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

[1] 4 6 10 14

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

[1] 4 9 30 14

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

[1] 4 9 25 49

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

[1] 1.50 0.40 1.75 1.00

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

[1] 4 9 10 21

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

[1] -2 -18 -8 -3

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

[1] 4 36 512 9 16 21 25

3

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

> x^2 + y^3 + z

[1] 27402

> ~~sqrt~~ sqrt (x^2+y)

[1] 20.73644

> x^2 + y^2

[1] 1300

Aim:- Basic of R Software

i) R is a software for statical analysis and data computing.

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

iii) It is capable of graphical display.

iv) It is a free software.

Q1 Solve the followings.

(i) $4+6+8\div2-5$

> 4+6+8/2-5

[1] 9

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

> 2^2 + abs(-3)+ sqrt (45)

[1] 13.7082

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

> 5^3 + 7 * 5 * 8 + 46/5

[1] 414.2

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

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

[1] 5.671567

(v) round off

$46\div7+9\times8$

> round (46/7+9*8)

[1] 79

Q4

```
> x <- matrix (nrow = 4, ncol = 2, data = c(1,2,3,
  4,5,6,7,8,9))
> x        [,1]  [,2]

[1]  [1,]   1    5
     [2,]   2    6
     [3,]   3    7
     [4,]   4    8
     [5,]   5    9
```

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

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

```
> x <- matrix (nrow = 3, ncol3, data = c(4,7,9,-2,0,-5,
  6,7,3))
> x      [,1]  [,2]  [,3]
  [1,]    4    -2     6
  [2,]    7     0     7
  [3,]    9    -5     3

> y <- matrix (nrow = 3, ncol 3, data = c(10,12,15,
  -5, -4, -6, 7, 9, 5))
> y      [,1]  [,2]  [,3]
  [1,]   10    -5     7
  [2,]   12    -4     9
  [3,]   15     6     5
```

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 $\Sigma P(x) = 1$

So,

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

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

$\therefore$ The given data is not a p.m.f 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) $\Sigma P(x) = 1$

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

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

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

code :-

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

Scanned with CamScanner

Q2(i) find the c.d.f for the following p.m.f. and sketch the graph.

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

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

$0.2 \quad 10 \le x < 20$

$0.4 \quad 20 \le x < 30$

$0.75 \quad 30 \le x < 40$

$0.96 \quad 40 \le x < 50$

$1.0 \quad 50 \le x$

Q2 (ii)

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

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

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

$= 0.40 \quad 2 \le x < 3$

$= 0.50 \quad 3 \le x < 4$

$= 0.70 \quad 4 \le x < 5$

$= 0.90 \quad 5 \le x < 6$

$= 1.00 \quad x \ge 6$

```
> 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)
```

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 i) 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 no 3

Topic :- Binomial distribution:

$P(X=x)$ = dbinom (x, n, p)

$P(x \le x$ - pbinom (x, n, p)

$P(X > x)$ = 1-pbinom (x, n, p)

If x is unknown

$p_1 = P(X < x) =$ qbinom (p, n, p)

1) find the probility of exactly 10 success in hundred trials with $p = 0.1$

2) suppose there are 12 mcq each question has 5 option out of having which 1 is correct. find the probility of having exactly 4 correct answer.
ii) Atmost 4 correct answers.
iii) more than 5 correct answer.

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

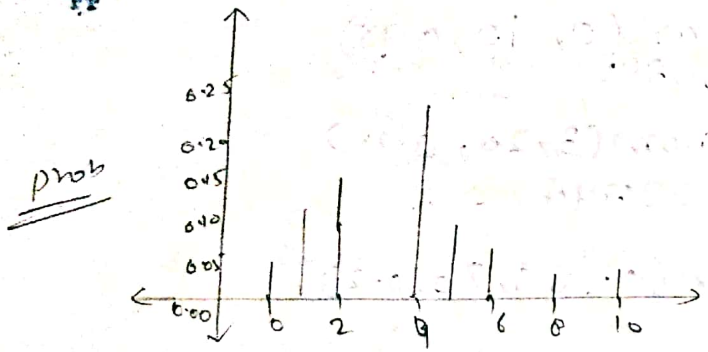
4) $n=12$ $p=0.25$ find following probility
i) $P(X=5)$ (iii) $P(X>7)$
ii) $P(X \le 5)$ (iv) $P(5 < x < 7)$

5) the probility of a salesman making a side to customer 0.15 find the probility of
i) No sales out of 10 customers.
ii) More than 3 sales out of 20 customers.

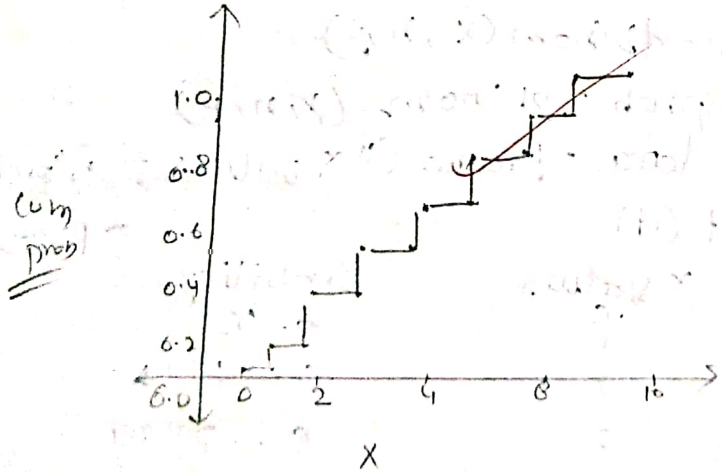
6) A sales man has 20% probility of making a sale to customer. Out of 30 customer what minimum number of sales he can make with 88% of procbility.

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

> plot (x, prob, 'h')



> plot (x, cumprob, "s")



```
p4 = rnorm (5,12,3)
>p4
[1] 15.254728  16.540505  11.200535  6.419944  12.272...
```

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

```
> a1 = pnorm (7, 10, 2)
> a1
[1] 0.0668072
> a2 = pnorm(5,10,2) - pnorm (12,10,2)
> a2
[1] -0.8351351
> a3 = 1- pnorm (12, 10, 2)
> a3
[1] 0.1586558
> a4 = rnorm (10, 10, 2)
> a4
[1] 11.600431  9.920417  12.637741  8.073354
    8.721380  9.193726  9.366024  11.707106
    9.537504  10.715006
> a5 = qnorm (0.4, 10, 2)
> a5
[1] 9.493306
```

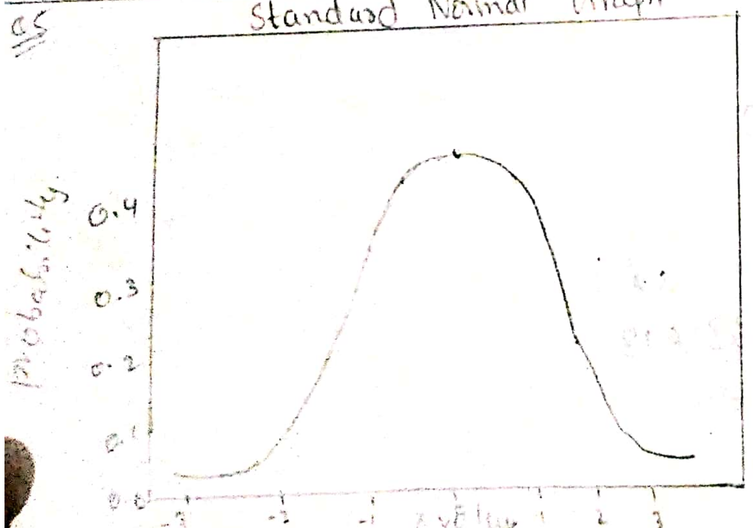
Q4 $X \sim N(30, 100)$. $\sigma = 10$

i) $P(X < 40)$
ii) $P(X > 35)$
iii) $P(25 < X < 35)$
iv) Find x such that $P(X < K) = 0.6$.

```
> f1 = pnorm(40,30,10)
> f1
[1] 0.0413447
> f2 = 1-pnorm (35, 30,10)
> f2
[1] 0.3085375
> f3 = pnorm (25,30. 10)- pnorm(35,30,10)
> f3
[1] -0.3029249
> f4 = qnorm (0.6,30,10)
> f4
[1] 32.53347
```

Q5

Standard Normal Graph



Q5 Plot the standard normal graph.

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

Practical no-5

(1) Let the hypothesis $H_0: \mu = 15$, $H_1: \mu \neq 15$
Random sample of size 400 drawn and it is collected the sample mean is 14 and the standard deviation is 3. Test the hypothesis at 5% level of significance.

```
> mo = 15
> mx = 14; sd = 3
> n = 400
> scal = (mx - mo) / (sd / sqrt(n)).
> scal.
[1] -6.66667
> pvalue = 2*(1-pnorm (abs(scal))
> pvalue
[1] 2.618796e-11
```

∴ The pvalue is less than 0.5 we reject the value H_0 $\mu = 15$.

(2) Test the Hypothesis $H_0: \mu = 10$ $H_1 = \mu \neq 10$.
Random sample mean 10.2 and Standard deviation 2.25 Test the hypothesis at 5% level of significance.

∴ P value is less than 0.05 we reject the value.

(9) Last year the formers lost 20% of their crop and 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 signification.

```
>p=0.2
>p=9/60
>n=60
> Q <-P
> zcal= (p-P)/ sqrt (P* Q/n))
> zcal
[1] -0.9682450
> pvalue = 2* (1-pnorm (zcal))).
> pvalue
[1] 0.3329216
```

∴ The pvalue is greater than 0.01 we accept the value of H0

Test the hypothesis H0 : u = 12.5 from the following sample at 5% level of significance.


```
>ma = mean(l)
>variance = (n-1) * var(n)/n
> sd = sqrt (variance)
> mo = 12.5
> t = (ma - mo) / (sd / sqrt (n))
> pvalue = 2* (1-pnorm (abs(t)))
> pvalue
[0]
[1] 8.04909
```

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

Practical no:- 6

Topic :- large sample test.

Q. Let the population mean (the amount spent per customer in a restorant) is 250. A sample of 100 customers selected. The standard deviation 30. test the hypothesis that the population is 250 or not at 5% level of significance.

Q. In a random sample of 1000 students it is found that 150 use blue pen. test the hypothesis that the population preportion 0.5 at 1% level of signification.

Ans. Ho = u = 275 against H1 = u = 275

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

Scanned with CamScanner

Q5 In a sample of 6000 students in a college 400 use blue ink In another College from a sample of 900 student 480 use blue ink. Test the hypothesis that the proportion of student using blue ink in 2 College are equal or not at 1% level of significance.

2)

Ho: u ≠ 0.8 against H₁: u ≠ 0.8:

```
>p=0.8; P=750/1000 ; n=1000; q=1-p
>zcal=(P-p)/sqrt(p*q/n)
>zcal
[1] -3.952847
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1] 7.72268e-05
```

Since the value is less than 1% level of significance we reject the hypothesis.

5 >n1 = 600
>n2 = 900
>p1 = 400/600
>p2 = 450/900
>p = ((n * p1) + (n2 * p2) / (n1 + n2))
> q = 1-p
>zcal = (p1-p2) /sqrt (p * q)/ (1/n1) + (1/n2))
>zcal
[1] 0.01772648
> pvalue = 2 * (1 - pnorm (abs(zcal)))
>pvalue
[1] 1.753222e-10
∴ The pvalue is less than 1% level of significance we reject the hypothesis.

⑥ Ho: $\mu_2 = p_1 = p_2$ against $\mu_1 \neq p_1 = p_2$
>n1 = 200 ; n2 = 200 ; p1 = 44/200 ; p2 = 30/200
> p1 = (n * p1 + n2 * p2) / (n + n2).
> p
[1] 0.185
> q = 1-p
[1] = 0.815
>zcal = (p1-p2)/ sqrt (p * q * (1/n1) + (1/n2))
>zcal
[1] 1.802441

> pvalue = 2 * (p1 - pnorm (abs(zcal)))
> pvalue
[1] 0.07
∴ p value is greater than 0.007 we accept the hypothesis.

practical no. 7

Topic : Small sample test.

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

Ho : $\mu = 66$

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

one sample t - test
data : x
t = 60319, df = 9, pvalue = 1.558e-13
alternative hypothesis
True mean is not equal to 0
95 percent confidence interval
85.65171 70.~~14~~829
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 group of student the following marks test the hypothesis that there is no significance diff^n 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

H0 : there is not diff b/w the 2 groups.

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

data : x and y
t = 2.2573  df = 16.376  p value = 0.0398
alternative hypothesis.
        true diff^n is mean is
not equal. to 0.95 percent
confidence interval:
0 : 1620205        5.0371795
sample estimates.
mean of x  mean of y
 20.1          17.5
> p-value = 0.3798
> if (pvalue > 0.05) {cat ("accept ho")}
else { cat ("reject ho")}
 reject H0.
```

(4) following are the weight 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

soln : H0: there is no significance diff?

> x = c(before)

> y = c(After)

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

paired t test

data: x and y

t = 4.3450, df = 5, pvalue = 0.9963

alternative hypothesis: true difference is mean less than 0.

95 percent ~~confidence~~ interval.

- inf 20.0295

sample estimates:

mean of the diffn.

19.83333

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

22

practical no. 8

Topic large and small test.

1. Ho: $\mu = 55$, H_1: $\mu \neq 55$

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

As pvalue is less than 0.05 we reject H at 5% level of significance.

2. Ho: $P=0.5$ against Ho: $P \neq 0.5$

```
> p=0.5
>q=1-p
> n=100
>zcal=(p-P)/(sqrt(P*q/n))
>zcal
[1] 0
> P value =2*(1-pnorm(abs(zcal)))
>P value
1]
```

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

(4) H0: $\mu = 100$ against H1: $\mu \neq 100$

> var = 64
> n = 400
> mo = 100
> mx = 900
> sd = sqrt (var)
> sd
[1] 8
> zcal = (mx - 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 H0 at 5% level of significance.

Q5 H0: $\mu = 66$ against. H1: $\mu \neq 66$

> x = c(63, 63, 60, 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
60.14286

since p value is less then 0.05 we reject H0 at 1% level of significance.

```
> pvalue
[1] 6.334248 e-05
```

∴ pvalue is less than 0.05 we reject Ho.

(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)
> p
[1] 0.2
> q = 1-p
[1] 0.8
> zcal = (p1-p2)/sqrt(p*q*(1/n1+1/n2))
> zcal
[1] 0.9128709
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.3613104
```

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

Q1 use the following data test whether the condition of home & the condition of child are indipendent or not.

→ condition of home

clean	Clean 70	Dirty 50
fairly clean	80	20
Dirty	35	45

→ Ho: condition of home & child are independent

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

m = 3

n = 2

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

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

pv = chisq. test (y)

pv

person'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

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
named (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, dinom df = 9,

pvalue = 0.00285

```
> anova = aov(values ~ ind, data = d)
summary (anova).
```

	df	sumsq	meansq	Fvalue	Pr(>F)
nd	3	91.44	30.479	6.845	0.00230
esiduals	20	89.00	4.453		

s Hao to import a csv file in
→ x=read csv(" C:|users|admin|Desktop/stats.csv")
print (x)

	stats	maths
1	46	68
2	45	40
3	42	42
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)
sd = sqrt ((n-1)* var $ stats) /n)
sd
[1] 12.6491
am = mean (x $ maths)
[1] 39.4