PRACTICAL - 01

Aim: Basics of R software

i) R software is used for statastical analysis and data computing

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

iii) It is capable of graphical display

iv) It is a free software

Q.1 Solve the following

1) $4+6+8\div2-5$

>4+6+8/2-5

[1] 9

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

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

[1] 13.7082

3) $5^3+7\times5\times8+46\div5$

5^3 + 7x5x8 +46 /5

[1] 414.2

4) $\sqrt{4^2+5\times3+7\div6}$

>sqrt (4^2+5 x3+7/6)

[1] 5.6716

5) round off [46 ÷ 7 + 9×8]

```
> round (46/7+9×8)
[1] 79
```

Q.2

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

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

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

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

> 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
```

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 $X = \begin{bmatrix} 4 & -2 & 6 \\ 7 & 0 & 7 \\ 9 & -6 & 3 \end{bmatrix}$ and $Y = \begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 8 \end{bmatrix}$


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

> Y <- matrix (nrow=3, ncol=3, data = c(10,12,15,-5,-4,-6,7,9,8))
> Y
     [,1]  [,2]  [,3]
[1,]  10    -5     7
[2,]  12    -4     9
[3,]  15    -6     8

> 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
```

Practical - 2

Aim : 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 given data is pmf then $\Sigma P(x) = 1$

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

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

$= 1.0$

As $P(2) = -0.5$; It is not a p.m.f

2)

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

Condition for pmf is $\Sigma p(x) = 1$

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

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

$= 1.1$

As $\sum P(x) \neq 1$; It is not a pmf

3)

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

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

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

$= 1$

$\therefore$ The given data is pmf

Condition for pmf

$\sum P(x) = 1$

$\forall P(x) \geq 0$

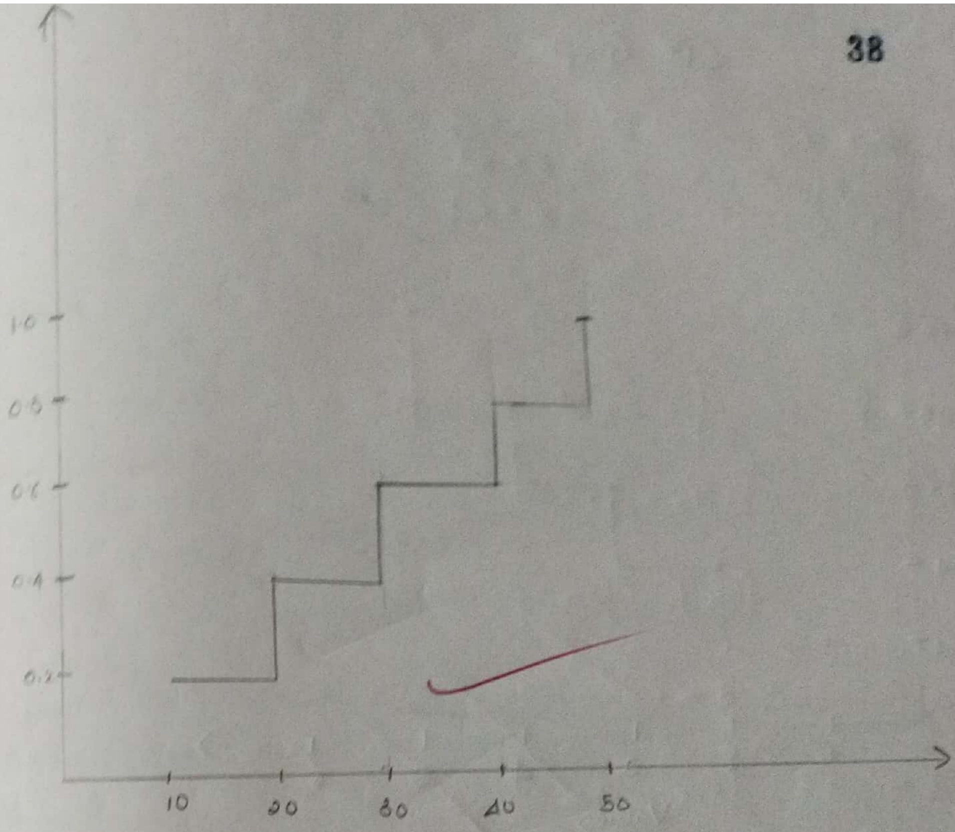
Code :

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

Q2 Find the c.d.f for the following p.m.f and sketch the graph.

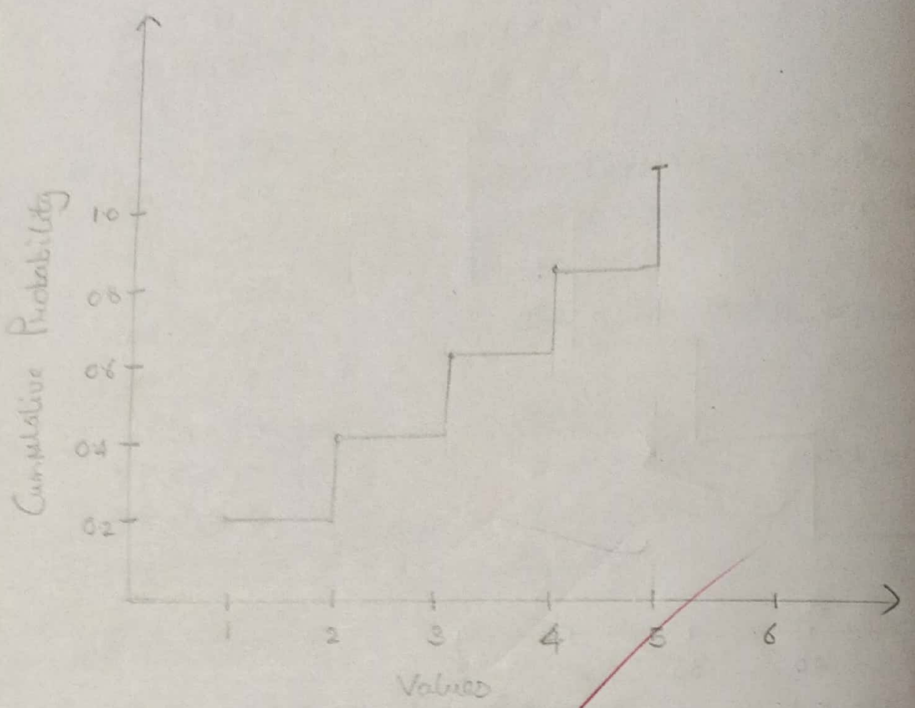
x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1

$F(x) = 0$	$x < 10$
$= 0.2$	$10 \leq x < 20$
$= 0.4$	$20 \leq x < 30$
$= \cancel{0.75}\ 0.75$	$30 \leq x < 40$
$= \cancel{0.95}\ 0.9$	$40 \leq x < 50$
$= \cancel{0.1}\ 0.1$	$x \geq 50$




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

CDF Graph



Find

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

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

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

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

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

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

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

$= 1 \quad x \geq 6$

```
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,  0.5, 0.7, 0.9, 1
> x = c(1,2,3,4,5,6)
> plot (x, cumsum (prob), "s", xlab = "value ,ylab= "cumulative pro
      main= " CDF Graph", col=" brown")
```

Q3 Check that whether the following is pdf or not

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

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

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

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

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

$$= \left[3x\right]_0^1 - \left[\frac{2x^2}{2}\right]_0^1$$

$$= 3 - 1$$

$$= 2$$

To be p.d.f ; $\int_0^1 f(x) = 1$

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

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

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

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

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

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

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

$$= 1$$

$\therefore \int_0^1 f(x)\,dx = 1$; It is a p.d.f

PRACTICAL - 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
$P1 = P(X \le x) =$ qbinom(p_1, 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 option out of which 1 is correct. Find probability of having exactly 4 correct answers
(ii) Atmost 4 correct answers.
(iii) More than 5 correct answers.

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

4) n=12, p=0.25 find the following probabilities

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

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

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] 0.4274445
iii) 1 - pbinom (5, 12, 0.2)
[1] 0.0194058

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
ii) pbinom (5, 12, 0.25)
[1] 0.9455978
iii) 1 - pbinom (7, 12, 0.25
[1] 0.0027851
iv) dbinom (6, 12, 0.25)
[1] 0.04014945

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

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

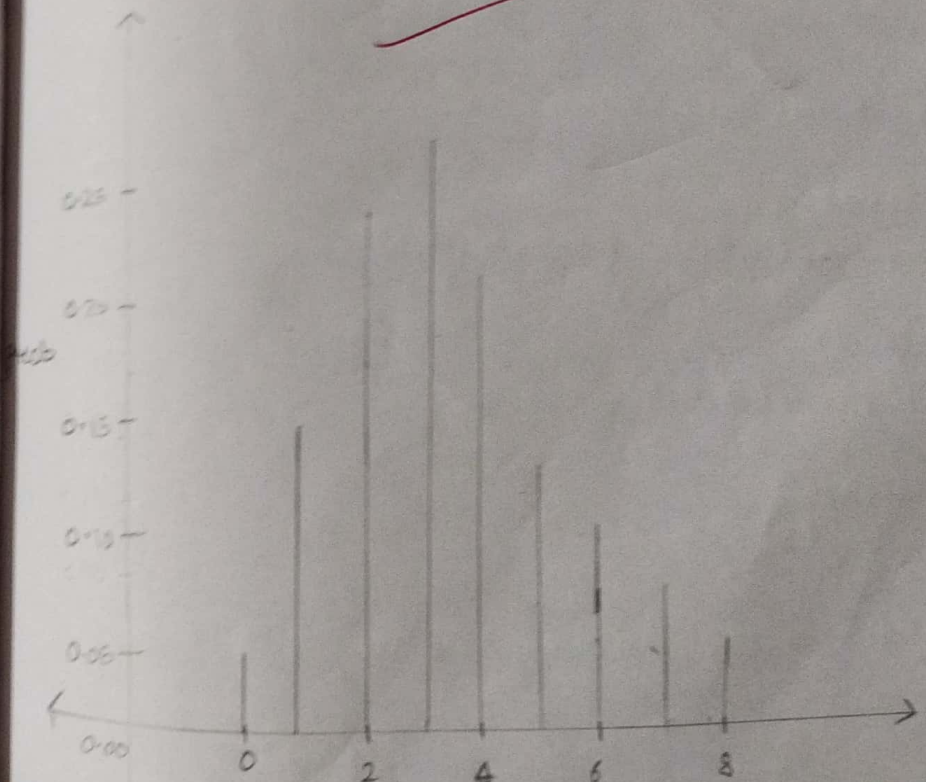
6) A salesman has 20% probability of making a sale to customer out of 30 customers. What is minimum no of sales he can make with 88% of probability.

7) X follows binomial distribution with $n=10$, $p=0.3$ Plot the graph of p.m.f & c.d.f


```
>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)
```

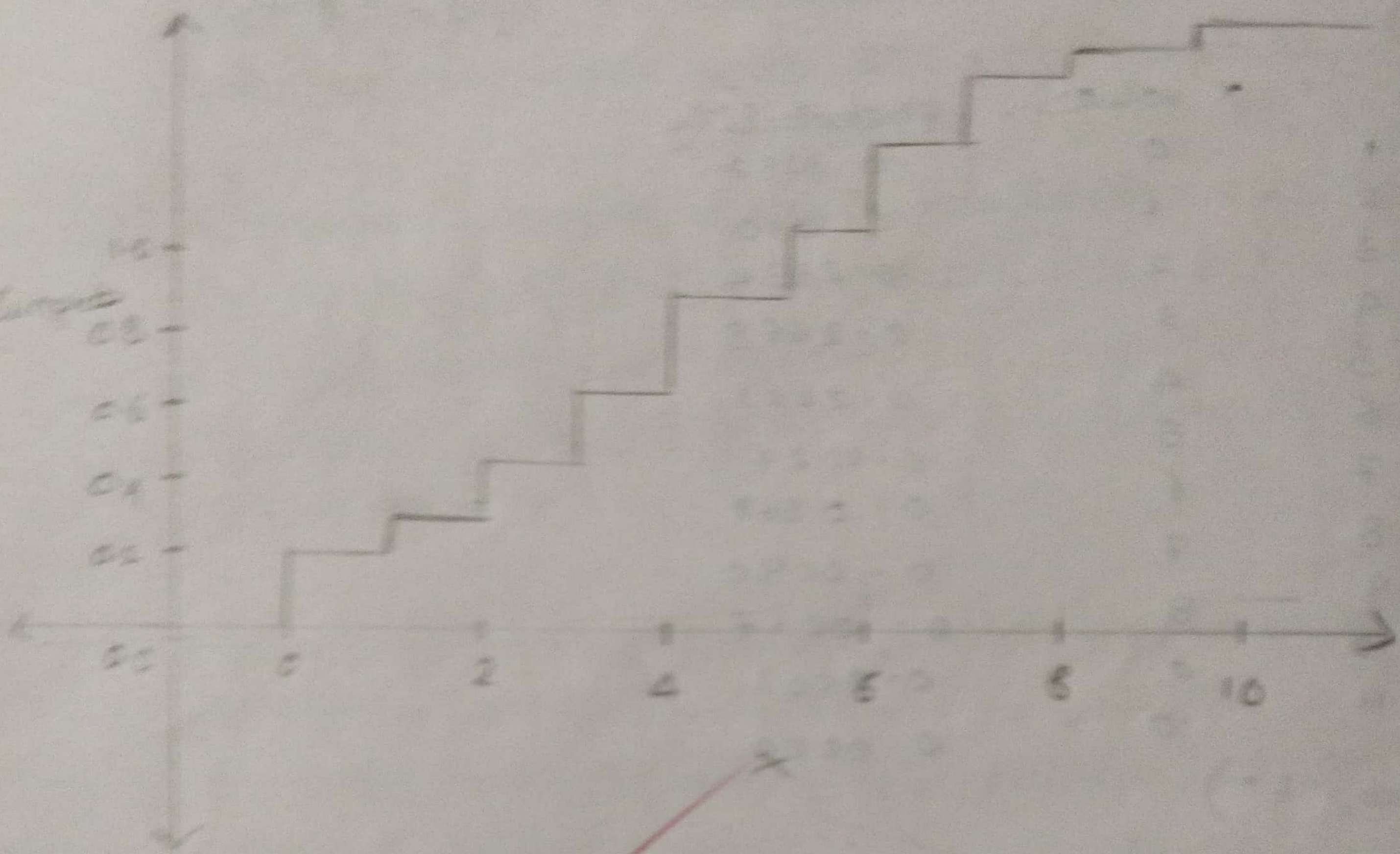
	X values	probability
1	0	0.0282
2	1	0.1210
3	2	0.2334
4	3	0.2668
5	4	0.2001
6	5	0.1029
7	6	0.0367
8	7	0.0090
9	8	0.0017
10	9	0.0001
11	10	0.0000

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



5#

```
> plot (x, cumprob, "s")
```



PRACTICAL - 4

TOPIC :- 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)

Q.1 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(10,12,3)
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  6.419944  12.272460
```

2) X follows normal distribution with $\mu = 10$, $\sigma = 2$. Find (i) $P(X \le 7)$ (ii) $P(6 < x < 12)$ (iii) $P(x > 12)$ (iv) 10 observation. Also, find K such that $P(X < k) = 0.4$

CODE:

```
> p1 = pnorm(.4,10,2)
p1
[1] 0.668072
> p2 = pnorm(6,10,2) - pnorm(12,10,2)
p2
[1] 0.8351351
> p3 = 1 - pnorm(12,10,2)
p3
[1] 0.1586553
> p4 = rnorm(10,10,2)
p4
[1] 17.608931  9.920417  12.637741  8.073354
    8.721380   9.193725  6.366824   11.707106
    9.537584   12.715006

p5 = qnorm(0.4,10,2)

p5
[1] 9.493306
```

Generate 5 random numbers 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     9.963444     13.345904     17.509668
> am = mean(x)
  am
[1] 11.87345
> cat ("Sample mean is = ", am)
  Sample mean  = 11.87345
> me = median (x)
  me
[1] 10.76499
> cat (" Sample median = ", me)
  Sample median = 10.76499
> n=5
> v= (n-1)* var(x)/n
> v
[1] 11.09965
> S.D = sqrt(v)
  S.D
[1] 3.33163
> cat ("S.D. = ", SD)
  S.D. = 3.33163
```

Q.4 $X \sim N(30, 100)$, $\sigma = 10$, (ii) $P(x > 35)$ (iii) $P(25 < x < 35)$

i) $P(x \leq 40)$

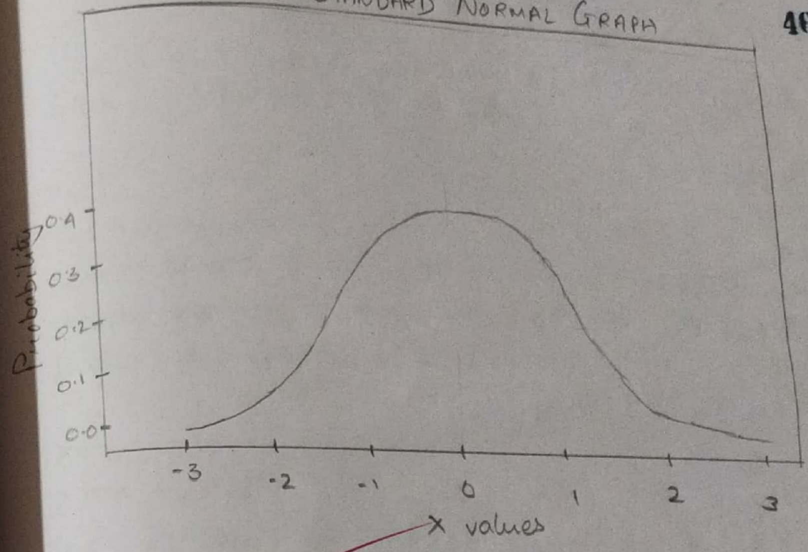
iv) Find K such that $P(x < k) = 0.6$

```
> 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
```

Q.5 Plot Standard normal graph

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


STANDARD NORMAL GRAPH



Topic : Normal and t test

Ho : μ=15   H1 : μ ≠ 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 signifance. (0.05> accept value , 0.05< reject value)

```
> mo = 15
> mx = 14
> n = 400
> sd = 3
> 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 Ho = μ = 15

2. Test the hypothesis $H_0 = \mu = 10$ 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

```
>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 is greater than 0.05; Value is accepted

Test hypothesis Ho: proportional of smokers in college is 0.2. A sample is collected and calculated sample proportional as 0.125. Test hypothesis at 5% level of significance

```
>p=0.2
>P0=0.125
>n= 400
>Q=1-p
>Zcal= (p-P0)/(sqrt(p*Q/n)
>Zcal
[1] -3.75
>pvalue = 2* (1-pnorm(abs(zcal)))
>pvalue
[1] 0.00017 68346
```

∴ Value < 0.05 ; Value is Rejected

Last year farmer's last 20% of their crops. A random sample of 60 fields are collected and it found that a field crops are polluted. Test hypothesis at 1% level of significance.

```
>p=0.2
>P0 = 960
>n=60
>zcal (P0-p)/(sqrt(p* q/n))
>zcal
[1] -0.968245
>pvalue = 2* (1-pnorm(abs(zcal)))
>pvalue
[1] 0.332916
```

∴ P Value > 0.1 ; Value is accepted.

Test the hypothesis Ho: μ=12.5 from 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] 0.019521
>t= (mx-mo) / (sd /sqrt(n))
>t
[1] -8.894909
>Pvalue = 2*(1-pnorm(abs(t)))
>Pvalue
[1] 0
```

Value < 0.05 ; Value is accepted

48

PRACTICAL - 6

Aim : Large Sample Test

Let the population mean (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. is 30. Test the hypothesis that the population mean is 250 or not on 5% level of significance

In a random sample of 1000 students it is found that 750 use blue pen. Test hypothesis that 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)))
>zcal
[1]8.33333
>pvalue = 2 * (1- pnorm (abs ( zcal)))
>pvalue
[1] 0
```

The value is less than 0.05 ; Value is rejected.

2. SOLUTION:

```
>p=0.8
>q=1-p
>P0 = 750/1000
>n=1000
>zcal (p-P0 )/(sqrt(p* Q/n))
>zcal
[1] -3.952847
>pvalue=2*(1-pnorm(abs(zcal)))
[1] 7.72268e-05
```

∴ Value is less than 0.01; Value rejected.

3. Two random sample of size 1000 & 2000 are drawn from two population with same S.D 2.5, sample means are 67.5 & 68. Test the hypothesis $\mu_1 = \mu_2$ at 5% < of significance

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

Hospital A	Hospital B
84	34
61.2	59.4
7.9	7.5

5. In a sample of 600 students, 400 use blue ink. In another college from a sample of 900 students, 450 use blue ink. Test the hypothesis that proportion of Students using blue ink in both college is equal or not at 1% level of significance

3. SOLUTION

```
>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] 2.41756e-07
```

As value is less than 0.05; Value is Rejected.

4. SOLUTION

```
>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] 0.245021
```

⟶ Value Accepted.

5. SOLUTION:

```
>n1 = 600
>n2 = 900
>P1 = 400/600
>P2 = 450/900
>P = (n1*P1+n2 * P2)/(n1+n2)
>pq=1-p
>zcal = (P1-P2) / sqrt (p*q* ((1/n1)+ (1/n2)))
>zcal
[1] 6.381534
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 1.75...
```

⟶ Value Rejected

PRACTICAL - 7

TOPIC : Small Sample test

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 population with
average 66.

$H_0 : \mu = 66$

> 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 = $1.558\,e^{-13}$

alternative hypothesis

True mean is not equal to 0

95% confidence interval

66.65171 70.14829

sample estimate

Mean of x

67.9

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

2. Two groups of students scored following marks. Test hypo that there is no significant difference between 2 groups.

GR1 - 18, 22, 27, 21, 17, 20, 17, 23, 20, 22, 21

GR2 - 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

H_0 : There is no difference between the two groups

```
> x = c(18, 22, 27, 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 = 2.2573 df = 16.376 p-value = 0.03798

alternative hypothesis :

True difference in means in not equal to 0. ~~95~~%

95% confidence interval :

0.1628205 5.0371795

Sample estimates :

Mean of x Mean of y

20.1 17.5

```
> p-value = 0.03798
> if (p-value > 0.05) {cat ("accept Ho")}
else { cat ("reject Ho")}
```

reject Ho

3. The sales data of 6 shops before & after a special campaign are:

Before : 53, 28, 31, 48, 50, 42

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

Test hypothesis that campaign is effective or not

H_0 : There is no significant difference in sales before & after 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 = ~~28~~ -2.7815 , df = 5 , pvalue = 0.9806

alternative hypothesis.

True difference in mean is greater than 0

95 percent confidence level :

-6.035547 inf

Sample estimates :

Mean of difference

-3.5

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

4) Following are weights before & after ... is diet eff...

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

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

H0 : There is significant 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 = 4.3458 , df = 5 , p-value = 0.9963

alternative hypothesis : True difference in mean is less than 0

95 percent confidence interval :

-inf 29.0295

Sample estimate :

mean of differences

19.83333

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

5. Two medicines are applied to two group of patient respectively

GR1 : 10, 12, 13, 11, 14

GR2 : 8, 9, 12, 14, 15, 10, 9

H0 : There is no significant difference

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

T-test

data : x and y

t = 0.80384 , df = 9.7594 , pvalue = 0.4406

alternative hypothesis : True difference in means is not equal to

95 % level confidence interval

-0.9698553 4.2981886

Sample estimates :

mean of x	mean of y
12.00000	10.33333

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

PRACTICAL - 8

Q.1 The arithmetic mean of a sample of 100 items from a lar[ge] population is 52. If the standard is 7, test the hypothesis [that] 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. Does the information supports that exactly half of males in the city are smokers? Test at 1% LOS

Q.3 Thousand articles 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 two factory are similar or not.

Q.4 A sample of size 400 was drawn at sample mean is 99. Test at 5% LOS that sample comes from population mean 100 and variance 64

Q.5 The flower stems are selected and heights are found to be (cm) 63, 63, 68, 69, 71, 71, 72. Test the hypothesis that mean height is 66 or not at 1% LOS

Q.6 Two random samples were drawn from 2 normal population and their 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 same variance at 5% LOS

Q.7 A company producing light bulbs finds that mean life span of population of bulbs is 1200 hours with 54 s.d. 125. A sample of 100 bulbs have mean 1150 hours. Test whether difference between population and sample mean is significantly different?

Q.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 rotten apples in two consignment are significantly different at 1% LOS?

TOPIC :- Large & Small Test

$H_0 = \mu = 55$, $H_1 : \mu \neq 55$

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

As pvalue is less than 0.05; we reject Ho

$H_0 : P = 0.5$ against $H_1 : P \neq 0.5$

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

As pvalue is greater than 0.05; we accept Ho

(3) $H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

```
>n1=1000
>n2=1500
>p1= 2/1000
>p2=1/1500
>p= (n1*p1 + n2*p2) /(n1+n2)
>p
[1] 0.0012
>q = 1-p
>zcal = (p1-p2) / (sqrt (p*q * (1/n1 + 1/n2)))
>zcal
[1] 0.9433752
>pvalue = 2* (1- pnorm( abs(zcal)))
>pvalue
[1] 0.345489
```

∴ pvalue is greater than 0.05; we accept Ho at 5% level of signifi

(4) $H_0: \mu = 100$ against $H_1: \mu \neq 100$

```
>var = 64
>n = 400
>m0 = 100
>mx = 99
>sd = sqrt (var)
>zcal = (mx-m0) / (sd /(sqrt(n)))
>zcal
[1] 2.5
>pvalue= 2* (1- pnorm(abs(zcal)))
>pvalue
[1] 0.01242
```

⇒ pvalue is less than 0.05 ; we reject Ho

(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 , 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

Since p-value is less than 0.05 ; we reject Ho at 1% LOS

(6) $H_0: \sigma_1 = \sigma_2$ against $H_1: \sigma_1 \neq \sigma_2$

```
>x = c(66,67,75,76,82,88,90,92)
>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 =0.788803 ,num df =7 , dnom.df =10 , pvalue=0.7737
alternative hypothesis : True ratio of variance is not equal to 1
95 percent confidence interval :
0.199509 3.751881
Sample estimates :
ratio of variances
0.7880255

Pvalue is greater than 0.05 ; we accept Ho at 5% LOS

22

(7) $H_0: \mu = 1150$ against $H_1: \mu \neq 1150$

```
> n = 100
> mx = 1150
> mo = 1200
> sd = 125
> zcal = (mx-mo) /(sd/(sqrt (n)))
> zcal
[1] -4
> pvalue = 2 * (1- pnorm (abs (zcal)))
> pvalue
[1] 6.334243 e-05
```

∴ pvalue is less than 0.05 ; we reject Ho

(8) $H_0 : P_1 = P_2$ 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
> 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% LOS

PRACTICAL - 09

TOPIC : Non-parametric Testing of Hypothesis using R Environment

① The following data represents earnings (in dollars) for a random sample of five common stocks listed on New York stock exchange. Test whether median earnings is 4 dollars.
Data : 1.68, 3.35, 2.50, 6.23, 3.24

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

data = s and n
number of success = 1 , number of trials = 5, p-value = 0.9688
alt hypothesis : True probability of success is greater than 0.5

95 percent confidence interval :
0.01020622   1.00000000

② The scores of 8 students in reading before & after lesson are as follows

Student No :-	1	2	3	4	5	6	7	8
Score Before :-	10	15	16	12	09	07	11	12
Score After :-	13	16	15	13	09	10	13	10

Test whether there is effect of reading

```
>b = c (10,15, 16,12,09,07,11,12)
>a = c (13,16,15,13,09,10,13,10)
>D = b - a
>wilcox - test (D , alternative = "greater")
          Wilcoxon signed rank test with continuity
```

Correlation data : D
V = 10.5 , p-value = 0.8722
alternative hypothesis : True location is greater than 0

∴ p-value is greater than 0.05, we accept it.

③ The diameter of ball bearing was measured by 6 inspectors, each using two different kinds of calipers. The results were stored below. Test whether average ball bearing

Inspector	1	2	3	4	5	6
Caliper 1	0.265	0.268	0.266	0.267	0.269	0.264
Caliper 2	0.263	0.262	0.270	0.261	0.271	0.260

CODE:

```
>x = c(0.265, 0.268, 0.266, 0.267, 0.269, 0.264)
>y = c(0.263, 0.262, 0.270, 0.261, 0.271, 0.260)
>wilcox.test(x, y, alternative = "greater")
```

Wilcoxon rank sum test

data: x and y

w = 24 , p = 0.17

alternative hypothesis: True mean shift is greater than 0

(4) An officer has 3 elective type writers A, B and C. In a study of machine usage, firm has kept records of machine usage rate of seven weeks

A	B	C
12.3	15.7	32.4
15.4	10.8	41.2
10.3	45.0	35.1
08.0	12.3	25.0
14.6	08.2	08.2
–	20.1	18.4
–	26.3	32.5

CODE :-

```
>x = c(12.3, 15.4, 10.3, 8.0, 14.6)
>n1 = length(x)
>n1
[1] 5
>y = c(15.7, 10.8, 45.0, 12.3, 8.3, 20.1, 26.3)
>n2 = length(y)
>n2
[1] 7
>z = c(32.4, 41.2, 35.1, 25.0, 8.2, 18.4, 32.5)
>n3 = length(z)
>n3
[1] 7

>X = c(x, y, z)
>g = c(rep(1,n1), rep(2,n2), rep(3,n3))
>Kruskal.test(X, g)
```

Kruskal-Wallis rank sum test

data : X and g

Kruskal Wallis Chi-squared = 5.217, df = 2, p-value = 0.07365