(8,9,12,14,15,10,9)
> t.test(x,y)

        Welch Two Sample t-test
data : x and y
t = 0.80384, df = 9.7594, p-value = 0.4406
alternative hypothesis : true difference in means
                         is not equal to 0.
95 percent confidence interval :
  -1.781171      3.781171
sample estimates :
  mean of x     mean of y
      12            11
> pvalue = 0.4406
> if (pvalue >0.05) {cat("Accept Ho")} else {cat("Reject Ho")}
Accept Ho
```

5) The followings are the weights before and after the diet program. Is the diet program **62** effective.

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

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

⇒ Ho - there is no significance diff. between before and after

```
>x=c(120,125,115,130,123,119)
>y=c(100,114,95,90,115,99)
>t.test(x, y, paired=T, alternative="less")

        Paired t-test
data: x and y
t= 4.3458 , df =5, p-value = 0.9963
alternative hypothesis: true difference in means
                        is less than 0.
95 percent confidence interval:
     -Inf     29.0295
sample estimates:
mean of differences
          19.8333

>pvalue = 0.9963
>if (pvalue >0.05) {cat("Accept Ho")} else {cat("Reject Ho")}
Accept Ho
```

A.H
6.7.2

Large and Small Sample tests

Q.1 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.

Q.2 In a big city 350 out of 700 males are found to be smokers. Ddoes the information supports that exactly half of the males in the city are smokers ? Test at 1% LOS.

Q.3 Thousand article from a factory:A are found to have 2% defectives ,1500 atricles from a 2nd factory:B are found to have 1% defective. Test at 5% LOS that the two factory are similar are not.

Q.4. 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.

Q.5. 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.

Q.6. 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.

7. A company producing light bulbs finds 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 difference between population and sample mean is significantly different?

8. 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

Scanned with CamScanner

PRACTICAL - 8

Topic :- Large and Small Sample Test

Q.1 $H_0 : \mu = 55$ against $H_1 : \mu \neq 55$

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

(pval is Rejected) pval<0.05

Q.2 Ho : P = 0.5 against H1 : P ≠ 0.5

```
> p = 350/700
> P = 0.5
> n = 700
> q = 1 - p
> zcal = (p - P) / sqrt (P * (q/n))
> zcal
[1] 0
> pval = 2 * (1 - pnorm (abs (zcal)))
> pval
[1]  1
```

(pval is Accepted) pval>0.01

Q.3 $H_0 : P_1 = P_2$ against $H_1 : P_1 \neq P_2$

```
> n1 = 1000
> n2 = 1500
> P1 = 0.02
> P2 = 0.01
> p = (n1 * P1 + n1 * P2) / (n1 + n2)
> q = 1 - p
> p
[1] 0.5
> q
[1] 0.986
> zcal = (P1 - P2) / sqrt(p * q * (1/n1 + 1/n2))
> zcal
[1] 2.084842
> pval = 2 * (1 - pnorm(abs(zcal)))
> pval
[1] 0.03708364
```

(Pval is Rejected) pval < 0.05

Q.4 $H_0 : \mu = 100$ against $H_1 : \mu \neq 100$

```
> mx = 100
> mo = 99
> n = 400
> v = 64
> sd = sqrt(v)
> zcal = (mx - mo) / (sd / sqrt(n))
> zcal
[1] 2.5
```

```
>pval = 2*(1 - pnorm(abs(zcal)))
>pval
[1] 0.01241933
```

(pval is Rejected) pval < 0.05

Q.5 $H_0 : \mu = 66$ against $H_1 : \mu \neq 66$

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

One Sample t-test

```
data:  x
t = 47.94 , df = 6, pvalue = 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
```

∴ pval is Rejected. (pval < 0.05)

Q.6 $H_0 : \sigma_1 = \sigma_2$ against $H_1 : \sigma_1 \neq \sigma_2$

```
>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
```

Scanned with CamScanner

sample estimates:
ratio of variances
0.7068567

→ pval is Accepted (pval > 0.05)

Q.7 Ho: $\mu = 1200$ against $H_1: \mu \neq 1200$

```
> n= 100
> mx = 1150
> mo = 1200
> sd = 125
> zcal = (mx - mo) / (sd / sqrt(n))
> zcal
 [1] 4
> pvalue
 [1] 6.334248e - 05
```

(pval is Rejected) pval < 0.05

Q.8 Ho: $P_1 = P_2$ against $H_1: P_1 \neq P_2$

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

(pval is Accepted) pval > 0.01

```
> P
[1] 0.2
```

——*——*——*——

PRACTICAL - 9

Topic - Chi Square Tests and ANOVA

Q.1) Use the following data to whether the condition of the home is condition of child are independent or not.

Condition of Home.

Condition of Child	Clean	Dirty
clean	70	50
Fairly cln	80	20
Dirty	35	45

→ Ho : Condition of home and child is 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
```

Scanned with CamScanner

> pv = chisq.test(y)
> pv

Person's Chi-Square test

data : y

χ - squared = 25.646 , df = 2 , p-value = 2.698e-06

∴ pvalue < 0.05 (Rejected)

Q.2) Test hypothesis that vaccination and disease are independent or not.

		Vaccine	
		Affective	Not Affective
Disease	Affective	70	46
	Not Affective	35	37

→ Ho: Diseases and Vaccine are independent

```
> x = c(70 , 35 , 46 , 37)
> m = 2
> n = 2
> y = matrix (x, nrow = m, ncol = n)
> y
```

Scanned with CamScanner

```
     [,1] [,2]
[1,]  70   46
[2,]  35   37
```

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

Person's Chi-Square test with Yates' continuity correction

data : y

x-squared = 2.0275, df = 1, p-value = 0.1545

$\therefore$ pvalue > 0.05 (Accepted)

Q.3) Perform a ANOVA for the following data.

Type	Observations
A	50, 52
B	53, 55, 53
C	60, 58, 57, 56
D	52, 54, 54, 55

→ Ho: The means are equal for A, B, C, D.

```
>x1=c(50,52)
>x2=c(53,55,53)
>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)

        One-way 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	18.17	2.019		

signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

∴ Pvalue < 0.05 (Rejected)

Q.4) The following data gives the life of the tires of 4 brands.

A	20, 23, 18, 17, 18, 22, 24
B	19, 15, 17, 20, 16, 17
C	21, 19, 22, 17, 20
D	15, 14, 16, 18, 14, 16

Ho: Average life of a tires is equal.

```
> x1 = c(20, 23, 18, 17, 18, 22, 24)
> x2 = c(19, 15, 17, 20, 16, 17)
> x3 = c(21, 19, 22, 17, 20)
> x4 = c(15, 14, 16, 18, 14, 16)
> d = stack(list(b1 = x1, b2 = x2, b3 = x3, b4 = x4))
```

Scanned with CamScanner

```
> names(d)
[1]  "values" "ind"
> anova = aov(values ~ ind , data = d)
> anova
 Call :
        aov (formula = values ~ ind , data = d)
 Terms :
                                      ind         Residuals
Sum of squares                    91.4381          89.0619
Deg. of Freedom                         3               20

Residual standard error : 2.110236
Estimated effects may be unbalanced

> oneway.test (values ~ind , data = d , var.equal = T)

        Oneway - way  analysis of means

data:  values  and  ind
F = 6.8445 , num df = 3, denom df = 20 , p-value =
                                                  0.002349
```

∴ Pvalue < 0.05   (Rejected)

Q.5)

```
> x = read.csv("c:/Users/tu6t/Desktop/
                marks.csv")
> x
> summary(x$stats)

 Min.   1st Qu.  Median   Mean   3rd Qu.   Max.
 15.00   29.25    38.50   37.00   44.25   59.00

> sd = sqrt((n-1)* var(x$stats)/n)
> sd
[1]   12.64911

> sd = sqrt((n-1)* var(x$maths)/n)
> sd
[1]   15.2
> summary(x$maths)

 Min.   1st Qu.  Median   Mean   3rd Qu.   Max.
 20     25.50    37.00    39.40  54.75     60
```

AM
Jan 2

27/02/20

PRACTICAL - 10

Aim :- Non - Parametric test :

1.) Following are the amount of sulphur oxide comitted by some industry in 20 days, apply sign test to test the hypothesis that the population medium is 21.5 at 5% level of significance.

data - 17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26.

→ Ho : Population medium is 21.5

```
> x = c (data)
> me = 21.5
> sp = length (x [x > me])
> sn = length (x [x < me])
> n = sp + sn
> pv = pbinom (sp, n, 0.5)
> pv
[1] 0.4119015
> n
[1] 20
```

∴ p-value > 0.05 (Accepted)

Scanned with CamScanner

2.) Following is the data of 10 observations apply sign test to the test the Hypothesis that the population median is 625 against the alternative it is more than 625.

612, 619, 631, 628, 643, 640, 655, 649, 670, 663.

→

Ho: population medium is 625.

```
> x = c (data)
> me = 625
> sp= length (x [x>me])
> sn = length (x [x<me])
> n = sp+sn
> pv =pbinom (sn, n, 0.5)
> pv
[1] 0.0546875
> n
[1] 10
```

∴ p-value > 0.05 (Accepted)

Scanned with CamScanner

3) The following are values of sample test the hypothesis that population medium is 60 70 against the alternative it is more than 60 at 5% LOS using
data: 63, 65, 60, 89, 61, 71, 68, 51, 69, 62, 63, 39, 72, 69, 48, 66, 72, 63, 87, 69

→ Ho: population medium = 60, H1: population medium > 60

```
> x = c (data)
> wilcox.test(x, alter ="greater", mu = 60)
```

data : x
V = 145, p-value = 0.02298
alternative hypothesis: true location is greater than 60

∴ p-value < 0.05 (Rejected)

Note: If the alternative is less ; alter ="less"
If the alternative is not equal to ; alter = "two.sided"
If the alternative is greater ; alter = "greater"

4) using WSRT test the population medium is 12 or less than 12.
15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26

→ Ho: population medium is 12, H1: population me < 12

```
> x = c (data)
> wilcox.test (x, alter = "less", mu= 12)
```

Wilcoxon signed rank test with continuity correction

data : x
V = 66, p-value = 0.9986
alternative hypothesis: true location is less than 12

∴ p-value > 0.05 (Accepted)

Scanned with CamScanner

5.) The weights of the students before and after stop smoking are given below using WSRT test that there is no significance change.

Weight Before	Weight After
65	72
75	74
75	72
62	66
72	73

Ho : before and after there is no change
H_1 : There is a change

```
> x = c(weight before)
> y = c (weight after)
> d = x-y
> wilcox.test (d, alter =" two:sides", mu = 0)
```

Wilcoxon signed rank test with continuity correction

data : d

V = 4.5, p-value = 0.4982

alternative hypothesis : true location is not equal to 0

∴ p-value > 0.05 (Accepted)

AN
27.07.