PRACTICAL :- 01

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 a free software

Q1) Solve

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

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

4. $\sqrt{4^2}+5\times 3+16$

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

5. round off $46\div 7+9\times 8$

```
> 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] 1.50 0.40 1.75 1.00
> 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 16 125
```

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

Q6) Marks of statistics of CS Batch B

x = c(58, 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 = 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

PRACTICAL :- 2

* Topic :- Probability distribution.

1) Check whether the foll 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 pmf then $\Sigma P(x) = 1$

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

$\therefore 0.1 + 0.2 + (0.5) + 0.4 + 0.3 + 0.5$

$\therefore 1.0$

$\therefore P(2) = -0.5$ it ~~can~~ be a probability mass function

$\therefore P(x) \geq 0 \quad \forall, x$

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 $\Sigma P(x) =$

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

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

$= 1.1$

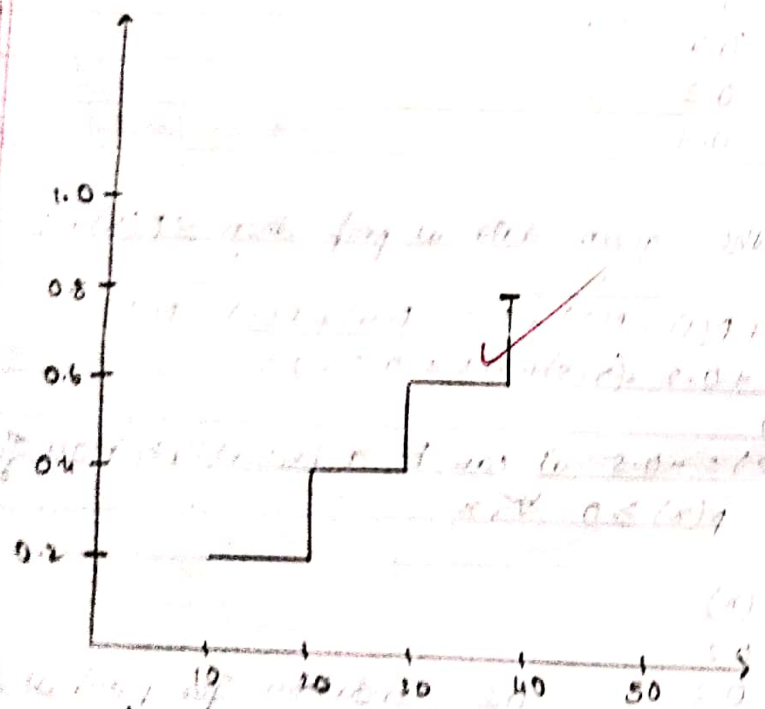
#8

∴ The given data is not a pmf because the $p(x) \neq 1$.

Q2) Find the c.d.f for the follo 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.

$$F(x) = \begin{cases} 0 & x < 10 \\ 0.2 & 10 \le x < 20 \\ 0.4 & 20 \le x < 30 \\ 0.75 & 30 \le x < 40 \\ 0.90 & 40 \le x < 50 \\ 1.0 & x \geq 50 \end{cases}$$




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

36

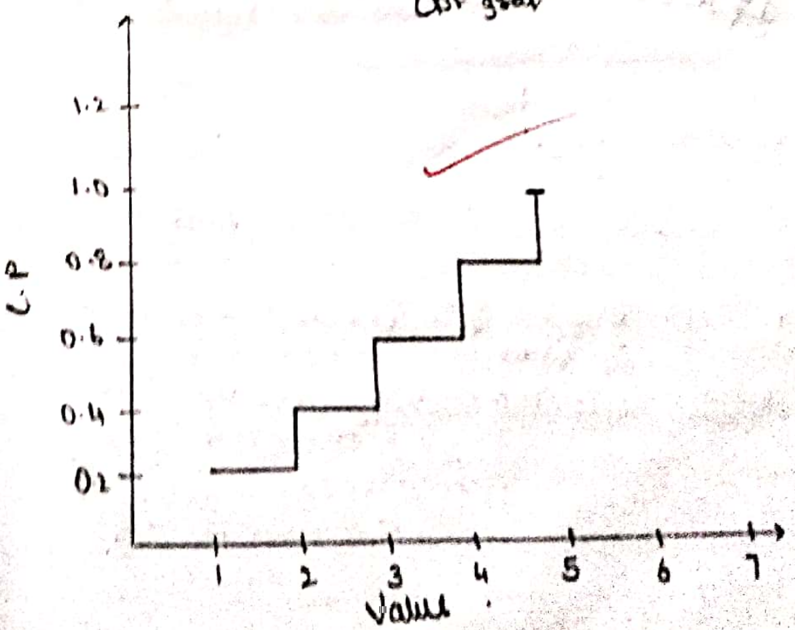
(ii)

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

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

```
> 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")
```

CDF graph



Q3) Check that whether the following is pdf or not

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

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

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

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

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

$\because$ The $\int_0^1 f(x) \neq 1$ $\therefore$ It is not a pdf.

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

$\int_0^1 f(x)$

$= \int_0^1 3x^2 \quad \therefore 3\int_0^1 x^2$

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

$\therefore x^3$

$\therefore 1$

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

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] 0.427445
[iii] - 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] 0.0027815)
4) dbinom (6,12,0.25)
[1] 0.04014945.

PRACTICAL:- 3

* Topic:- Binomial distribution

P(x = x) = dbinom (x, n, p)
P(x ≤ 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 = 0.4.

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) at most 4 correct answer
(iii) More than 5 correct answer.

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

4) n = 12, p = 0.25 find the foll probabilities
(i) P(X = 5)
(ii) P(X ≤ 5)
(iii) P(X > 7)
(iv) P(5 < X < 7)

5) The probability of salesman making a sale to customer 0.15. Find the probability of
(i) No sales out of 10 customer.
(ii) More than 3 sales out of 20 customer.

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

7) X follows binomial distribution with n = 10, p = 0.3 plot the graph of pmf and cdf.

Answer

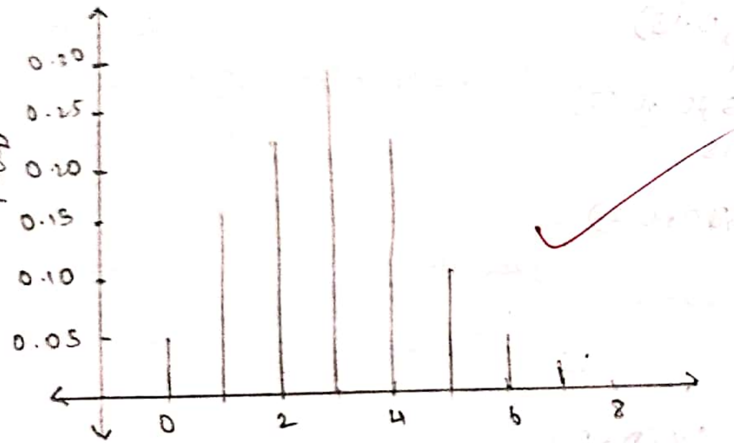
5)
```
dbinom (0,10,0.15)
[1] 0.1968744
>1- dbinom (3,20,0.15)
[1] 0.3522748
```

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

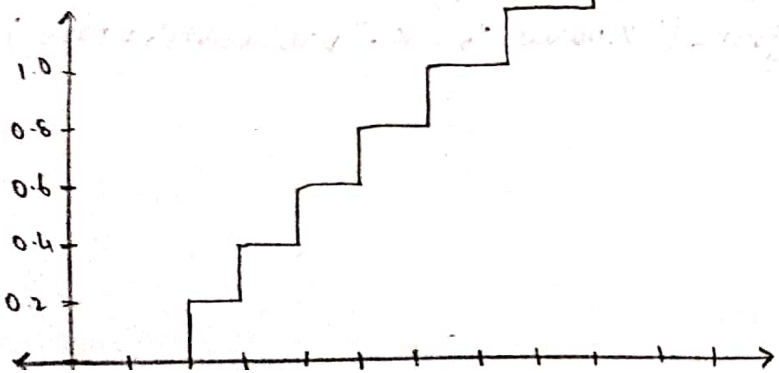
7)
```
>n = 10
>p = 0.3
>x = 0:n
>prob = dbinom (x,n,p)
>cumprob = pbinom (x,n,p)
>d = data.frame("x values", = x,"probability", = prob)
>print(d)
```

	x values	probabilities
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

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) To generate random no from a normal distribution (n random variables) the R code is rnorm(n, μ, σ).

Q) A random variable x follows normal distribution with mean μ = 12 and S.D = σ = 3 find
(i) $P(x \le 15)$
(ii) $P(10 \le x \le 15)$
(iii) $P(x > 14)$

CODE:-

```
> P1 = pnorm (15,12,3)
> P1
[1] 0.8413447
> cat ("P(x<=15) =", P1)
P(X<=15) = 0.8413447
> P2 = pnorm (15,12,15) - pnorm (10,12,3)
> P2
[1] 0.5780661
> cat ("P(10 <= x <=13) =", P2)
P(10 ≤ x ≤ 15) = 0.3780661
> P3 = 1 - pnorm (14,12,3)
> P3
```

```
> cat ("P(x>14)= " P3)
P(x>14) = 0.2524925
> P4 = rnorm (5,12,3)
> P4
[1] 15.254713    16.548505   11.280515   6.419944   12.272960
```

2) X follows normal distribution with $\mu = 10$, $\sigma = 2$. Find (i) $P(x \le 7)$ (ii) $P(5 < x < 12)$ (iii) $P(X > 12)$ (iv) generate 10 observations (v) find K such that $P(X < K) = 0.4$

Code :-

```
> a1 = pnorm (7,10,2)
> a1
> a2 - pnorm(5,10,2) - pnorm (12,10,2)
> a2
[1] -0.8351351
> a3 = 1 - pnorm (12,10,2)
> a3
[1] 0.1586553
> a4 = qnorm (10,10,2)
> a4
[1] 11.608931   9.920417   12.637741   8.073354
    8.721350    9.193776   9.366824    11.707106
    9.537584    10.715006
> a5 = qnorm (0.4,10,2)
> a5
[1] 9.493306.
```

Q.3) Generate 5 random no from a normal distribution $\mu = 15$, $\sigma = 4$ find sample mean, median, S.D and print it.

Code :-

```
> rnorm (15,15,4)
[1] 10.7649    7.793249   9.953444   13.345904
    17.509668
> am = mean (x)
[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] 3.33163
> cat ("S.D is = ", SD)
SD is = 3.331613
```

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

(i) $P(x \le 40)$

(ii) $P(x > 35)$

(iii) $P(25 < x < 35)$

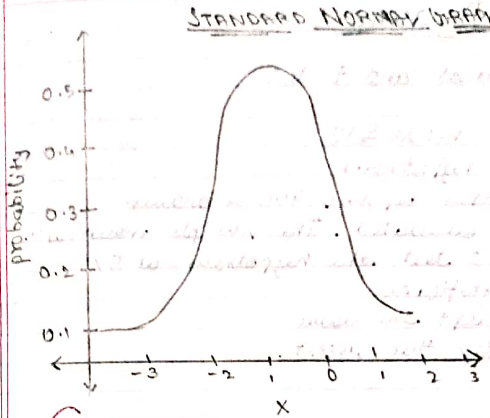
(iv) find k such that $P(x < K) = 0.6$

```
> 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")
```


STANDARD NORMAL GRAPH



II

PRACTICAL:-5

* TOPIC :- Normal and t-test

1) Ho: $\mu = 15$ $H_1: \mu \neq 15$
Test the hypothesis
Random sample of size 400 is drawn and it is calculated. The sample mean is 14 and S.D is 3 test the hypothesis at 5% level of significance.
0.05 > accept the value.
0.05 < less than reject.

```
>mo = 15
>mx = 14
>sd = 3
>n = 400
>zcal = (mx - mo)/(sd/(sqrt(n)))
>zcal
[1] -6.66667
>cat("calculated value of z is = ",zcal)
calculated value of z is = -6.666667
>pvalue = 2*[1-pnorm(abs(zcal)))
>pvalue
[1] 2.616796 e-11
```

∵ The value is less than 0.05 we will reject the value of Ho = μ = 15

2) Test the hypothesis Ho: $\mu = 10$ against $H_1: \mu \neq 10$
A random sample size of 400 is drawn with sample mean = 10.2 & S.D = 2.25 Test the hypothesis at.

```
>mo = 10
>n = 400
>mx = 10.2
>sd = 2.25
>zcal = (mx - mo)/(sd/(sqrt(n)))
>zcal
[1] 1.77778
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.07544036
```

∵ The value pvalue is greater than 0.05
∴ The value is accepted.

3) Test the hypothesis Ho: proportional of smokers in colleage is 0.2
A sample is collected and calculated the sample proportional as 0.125. Test the hypothesis at 5% level of significance (sample size is 400)

```
>P = 0.2
>p = 0.125
>n = 400
>Q = 1 - P
>zcal = (p-P)/(sqrt(P*Q/n))
>cat ("calculated value of z is = ",zcal)
[1] calculated value of z is = -3.75
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
```

[1] 0.0001768346 (Reject).

4) Last year farmers lost 20% of their crops A random 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 = 66
> z cal = (P-p) / (sqrt (p* Q/n))
> z cal
[1] -0.9682458
> pvalue = 2* (1-pnorm (abs (z cal)))
> pvalue
[1] 0.3329216
```

∴ The value is 0.1 so value is accepted.

5) Test the hypothesis $H_0: \mu = 12.5$ from the foll sample at 5% level of significance:

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



```
> variance = (n-1) * var (x)/n
> variance
[1] 0.019521
> sd = sqrt (variance)
> sd
[1] 0.1397176
> mo = 12.5
> t = (mx - mo) / (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.

43

PRACTICAL:- 6

* AIM:- 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 the population mean is 250 or not on 5%. level of significance.

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.

1) Solution:-
>mo = 250.
>mx = 275
>sd = 30
>n = 100
>zcal = (mx - mo) / (sd / (sqrt (n)))
>cat ("calculated value of z is =", zcal)
[1] calculated value of z is = 8.33333
>pvalue = 2 * (1- pnorm (abs (zcal)))
>pvalue
[1] 0
∴ The value is less than 0.05 we will reject the value of $\mu_0 = \mu = 250$

44

2) Soln :-
>P = 0.8
>Q = 1-P
>p = 750/1000
>n = 1000
>zcal (p-P) / (sqrt (P * Q/n))
>cat ("calculated value of z is :", zcal)
[1] calculated value of z is : -3.952847
>pvalue = 2 * (1- pnorm (abs (zcal)))
>pvalue
[1] 7.72268 e-05
∴ The value is less 0.01 we reject!

3) To random sample of size 1000 & 2000 are drawn from two population with same S.D 2.5 the sample mean are 67.5 and 68. Test the hypothesis $\mu_1 = \mu_2$ at 5%. LOF significance.

4) A study of noise level in 2 hospital is given in given below test the claim that 2 hospital have same level of noise at 1%. level of significance.

Hos A	Hos B
84	34
61.2	59.4
7.9	7.5

3) > 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
> pvalue = 2* (1-pnorm(abs(zcal)))
> pvalue
[1] 2.417564e-07 ∴ (Rejected).

4) > 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
> pvalue = 2* (1-pnorm(abs(zcal)))
> pvalue
[1] 0.2450211

Pratical :- 07

* Topic :- Small sample test.

Q1) 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 marks 66.

→ Ho : μ = 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 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.8) Two groups of student score the foll marks. Test the hypothesis that there is no significance difference between the 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 betw two grps :-

```
> 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 & y
t = 2.2513, df = 16.316, p-value = 0.03798
alternative hypothesis :- true diff means it 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 Ho")}
  else {cat ("reject Ho")}
reject Ho>
```

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

Q.9) The sales data of 6 shops before and 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 effective or not.

→ Ho :- There is no significant difference of sales before and after the campaign

```
> 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 :- the diff in mean is greater than 0.
95 percent confidence interval
-6.035547 in
sample ~~estimates~~ :
mean of the difference
-3.5

```
> pvalue = 0.9806
> if (pvalue > 0.05) {cat ("accept Ho")}
  else {cat ("reject Ho")}
accept Ho>
```

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

Q.4) Two medicines are applied to the two grp of patients respectively.

grp1 :- 10,12,13,11,14.

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

Is there any significant difference between two medicines

Sol^n :- Ho: There is no significant diff between two 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
alternative hypothesis :- true difference in
   mean is not equal to 0.
95 percent confidence interval:
   -1.781171    3.781171
Sample estimates:
Mean of x & y
    12    11

> pvalue = 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.

(5.)

Pratical :- 08

* Topic :- large and small sample test

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 means is 55 against the alternative it is more than 55 at 5% LOS.

→ Ho : $\mu = 55$

H1 : $\mu \neq 55$

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

Since p value is less than 0.05 we reject the hypothesis at 5% LOS.

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

→ Ho: U

```
> 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% LOS.

Q) 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.

→ Ho: $P_1 = P_2$ vs $H_1: P_1 \neq P_2$

```
> n1=1000
> n2=1500
> p1=0.02
> p2=0.01
> p=(n1*p1+n2*p2)/(n1+n2)
> p
[1] 0.014
> q=1-p
> q
[1] 0.986
> zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
> 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 z value is =",zcal)
calculated z values 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.

Q4) 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.

→ Ho: μ = 100

```
> mn = 99
> mo = 100
> sd = 8
> n = 400
> zcal = (mn - mo) / (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 the hypothesis at 5% los.

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

→ Ho : μ = 66

```
> x = c(63, 63, 68, 69, 71, 72)
> t.test (x)
        One sample t-test )
data : x
```


t = 71.7.94, df = 6, p-value = 5.522e-09
alternative hypothesis the mean is n

CR

PRACTICAL :- 9

• 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 listen

Data :- 168, 3.35, 2.50, 6.25, 8.25

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

data 95 confidence interval

0.01020622 1.0000

Sample estimates
probability of success 0.2 .

2) The scores of 8 students in reading before & after lessor are as follows Test whether there is effect of reading.

→

Inspector	1	2	3	4	5
caliper 1	0.265	0.268	0.266	0.269	0.264
caliper 2	0.263	0.262	0.270	0.271	0.260

caliber 1 & caliber 2 are are .

Code :-

```
> x <- c (0.263, 0.268, 0.266, 0.267, 0.269,
                                   0.264)
> y < - c (0.263, 0.262, 0.261, 0.271, 0.260)
> will con-test (x, y, alternative = "greater")
        wilcox. rank test .
  data : x and y
  W = 24 , p = 0.197
```

alternative hypothesis
∴ three location shift is than 0,

∴ p value is greater than 0.05 we accept it .

2) The scores of 8 students in reading before and after lessons are as follows.

student No	1	2	3	4	5	6
score before	10	15	16	12	09	07
score after	13	16	15	13	09	10

Code :-

```
> b <- c (10, 15, 16, 12, 09, 07, 11, 12);
> a <- c (3, 16, 15, 13, 08, 13, 10)
> b <- b - a;
wilk w - x test (D, alternative "greater")
```

will wox - or signed rank list with will continuity.

In will w - x test default alternative - "greater", cannot compute exact p - value value with ties

∴ p - value is greater than 0.05

∴ We accept it :-

—— x ——

PRACTICAL :- 10

* AIM :- chi square test & ANOVA . (analysis of variances)

Q) Use the foll data to test whether the condition of home & condition of child are not .

child	Home clean	Dirty
clean	70	50
fairy 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
> matrix (x, nrow = m , ncol = n)
> y .
```


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

```
> Pu = chisq . test [y]
> Pu
```

Pearson : chi - squared test .

data : y

X - square = 25 . 646

df = 2

p - value = . 2.6986e 06 .

∴ They are dependent .

∴ Test the hypothesis that variation & disease are indepent or Not .

Q) Test the hypothesis that vaccination & disease are independent or not

vaccine	Affected	Not
disease	70	46
affect	35	37

Ho :- disease & vaccine are independent

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

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

> PV = chisq.test (y)
> PV
Pearson's chi squared 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

" They are Independent .

14/03/20