★★ INDEX ★★

PRACTICA – 1

Aim:- Basic of R software

1) R is a software for statical 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 an free software

Solve the following.

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

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

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

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

$5^3 + 7 \times 5 \times 8 + 46/5$

```
> 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] 5.671517
```

5. round off

$46 \div 7 + 9\times8$

```
>round (46÷7 + 9*8)
[1] 79
```

2)

```
>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)*(-2,-3,-4,
[1] -2, -18 -8 -3

>c(4,6,8,9,4,5)^c(1,2,
[1] 4 36 512 9162
```

3)

```
>x= 20 >y =30 >z =2
>x^2 +y^3 +z
[1] 27402
>sqrt (x^2 +y)
[1] 20.73649
>x^2 +y^2
[1] 1300
```

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

$y = \begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 1 \\ 15 & -6 & 5 \end{bmatrix}$

> x <- matrix (nrow = 3, ncol = 3, 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, 7,1,5))

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

```
> x+y      [,1]   [,2]   [,3]
[1,]       14     -7     13
[2,]       17     -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.6) Marks of Statics of CS Path A

```
x = c (50, 20, 35, 24, 46, 56, 55, 45, 27, 22,
       58, 54, 40, 50, 32, 36, 29, 35, 39)
> x = c (data)
> breaks 6 = seq ( 20, 60, 5 )
> a = cut (x, breaks, right = FALSE)
> b = table(a)
> c = transform (b)
> c
```

	a	freq
	[20, 25)	3
1	[25, 30]	2
2	[30, 35]	1
3	[35, 40]	4
4	[40, 45]	1
5	[45, 50]	3
6	[50, 55]	2
7	[55, 60	4
8		

PRACTICAL : 02

TOPIC: probability Distribution.

1) Check whether the followings are p.m.f

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
$\Sigma P(x) = 1$

$\therefore P(0) + P(1) + P(2) + P(3) + P(4) + P(5) =$
$= 0.1 + 0.2 + -0.5 + 0.4 + 0.3 + 0.5$
$= 1.0$

$\because P(2) = -0.5$, it cant a probability m Function.

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

2)

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

Scanned with CamScanner

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

so,

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

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

$= 1.1$

∴ 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) $P(x) \geq 0 \quad \forall x$ satisfy

2) $\Sigma P(x) = 1$

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

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

$= 1$

∴ The given data is p.m.f

code:

```
>prob = c(0.2,0.3,0.35,0.15,0.1)
>sum (prob)
[1] 1
```

check that whether the following is p.d.f or not

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

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

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

$= \int_0^1 f(x)$

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

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

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

$\therefore$ the $\int_0^1 f(x) \neq 1$ $\quad \therefore$ It is not a p.d.f

9) $F(x) = 3x^2 ; \quad 0 < x < 1$

$\int_0^1 F(x)$

$= \int_0^1 3x^2$

$= 3\int_0^1 x^2$

$= \int_0^1 x^2$

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

$= x^3$

$= 1$

The $\int_0^1 F(x) = 1 \quad \therefore$ It is a p.d.f

1) >x = dbinom (10, 100, 0.1)

>x

[1] 0.1318653

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)

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] 0.04014945

PRACTICAL - 3

Topic :- Binomial distribution

\# $P(X = x) = \text{dbinom}(x, n, p)$

\# $P(X \le x) = \text{pbinom}(x, n, p)$

\# $P(X > x) = 1 - \text{pbinom}(x, n, p)$

\# if x is unknown

$p_1 = P(X \le x) = \text{qbinom}(p, n, p)$

1) Find the probability of exactly 10 success is 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 answer

ii) Atmost 4 correct answers.

iii) more than 3 correct answer.

3) Find the complete distribution when $n = 5$ and $p = 0.1$

4) $n = 12$, $p = 0.25$ Find the following problem

i) $P(X = 5)$

ii) $P(X \le 5)$

iii) $P(X > 7)$

iv) $P(5 < X < 7)$

5.) The probability of a salesman making a sale of cu
0.15 find the probability of.
i) No sles out of 10 customer
ii) more than 3 to sales out of 20 customer.

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

7.) X follows binomial distribution with $n = 10$, $p = 0.$
Plot the graph of p.m.f and c.d.f.

Scanned with CamScanner

5) dbinom (0, 10, 0.15)

[1] 0.1968744

>1- pbinom (3,20, 0.15)

[1] 0.352248.

6) qbinom (0.88 130, 0.2)

[1] 9

7) >n=10

>p = 0.3

>x = 0:n

> prob = dbinom (x, n, p)

> cumprob = pd pbinom (x, n, p)

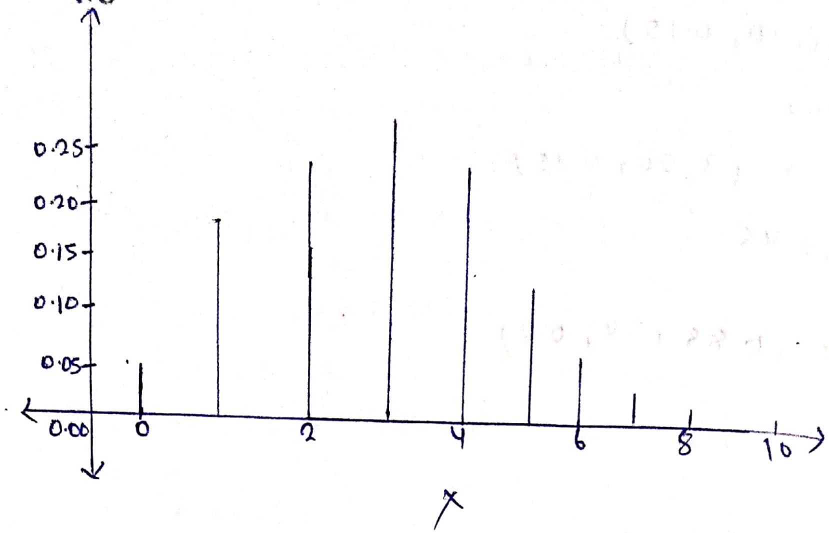
> cumprob = pbinom (x, n, p)

> d = data.fram ("x values"= x, "probability"= prob)

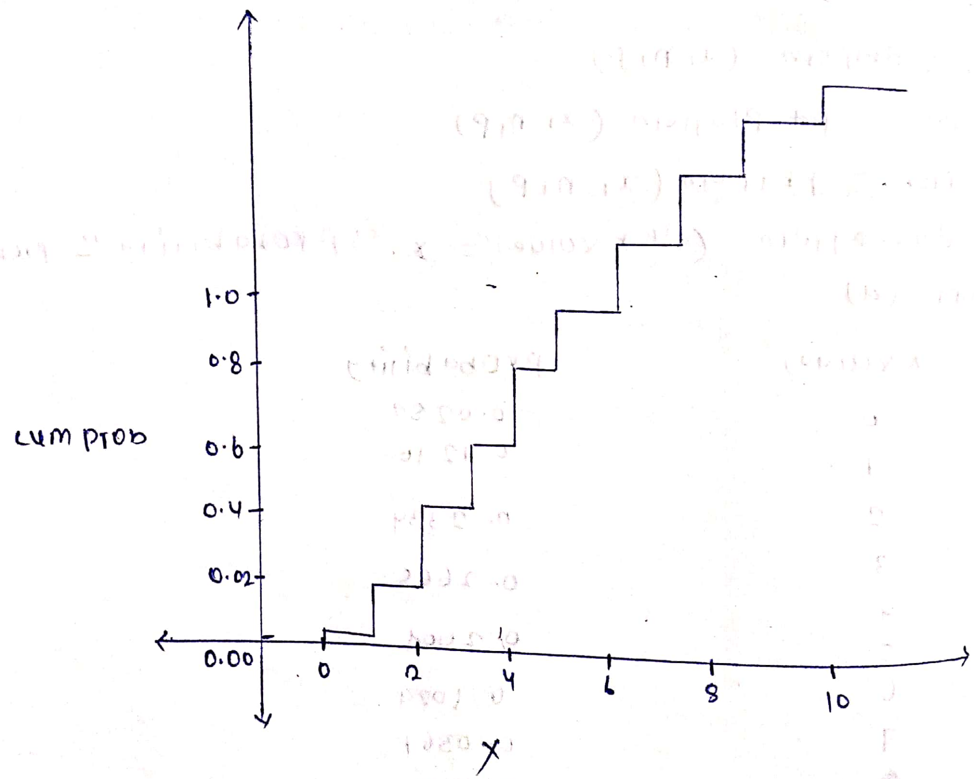
> print (d)

.	x values	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

> plot (x , prob,"h")



> plot (x, cumprob , "s")



PRACTICAL – 4

Aim:- Normal Distribution

\# $P(X=x)$ = dnorm (x, μ, σ)
\# $P(X \le x)$ = pnorm (x, μ, σ)
\# $P(X > x)$ = 1 − pnorm (x, μ, σ)

\# To generate random number from a normal distribution (in Random numbers)
The r code is rnorm (n, μ, σ)

1)A) A Random variable X follow normal distribution with mean = $\mu = 12$, & S.D = 6 = 3. Find

i.) $P(X \le 15)$
ii.) $P(10 \le X \le 13)$
iii.) $P(X > 14)$
iv.) Generate 5 observations (RANDOM NUMBERS)

Ans

(i) p1 = pnorm(15, 12, 3)
\> p1
[1] 0.8413447
\> cat ("P(X <= 15) =", p1)
\> P (X < = 15) = 0.8413447.

ii.) P2 = pnorm (13, 12, 3) − Pnorm (10, 12, 3)
\> P2
[1] 0.3780661

```
> 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

> rnorm (5,12,3)
[1] 14.98663   12.51618   13.14904   14.98075  19.
```

2) X follows normal distribution with
$\mu = 10$ $\sigma = 12$. Find = ?

i) $p(x \leq 7)$
ii) $p(5 < x < 12)$
iii) $p(x > 12)$
iv) Generate 10 Random observation
v) Find k such that probability $p(x<k) = 0.4$.

Ans i)
```
p1 = pnorm (7,10,2)
p1
[1]  0.0668072
> cat ("p(x<=7) =", p1)
p (~~x<~~ x<=7) = 0.0668072
```

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

```
> cat ("P( X < X < 12 ) =", p2 )
P( 15 < X < 12 ) = 0.8351351
```

iii)
```
p3 = 1 - pnorm (12,10,2)
p3
[1]     0.1586553
> cat ("P(P(X > 12) =", p3 )
P(P(X>12) = 0.1586553
```

```
> rnorm (10,10,2)
[1]  11.986324   12.155504   7.130492
     10.519359   5.637988    10.959963
     7.988735    9.336612    8.115370
     13.052133
```

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

1) Generate 5 Random number from a normal distribution with mean = 15 & S.D = 4

Find . sample mean, median, S.d and print (cat)

4) $X \sim N(30,100)$ ($6^2 = 6 = 10$) Find only $P_1 . P_2$ -- (no cat command)

52.

Ans

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

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

> p3 = pnorm (35, 30, 10) - pnorm (25, 30, 10)
> p3
[1] 0.3829249
> qnorm (0.6, 30, 10)
[1] 32.53347
```

B) Ans

```
> x = rnorm (5, 15, 4)
> x
[1] 16.43602  12.58878  16.91096  17.37383
    15.01140
> am = mean (x)
> am
[1] 15.67034
> med = median (x)
> me
[1] 16.43602

> varience = (n-1) * var (x) / n
> varience
[1] 2.9636
```

```
> sd = sqrt (varience)
> sd
[1] 1.72731
> cat (" sample mean is =", am)
sampr mean is = 15.67034
> cat (" sample median is = ", me)
sample mean is = 15.67034
> cat ( " sample is median is =", me)
sample median is=16.43602
> cat ( " sample sd is = ", sd)
sample sd is = 1.72731
```

3.) plot the standard normal graph.

Ans
```
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.

Q.1) Test the Hypothesis $H_0: \mu = 15$, $H_1: \mu \neq 15$. 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 significance.

m → mean   o → population

σ → s.d   n → sample size.

Sol.

```
> mo =15
> mx =14
> sd = 3
> n =400
> zcal (mx=mo)/(sd/(sqrt(n)))
> zcal
[1] -6.666667
> cat(" calculate value at z is =,")
calculate value at z is =,
> cat ( " calculate value at z is = -6.666667
> pvalue = 2* (1- pnorm (abs (zcal)))
> pvalue
[1] 2.61679 be -11
```

Since p value is less than 0.05 we reject H_0. $\mu = 15$.

2) Test the Hypothesis $H_0: \mu = 10$, $\mu \neq 10$
Random sample of size 400 is drawn with sample 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 of z is =", zcal)
calculate =2 * (1-pnorm (abs (zcal)))
>pvalue
[1] 0.07544036.
```

3) Test the Hypothesis No proportion of smokers in a college is 0.2 a sample is collected and sample proportion is calculated as 0.125. Test the Hypothesis at 5% Level of significance (sample size = 400)

P → population
p → sample.

Solution : > p = 0.2

> P = 0.125

> n = 400

> Q = 1-P

> zcal = (p-P) / (sqrt (P*Q/n))

> zcal

[1] -3.75

Q.4.) Last year farmers lost 20% of theirs crops. a Random sample of 60 fields are collected wi found that 9 field crops are insect popula Test the Hypothesis at 1% Level of significanc

p - 0.2 (capital)

P - 9/60

Soln : > p = 0.2

> P = 9/60

> n = 60

> 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.

5) Test the Hypothesis Ho: $\mu = 12.5$ 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
> mo = 12.5
> t= (mx - mo) /(sd / sqrt (n))
> t
[1] -8.894909
R> pvalue = 2*(1- pnorm (abs(t)))
> pvalue
[1] 0
```

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

PRACTICAL NO-06

Aim:- large sample test.

Q.1.) Let the population mean (the amount spent by customer in a Restaurent in 250. A sample of 100 customer selected. Sample mean is calculated as ~~275~~ 272 Rs and SD as 30 Rs. Test the hypothesis that population mean is 250 or not at 5% level of significance.

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

Q.1) Solution:-

~~Ho~~ = Ho : μ = 275 against $H_1 : \mu \neq 275$

```
> mo = 250
> mx = 275
> n = 100
> sd = 30
> zcal = (mx - mo) / (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
```

∴ p value = 0 < 0.05 ~~we~~ are reject Ho at 5% level of significance.

Q.2) 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))
> zcal
[1] -3.952847
p value = 2 * (1- pnorm (abs (zcal)))
>pvalue
[1] 7.72268e - 05
```

∴ The value is less 0.01 we reject

Q.3) 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 Ho: $\mu_1 = \mu_2$ against H1: $\mu_1 \neq \mu_2$ at 5% level of significance population.

Q.4) The Study of Noise level in two Hospitals is given below. Test the claim that the two Hospitals have same level of Noise at One (1%) level of significance.

	Hospital (A)	Hospital (B)
size	84	34
mean	61.2	59.4
S.D	7.9	7.5

Q.5) 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.

Q.3) $H_0: \mu_1 = \mu_2$ as $H_1: \mu_1 \neq \mu_2$

> n1 = 1000

> n2 = 2000

> mx1 = 67.5

> mx2 = 68

> sd1 = 2.5

> 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
> cat (" pvalue : ", pvalue)
> pvalue : 2.417564e-07
```

∴ Rejected.

```
> 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 ("pvalue :", pvalue)
> pvalue : 0.2450211
```

Scanned with CamScanner

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

Q.5)
```
> n1 = 600
> n2 = 900
> p1 = 400/600
> p2 = 450/900
> zcal = (p1 - p2) / sqrt (p* q*(1/n1+ n2))
> zcal
[1] 6.381534
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 1.733222e -10
```
∴ value is less than 0.01 the value is rejected.

Q.6) First sample size.

$H_0 : p_1 = p_2$ as $H_1 : p_1 \neq p_2$

$n_1 = 200$
$n_2 = 200$
$p_1 = 44/200$
$p_2 = 30/200$

$n_1 = 200$
$n_2 = 200$
$p_1 = 44/200$
$p_2 = 30/200$
$P = (n_1 * p_1 + n_2 * p_2) / (n_1 + n_2)$
$q = 1 - p$

```
zcal
[1] 1.802741
> pvalue =2*(1-pnorm(abs(zcal)))
pvalue
[1] 0.07142888
```

Accept: greater than 0.05.

PRACTICAL 7

(1) Topic: small sample Test.

Q.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 a population with average ~~meth~~ marks 66.

Soln.:

$H_0 : \mu = 66$

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

One sample test.

data: x.

t = 68.319, df = 9, p-value = 1.558e-13

alternative hypothesis: true mean is not. equal

95 percent confidence interval.

65.65171 70.14829.

Sample estimates.

mean of x.

67.9

Since p-value is less than 0.05 we reject hypothesis at 5% level of significance

```
> los = 0.05
> pvalue = 1.558e-13
```

```
>if (pvalue > 0.05) {cat ("accept Ho ") {
else {cat (" reject Ho ") {
reject Ho>
```

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

group 1 = 18, 22, 21, 17, 20, 17, 23, 23, 20, 22, 21,

group 2 = 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

Soln: 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 = 2.2513 , df = 16.316, p-value = 0.03798.
alternative hypothesis : true difference means, it is
not equal to 0.
95 present 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 Ho") {
```

else {cat ("reject H0") }.
reject H0>
Since p - value is less than 0.05 we reject hypothesis at 5% level of significance.

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

Soln: H0: There is no significant different for sa before and after the compaign.

```
> 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 is
means is greater than 0

95 percent confidence interval :
 -6.035547   Inf
sample   estimates :
   -3.5
> p - value = 0.9806
```

```
> if (p-value > 0.05) { cat ("accepts Ho") }
else { cat ("reject Ho") }
accept Ho>
```

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

iv) Two medicines are applied to the two patients respectively.

group 1 : 10, 12, 13, 11, 14

group 2: 8, 9, 12, 14, 15, 10, 9

Is there any significant difference between two mediciens.

Solution: Ho: There is no significant difference b/w medicines of two groups.

```
> 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

~~at to.~~ alternative hypothesis : true difference in mean is not equal to 0.

95 percent confidence interval:

-1.781171, 3.781171

sample estimates:

mean of x and y:

P - value = 0.4406

> if (p-value > 0.05) {cat ("accept Ho")}
else {cat ("reject Ho")}
accept Ho>

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

Q.5) The following are the weights before and after the diet program. Is the diet program effective.

Before: 120, 125, 115, 130, 123, 119,
After: 100, 114, 95, 90, 115, 99.

Solution: Ho: There is no - significance difference.

>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 = 5, p-value = 0.9963
4.3458

alternative hypothesis: True difference in means is less than 0
95 percent confidence interval:
-Inf 24.0295

sample estimates:
mean of the difference
= 19.83333

```
> p-value = 0.9963
> if (pvalue > 0.05) { cat ("accept H0") } else { cat ("reject H0") }
accept H0 >
```

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

PRACTICAL 8

TOPIC : Large and small sample test.

Q.1) The arithmetic mean of a sample of 100 items from large population is 52. If the standard is 7 the hypothesis that the pol population mean is against the alternative it is more than 55% at

Ans: $H_0 : \mu = 55$

$H_1 : \mu \neq 55$

```
> n = 100
> mx = 52
> mo = 55
> sd = 7
> zcal = (mx - mo) / (sd / (sqrt (n)))
> zcal
[1] -4.28514
> cat ("calculated z value is = "zcal )
calculated z value is = -4.285714.
> p-value = 2* (1- pnorm (abs (zcal)))
> pvalue
[1] 1.82153e-05
```

Since p value is less than 0.05 we reject the hypothesis 5% level of significance

Q.2) In a big city 350 out of 700 males are found be smokers. does information supports that exactly half of the males in the city are smokers. Test at 1%. los.

Ans H0 : μ

```
> P = 0.5
> Q = 1-P
> p = 350 / 700
> n = 700
> zcal = (p - P) / (sqrt (P * Q / n))
> cat (" calculated z value is = ", zcal)
calculated z value is = 0
> pvalue = 2 * (1 - pnorm (abs (zcal )))
> pvalue
[1] 1
```

Since pvalue is greater than 0.05 we accept hypothesis at 1% level of significance.

3) Thousand article 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 are not.

Ans H0 : $p_1 = p_2$ as H1: $p_1 \neq p_2$

```
> n1 = 1000
> n2 = 1500
> p2 = 0.01
> p = (n1 * p1 + n2 * p2) / (n1 + n2)
> p
[1] 0.014
```

```
> p
[1]   0.014
> q = 1-p
> q
[1]   0.986
> zcal = (p1-p2) / sqrt (p*q * (1/n1 + 1/n2))
> zcal
[1]   2.084842
> cat ("calculated zvalue is =", zcal)
   calculated zvalue is = 2.084842
> pvalue = 2* (1- pnorm (abs (zcal)))
> pvalue
[1]   0.03708364
```

Since pvalue is less than 0.05 we reject the hypothesis at 5% level of significance

Q.4) A sample of size 400 was drawn at a sample mean is Test at 5% LOS that the sample comes from a population with mean 100 and variance 64.

Ans: $H_0: \mu = 100$

```
> mx = 99
> mo = 100
> sd = 8
> n = 400
```

```
>zcal = (mx - mo) | (sd / (sqrt (n)))
>zcal
[1] -2.5
> p-value = 2 * (1 - pnorm (abs(zcal)))
> p-value
[1] 0.01241933
```

Since p-value is less than 0.05 we reject the hypothesis at 5% level of significance.

) 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.

Ans: Ho: $\mu = 66$

```
> 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
```

Scanned with CamScanner

Since p-value is less than 0.05 we reject hypothe- at 1% level of significance.

Q.6) Two random samples were drawn from 2 normal populations and there values are (A) → 66, 67, 75, 76, 82, 88, 90, 92. (B) → 64, 66, 74, 78, 82, 85, 87, 92, 93, 95, Test ~~at 5% LOS that the sample comes from a~~ whether the population have the same variance at 5% LOS

Ans
$H_0: \sigma_1 = \sigma_2$

> x=c(66, 67, 75, 76, 82, 84, 88, 90, 92)
> y= c (64, 66, 74, 78, 82, 85, 87, 92, 93, 9
> var. test (x, y)

F test to compare to variance.

data : x and y.
F = 0.70686 , num df = 8 , denom df = 10,
p-value = ~~0.634~~ 0.6359.
alternative hypothesis : true ratio of 95 perce ~~is not equal to~~ confidence interval:
0.1833662 3.0366393
sample estimates.
ratio of variance
0.7068567.

> p-value = 0.6359.
> if (p-value > 0.05) { cat ("accept H0") }
else { cat ("reject H0") }

accept Ho >

since p-value is greater than 0.05 we accepts hypothesis at 5% level of significance.

Q.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 the wheter the difference b/w population and sample mean is significantly different ?

Ans
~~Ho=~~ Ho: μ = 1200

> mo = 1200
> mx = 1150
> n = 100
> sd = 125
> zcal = (mx - mo) / (sd / (sqrt (n)))
> zcal (" calculated z value is = -4
> pvalue = 2 * (1 - pnorm (abs (zcal)))
[1] 6.334248 e - 05

Since pvalue is less than 0.05 are reject hypothesis

Q.8) From each of two consignment of apples, a sample of size 200 is drawn and number of bad apples are counted. Test whether the proportion of rotten ~~appl~~ apples in two assignments are significantly different at 1% Los ?

Sample size	No. of bad apples
200	44
	56

Scanned with CamScanner

<u>Solution</u> $H_0: \mu = p_1 \neq p_2$

```
> n1 = 200
> n2 = 300
> p1 = 44/200
> p2 = 56/300
> p = (n1*p1 + n2*p2) / (n1 + n2)
> p
[1] 0.2
> q = 1-p
> zcal = (p1 - p2) / sqrt (p*q*(1/n1 + 1/n2))
> zcal
[1] 0.9128709.
> cat (" calculated z values is =", zcal)
calculated = value is = 0.9128709
> pvalue = 2*(1- pnorm (abs (zcal)))
> pvalue
[1] 0.3613104
```

Since pvalue is greater than 0.05 we accept the hypothesis.

PRACTICAL NO-09

chi - square tests and ANOVA (Analysis variances

1) use the following data test whether the condition of home and the condition of child are indipendent or not.

→ condition of ~~b~~Home.

clean	clean	70	Dirty	50
fairly clean		80	~~50~~	20
Dirty		35	~~20~~	45
			4	

→ Ho: condition of Home and child are indipendent

$x = c(70, 80, 35, 50, 20, 45)$

$m = 3$

$n = 2$

y = matrix (x, nrow = m , ncol = n)

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

pv = chisq.test (x).

pv

pearson's chi squared test.

data y

X - squared = 25.646 , df = 2 , pvalue = 2.698e..

∴ Ho is rejected , since pv is less than 0.05

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

vaccine.

Disease	Affected	Not Affected.
Affected	70	46
Not Affected	35	37

Ho : condition of disease and vaccination are

→ x = c(70, 46, 35, 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

X - squared = 2.0275 , df = 1 , pv = 0.1545

Ho : is Accepted , since pv is more than 0.05.

Q.3) perform ANOVA for the following data.

Type	observation
B	50, 52.
C	53, 55, 53
D	52, 54, 54, 55.

> H0: 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 (b) = x1, b2 = x2, b3 = x3, b4 = x4))
names (d)
[1] "value" "ind"

> oneway.test (value ~ ind, data = d, var. equal =

one-way analysis of means
data : values ind
F = 11.735, num df = 3, denom df = 9, p-value

> 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 ''

22

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

Type	Life.
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.

→ H0: The average life of A, B, C, D are 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, b2 = x2, b3 = x3, b4 = x4))
    names (d)
> oneway.test (value and ind, data = d, var.equal = T)
            one - way analysis of means.
data:
    F = 6.8445 , num df = 3, denom df = 20,
    p-value = 0.002344.
> anova = aov (values ~ ind , data = d)
    summary (anova).
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
trt	3	91.44	30.479	6.845	0.00235
Residuals	20	39.06	4.453		

Q.5.) How to import a . csx file in R software.

→ x = read . csv ("c:/ users / admin / Desktop / stats . csv")

print(x)

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

```
> am = mean (x $ stats)
> am
[1] 37
n = length (x$ stats)
~~>am~~     sd = sqrt (n-1) * var ( x$ stats ) /n )
~~[1] 37~~  sd
      [1]  12. 64911
      am = mean (x$ maths)
      am
      [1] 39. 4
      sd = sqrt ((n-1) * var (x$ maths) /n )
      sd
```

r.a.

[1] 15.2

cov (x$stats ; x$ maths

~~[1]~~ 0. ~~830613~~

[1] 0.830618.