PRACTICAL No - 1

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

> 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) $\sqrt{4^2+5\times 3+7/6}$

> sqrt (4^2 +5 * 3+ 7/6)

[1] 5.671567

Q1] 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)
[1] 4 9 10 21

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

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

> c(2,3,5,7) ^2
4 9 25 49

> c(4,6,8,9,4,5) ^ c(1,2,3)
4 36 512 9 16 ~~~~ 215

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

Q.3 > x - 20 > y - 30 > z = 2
> x^2 + y^3 + Z
[1] 27402
> sqrt (x^2+y)
[1] 20.73644
> 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,] 1 5
[2,] 2 6

TSO

[3,
[4
Q5 fin
y -
> x
6
> x
[
[
[
y <
>
[
[
[
>

[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 = $\begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 5 \end{bmatrix}$

> 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,
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

PRACTICAL - 02

Topic :- Probability distribution

① check whether the following are pmf 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

$\Sigma p(x) = 1$

$= p(0) + p(1) + p(2) + p(3) + p(4) + p(5)$

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

$= 1.0$

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

$p(x) \geq 0$, x

②

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

eso

If the Given data is pmf then

$\Sigma p(x) = 1$

$= p(1) + p(2) + p(3) + p(4) + p(5)$

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

$= 1.1$

$\neq 1$

The Given data is not a pmf because $\Sigma p(x) \neq 1$

(3)

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

The condition for pmf is

1) $p(x) \geq 0 \ \forall x$ satisfy
2) $\Sigma p(x) = 1$

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

$= 1$

$\therefore$ The Given data is pmf.

code:-

```
prob= c(0.2, 0.2, 0.35, 0.15, 0.1)
Sum(prob)
1
```

Div.:________________ Roll No.:__________


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

```
prob = c (0.15 , 0.25 , 0.1 , 0.2 , 0.2 , 0.2, 0.1)
sum (prob)
= 1
cumsum (prob)
= 0.15, 0.4 , 0.5 , 0.7 , 0.9 , 1
x = c (1, 2 , 3, 4, 5, 6)
plot (x, cumsum (prob) , "s" , xlab = value, ylab = "cummlative
Probability . main = 'CDF' , col = "brown")
```

Scanned with CamScanner

[illegible]

Q. [illegible] the following [illegible] the [illegible]

x	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
$f(x)$	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

Find

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

$F(x) = 0, \quad x < 1$

$= 0.15, \quad 1 \le x < 2$

$= 0.4, \quad 2 \le x < 3$

$= 0.5, \quad 3 \le x < 4$

$= 0.7, \quad 4 \le x < 5$

$= 0.9, \quad 5 \le x < 6$

$= 1, \quad x \ge 6$

Q.3 check whether the following is pdf or not.

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

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

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

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

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

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

$$= 3 - 1$$

$$= \underline{\underline{2}}$$

$\therefore f(x) = 3 - 2x$ is not pdf.

(ii) $f(x) = 3x^2$

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

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

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

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

$$= 1$$

$\therefore f(x)$ is a pdf.

180 Practical No. 3

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 in hundred trials with $p = 0.1$

2) Suppose there are 12 mcq. Each question has 5 options out of which is correct. Find the probability of having exactly 4 correct answers.

ii) Almost 4 correct answer.
ii) more than 5 correct answer.

(

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

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

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

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

x

0.131865 3

(2) (i) dbinom (4, 12, 0.2)

0.1328756

(ii) pbinom (4, 12, 0.2)

0.4274445

(iii) 1 - pbinom (5, 12, 0.2)

0.0194 6528

(3) dbinom (4, 12, 0.2)

0.112

(3) dbinom (0:5, 5, 0.1)

0 - 0.59049

1 - 0.32805

2 - 0.7290

3 - 0.00510

4 - 0.0045

5 - 0.00001

(4) (i) dbinom (5, 12, 0.25)

0.1032414

ii) pbinom (5, 12, 0.5)

0.9455978

(iii) 1 - pbinom (7, 12, 0.5)

0.0027815

iv) dbinom (6, 12, 0.25)

0.04014945

(5)(i) dbinom (0, 10, 0.15)

0.1968744

(ii) 1- pbinom (3, 20, 0.15)

0.3522748

(6) qbinom (0.98, 30, 0.2)

9

(7) n=10

p=0.3

x =0:n

prob = dbniom (x,n,p)

comprob = pbinom (x, n, p)

d= data.frame (" ~~x~~ values "= x, " Probability " = prob)

Print(d)

plot (x, prob, "h")

plot (x, comprob, "s")

Scanned with CamScanner

5) The probability of salesman making a sale to customer is 0.15

find
(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 is 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 & cdf

Q

Practical No. 4.

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 numbers from a normal distribution (n random numbers), the R code is
rnorm(n, μ, σ)

Q1 A random variable x follows normal distribution with mean = 12. S.D = 3. Find: i) $P(x \le 15)$ ii) $P(10 \le x \le 13)$ iii) $P(X > 14)$ iv) Generate 5 random numbers

CODE:-

```
> P1 = pnorm (15, 12, 3)
  P1
[1] 0.8413447
> P2 = pnorm(13, 12, 3) - pnorm (10, 12, 13)
  p2
[1] 0.3780661
> p3 = 1 - pnorm (14, 12, 3)
  p3
[1] 0.2524925
> p4 = rnorm (5, 12, 3)
  p4
[1] 15.254723 . 16.548505   11.280515  8.419
12.272460.
```

Div: ______________ Roll No.: ______________

No. ______________

031

Q) 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) 10 observation. Also, find K such that $P(x < K) = 0.4$

Code:

p1 = pnorm (4, 10, 2)
p1
0.668072

p2 = pnorm (5, 10, 2) - pnorm (12, 10, 2)
p2
0.835135 1

p3 = 1 - pnorm (12, 10, 2)
p3
0.1566553

p4 = rnorm (10, 10, 2)
p4

11.608931	9.920417	12.37741	9.073354
8.721386	9.143725	0.366924	11.70710 6
4.537584	12.715006		

p5 = qnorm (0.4, 10, 2)
p5
9.493306

Scanned with CamScanner

③ . Generate 5 random numbers form a normal distribution
$\theta = 15$, $\sigma = 4$. find sample mean, meadian, SD
& print it.

Code :-

```
rnorm ( 15, 15, 4)
10.7649 .   7.793249        9.93344    13.345904
17.509608

> am = mean (n)
> am
11.87345
cat ("sample mean is - ", am)
sample mean = 11.87345
> me = median (n)
> me
10.76499
> cat ("sample median is - ", me)
Sample Median : 10.76499.
n = 5
> v = (n - 1) * var (n) /n
> v
11.69965
SD = Sqrt (v)
SD
3.33163
```

Scanned with CamScanner

cat ("SD :", SD)

SD = 3.33163

(19) X ~ (30, 100) , σ = 10

(i) P(x ≤ 0) (ii) P(x > 35) (iii) P(25 < x < 35)

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

```
> p1 = pnorm (40, 30, 10)
> p1
[1] 0.841344
> p2 = 1 - pnorm (35, 30, 10)
> p2
[1] 0.3829249
> p4 = qnorm (0.6, 30, 10)
> p4
32.53347.
```

PRACTICAL No-5

Topic:- Normal and t-Test.

1. $H_0: \mu = 15$. $H_1: \mu \neq 15$
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

```
>m0 = 15
>mx = 14
>n = 400
>Sd = 3
>zcal = (mx -m0) /( sd)/ (sqrt (n)))
>zcal
[1] -6.666667
>cat ("calculated value of z is = ", zcal)
calculated value of z is = -6.666667
>pvalue = 2 * (1 - pnorm (abs (zcal)))
>pvalue
[1] 2.616796e -11
```

∴ The value is less than 0.05 we will reject the value of $H_0 = \mu = 15$

2. Test the hypothesis $H_0 = \mu = 10$ against $H_1: \mu \neq$
A random sample size of 400 is drawn with sample mean = 10.2 and SD = 2.25.
Test the hypothesis at.

```
> m0 = 10
> n = 400
> mx = 10.2
> sd = 2.23
> zcal = (mx - m0) / (sd / (sqrt(n)))
> zcal
[1] 1.77778
> pvalue = 2 * (1 - pnorm (abs (zcal)))
> pvalue
[1] 0.07544036
```

∴ The value pvalue is greated than 0.05
∴ The value is accepted.

3. Test the hypothesis H0 : proportional of smokers in college 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] 0.000176 8346
```
Reject

4. last Year farmer's last 20% of their crops. A random sample of 60 fields are collected and it is found that 9 field crops are insect polluted. Test the hypothesis at 1% level of significance.

```
> 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 $H_0: \mu = 12.5$ 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
```

```
>variance = (n-1)* var(x)/n
>variance
[1] 0.0919521
>sd = sqrt (variance)
>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 No - 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 SD 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

Solution!:

```
>m0=250
>mx=275
>sd=30
>n=100
>zcal=(mx-m0)/(sd/(sqrt(n)))
>cat("Calculated value of z is=",zcal)
[1] calculated value of z is: 8.333333
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1]0
```

∴ The value is less that 0.05 we will reject the value of $H_0 = \mu = 250$.

2. Solution

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

3. To random sample of size 1000 & 2000 are drawn from two population with same SD 2.5 the sample means are 67.5 and 68 test the hypothesis $\mu_1 = \mu_2$ at 5% of significance

4. A study of noise level in 2 hospital is 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	7.8

3. Solution :-

>n1 = 1000
> n2 = 2000
>mx1 = 67.5
> mx2 = 68
> sd1 = 2.5
> sd2 = 2.5

380

```
>z cal = (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))
>zcal
[1] 1.16252
>pvalue = 2* (1-pnorm (abs (zcal)))
>pvalue.
[1] = 0.245021
```

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

Name : ____

Div.: ____ Roll No.: ____

PRACTICAL NO - 7 -

039

Topic :- Small sample Test

① 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 average 66

=> Ho : μ = 66

x = c (66, 63, 66, 67, 68, 69, 70, 71, 72)

t. test (x)

One sample t-test.

data: x

t = 68.319, df = 9, pvalue = 1.558e-13

alternative hypothesis.

True mean is not equal to 0 at ~~at~~

95 percent confidence interval.

65.6517 1 70.14829

Sample estimates.

Mean of x

67.9

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

② Two groups of student scorred the following marks. Test the hypothesis that there is no significant difference between 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

Ho : There is not difference b/w the 2 groups

```
>x = c(19,22,21,17,20,17,23,20,22,21)
>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 ~~con~~ confidence interval.

0.1628205 . 5.071795

sample estimates

mean of x mean of y

20.1 17.5

```
>pvalue = 0.03798
> if (pvalue > 0.05) {Cat ("accept ho")}
else {cat ("reject ho")}
reject Ho
```

Div.: Roll No.:

(PAIRED t - test)

3) The sales data of 6 shops before and after a special campaign are given below.

Before : 53, 28, 31, 47, 50, 42

After : 58, 29, 30, 55, 56, 45

Test the hypothesis that the campaign is effective or not.

H0: There is no significance difference of sales before and after campaign.

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.

MAD

4] 2 medicines are applied to two greater groups of patient respectively.

gr1 : 10, 12, 13, 11, 14

gr2 : 8, 9, 12, 14, 15, 10, 9

Is there any significance difference b/w 2 medicies.

Ho: 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 10.53333

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

Practical No. 8

Topic: Large and Small test.

1) Ho = μ = 55 H_1 : μ ≠ 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
```

: The value is less than 0.05 we will reject the value Ho = μ - 55

Ho : μ - 0.5

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

∴ The value of 1 is greater than 0.01 we accept the Hypothesis.

140

5) Ho : $P_1 = P_2$. or $H_1 : P_1 \neq P_2$

```
>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))
>zcal
[1] 2.084842
>pvalue = 2 * (1-pnorm (abs (zcal)))
>pvalue
[1] 0.03708364
```

∴ pvalue is less than 0.05 we reject the hypothesis

4] Ho : μ = 100

```
>mx = 99
>m0 = 100
>sd = 8
>n = 400
>zcal = ((mx - m0) / (sd / sqrt (n)))
>zcal
[1] -2.5
> pvalue = 2 * (1 - pnorm (abs (zcal)))
>pvalue.
```

8

Practical 9

[1] 0.01241933

∴ The value is less than 0.05 we reject the hypothesis.

2) Ho : μ = 66

\> x = c (65, 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

∴ p value is '0.05' less than we reject the hypothesis.

3) > n1 = 200

\> n2 = 300

\> p1 = 44/200

\> p2 = 56/300

\> p = (n1 × p1 + n2 × p2) / (n1 + n2)

\> p

[1] 0.2

, (p × q × (1/n1 + 1/n2)

```
> Zcol
[1] 0.9128709
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.3613104
```

∴ pvalue is less than 0.05 we rejet hypothesis.