Practical no. 1

- Basics of R software

1. R is a software for data analysis and computing
2. This software is used for effective data handling and output storage is possible
3. Capable of graphical display
4. It is a free software

Q1. 1. $2^2 + \sqrt{25} + 35$

```
> (2**2) + sqrt (25) + 35
[1] 44
```

2. $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$

```
> (2*5*3) + (62/5) + sqrt (49)
[1] 49.4
```

3. $\sqrt{76 + 4 \times 2 + 9 \div 5}$

```
> sqrt (76 + (4*2) + (9/5))
[1] 9.262829
```

4. $42 + |-10| + 7^2 + 3 \times 9$

```
> 42 + abs (-10) + (7**2) + (3*9)
[1] 128
```

Q2. $x = 20$

$y = 30$

find $x+y$, x^2+y^2, $\sqrt{y^3-x^3}$, $|x-y|$

```
> x + y
[1] = 50
> (x** 2) + (y** 2)
[1] = 1300
> sqrt ((y** 3) - (x** 3))
[1] = 137.8405
> abs (x - y)
[1] 10
```

Q3. `c(2,3,4,5)^2`

```
[1] 4 9 16 25
```

`c(4,5,6,8)* 3`

```
[1] 12 15 18 24
```

`c(2,3,5,7)* c(-2,-3,-5,-4)`

```
[1] -4 -9 -25 -28
```

`c(2, 3, 5, 7)* c(8,9)`

```
[1] 16 27 40 63
```

```
c(1,2,3,4,5)* c(2,3,4)
error

c(1,2,3,4,5,6)^c(2,3)
[1] 1 8 9 64 25 216
  1^2 2^3 3^2 4^3 5^2 6^3
```

Q4. Find the sum, product, maximum, minimum of the values

5, 8, 6, 7, 9, 10, 15, 5.

```
> x = c(5,8,6,7,9,10,15,5)
> length
[1] 8
> sum(x)
[1] 65
> prod(x)
[1] 11340000
> max(x)
[1] 15
> min(x)
[1] 5
```

Q5.
```
> 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
```

Q6 $x = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$

$y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ 10 & 6 & 12 \end{bmatrix}$

```
>x<- matrix (nrow=3, ncol=3, data = c(1,2,3,4,6,7,8,9))
>4<- matrix (nrow=3, ncol=3, data = c 2,-2,10,4,8,6,10,-1
```

```
> x+y
     [1] [2] [3]
[1]    3   8  17
[2]    0  13  -3
[3]   13  12  21
```

```
> x*y
     [1] [2] [3]
[1]    2  16  70
[2]   -4  40 -88
[3]   50  36 108
```

```
> x*2+y*3
     [1] [2] [3]
[1]   8  20  44
[2]  -2  34 -17
[3]  36  30  54

> x=c(2,4,6,1,3,5,7,18,16,14,17,19,3,3,2,5,0,15,9,14,18,
    10,12)
> length(x)
[1] 23
> a= table(x)
> a
x
 0  1  2  3  4  5  6  7  9 10 12 14 15 16 17 18 19
 1  1  2  3  1  2  1  1  1  1  1  2  1  1  1  2  1

> transform
[1]  x  freq
     0    1
     1    1
     2    2
     3    3
     4    1
     5    2
     6    1
     7    1
     9    1
    10    1
```

12	1
14	2
15	1
16	1
17	1
18	2
19	1

```
> breaks = seq (0,20,5)
> s= cut (x, breaks, right = FALSE)
> c = table (s)
> transform (c)

[1]
  s     freq
1 [0,5)   8
2 [5,10)  5
3 [10,15) 4
4 [15,20) 6
```

1. Can the following be p.d.f

1) $f(x) = \begin{cases} 2-x & ; 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$

$\rightarrow \int_1^2 (2-x)\,dx$

$= \int_1^2 2\,dx - \int_1^2 x\,dx$

$= 2x \Big|_1^2 - \frac{x^2}{2}\Big|_1^2$

$= (4-2) - (2-0.5)$

$\neq 1$

$\Rightarrow$ Not a p.d.f

2) $f(x) = \begin{cases} 3x^2 & ; 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

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

$= \int_0^1 \frac{\cancel{3}x^3}{\cancel{3}}$

$= x^3 \Big|_0^1$

$= 1$

$\Rightarrow$ It is a p.d.f

3) $f(x) = \begin{cases} \frac{3x}{2}\left(1 - \frac{x}{2}\right); & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$

$\longrightarrow = \int_0^2 \left(\frac{3x}{2} - \frac{3x^2}{4}\right) dx$

$= \frac{3}{2}\int_0^2 x\, dx - \frac{3}{4}\int_0^2 x^2\, dx$

$= \frac{3}{2}\left[\frac{x^2}{2}\right]_0^2 - \frac{3}{4}\left[\frac{x^3}{3}\right]_0^2$

$= \frac{3}{4}\left[x^2\right]_0^2 - \frac{1}{4}\left[x^3\right]_0^2$

$= \frac{3}{\cancel{4}}\left[\cancel{4}\right] - \frac{1}{4}\left[\cancel{8}^{\,2}\right]$

$= 1$

$=$ It is a p.d.f

Can the following be p.m.f

1.

x	1	2	3	4	5
$p(x)$	0.2	0.3	-0.1	1.5	0.1

Since, 1 of the probability is negative, it is not a p.m.f

2.

x	0	1	2	3	4	5
$p(x)$	0.1	0.3	0.2	0.2	0.1	0.1

Since $p(x) \geq 0 \ \forall x$
and $\Sigma p(x) = 1$
$\therefore$ It is a p.m.f

3.

x	10	20	30	40	50
$p(x)$	0.2	0.3	0.2	0.6	0.2

Since $p(x) \geq 0 \ \forall x$
and $\Sigma p(x) \neq 1$
$\therefore$ It is not a p.m.f

Find $P(x \leq 2)$, $P(2 \leq x < 4)$, P (atleast 4), $P(3 < x < 6)$

x	0	1	2	3	4	5	6
$p(x)$	0.1	0.1	0.2	0.2	0.1	0.2	0.1

$\longrightarrow p(x \leq 2)$

$= p(0) + p(1) + p(2)$

$= 0.1 + 0.1 + 0.2$

$= 0.4$

$p(2 \leq x < 4)$

$= p(2) + p(3)$

$= 0.2 + 0.2$

$= 0.4$

p (atleast 4)

$= p(4) + p(5) + p(6)$

$= 0.1 + 0.2 + 0.1$

$= 0.4$

$p(3 < x < 6)$

$= p(4) + p(5)$

$= 0.1 + 0.2$

$= 0.3$

Probability distribution

Find c.d.s of the following p.m.f & draw the graph

x	10	20	30	40	50
p(x)	0.15	0.25	0.3	0.2	0.1

$f(x) = 0$ if $x < 10$

$= 0.15$ $10 \le x < 20$

$= 0.40$ $20 \le x < 30$

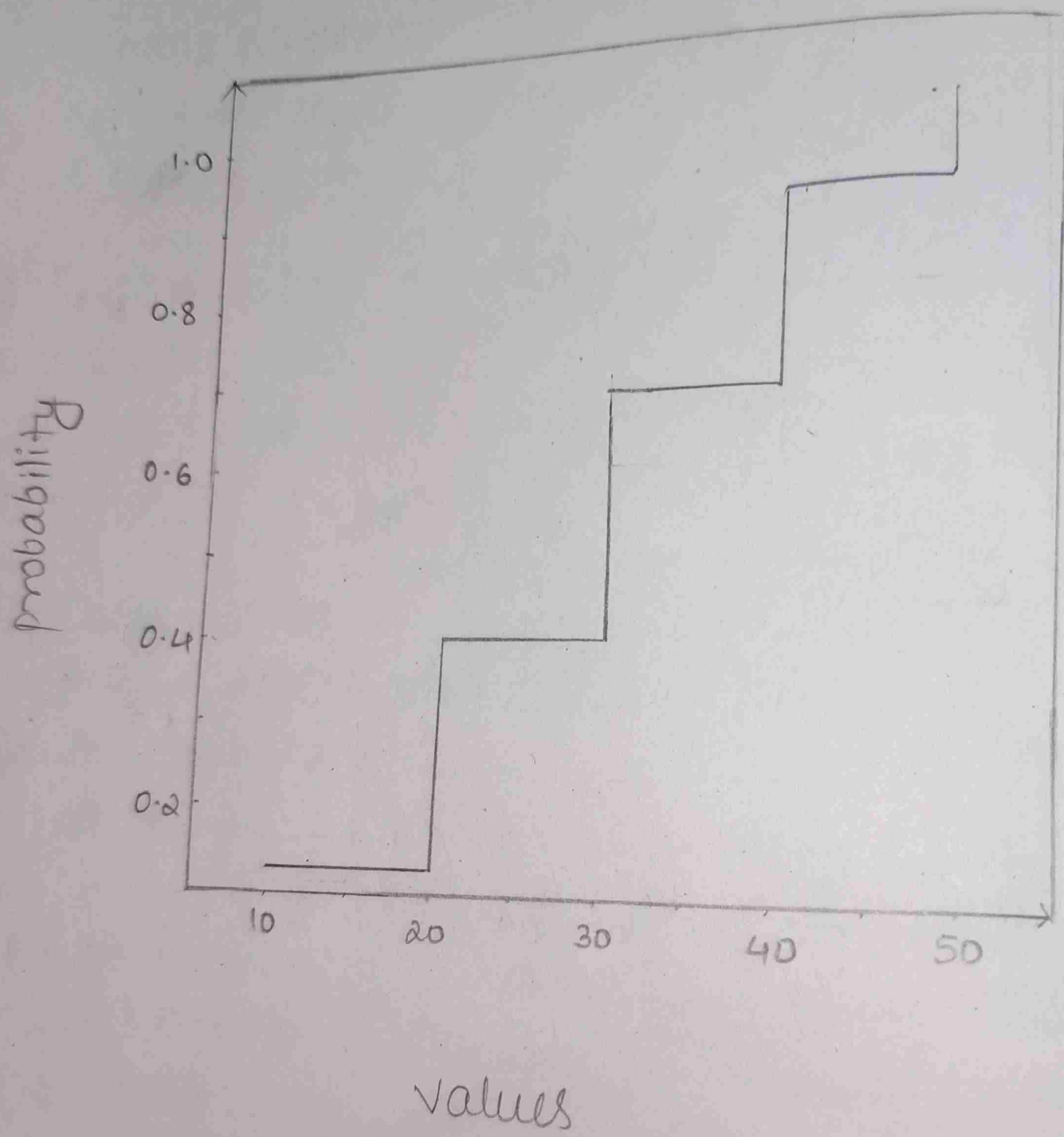
$= 0.70$ $30 \le x < 40$

$= 0.90$ $40 \le x < 50$

$= 1.0$ $x \geq 50$

```
> x = c(10,20,30,40,50)
> prob = c(0.15,0.25,0.3,0.2,0.1)
> cumsum(prob)
[1] 0.15 0.40 0.70 0.90 1.0
> plot(x,cumsum(prob),xlab="values",ylab="probability",main="graph
of c.d.f","s")
```

graph c.d.f



Binomial distribution

Suppose there are 12 mcq's in a test each question has 5 option and only one of them is correct. Find the probability of having (1) 5 correct answer (2) Atmost 4 correct answer

→ Given that :-

$n = 12$, $p = 1/5$, $q = 4/5$

X = Total no. of correct answer

$X \sim B(n, p)$

$n = 12$, $p = 1/5$, $q = 4/5$, $x = 5$

```
>n = 12; p = 1/5; q = 4/5; x = 5
>sum(n, p, q, x)
[1] 18
>dbinom(5, 2, 1/5)
[1] 0.05315022
>pbinom(4, 12, 1/5)
[1] 0.9274445.
```

3. There are 10 member in a comitte, the probability of any member attending a metting i 0.9. Find the probability
(1.) 7 member attended (2.) Atleast 5 members attended
(3.) Atmost 6 member attended

→ Given that
$n = 10$, $p = 0.9$, $q = 0.1$
x = Total no of member attended
$x \sim B(n, p)$
$n = 10$, $p = 0.9$, $q = 0.1$

```
> n=10; p=0.9; q=0.1
> dbinom (7,10,0.9)
[1] 0.05739563
> 1- pbinom (5,10,0.9)
[1] 0.9983651
> pbinom (6,10,0.9)
[1] 0.0127952
```

4. Find the c.d.f and draw the graph ~~x and~~

x	0	1	2	3	4	5	6
p(x)	0.1	0.1	0.2	0.2	0.1	0.2	0.1

```
> x = c (1,2,3,4,5,6)
> prob = c (0.1, 0.1, 0.2, 0.2, 0.1, 0.2, 0.1)
> cumsum (prob)
[1] 0.1 0.2 0.4 0.6 0.7 0.9 1.0
> plot (x, cumsum (prob), xlab = "values", ylab = "probability
main = "graph of c.d.f", "s")
```

Practical no:4

Binomial Distribution

1. Find the complete binomial distribution when $n = 5$ and $p = 0.1$

2. Find the probability of 10 success in 100 times where $p = 0.1$

3. X follows binomial distribution with $n = 12$, $p = 0.25$
find (1) $P(X=5)$ (2) $P(X \leq 5)$ (3) $P(X>7)$ (4) $P(5<X<7)$

4. Probability of a salesman makes a sell to a customer is 0.15. Find the probability
(1.) No sale for 10 customer
(2.) More than 3 sales in 20 customer

5. A student writes 5 mcq's each question has 4 option out of which one is correct. calculate the probability for atleast 3 correct answer

6. X follows binominal distribution with $n =$, $p =$
plot the graph of p.m.f and c.d.f

Note

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

$P(x \leq x)$ = pbinom (x, n, p)

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

To find the value of x for which the probability is given as P_1, command is qbinom

⇒ qbinom (p, n, p)

* Answers:

1. $n = 5$, $p = 0.1$

dbinom (0:5, 5, 0.1)

[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001

2. $x = 10$ $n = 100$ $p = 0.1$

dbinom (10, 100, 0.1)

[1] 0.1318653

3. $n = 12$ $p = 0.25$

(1) $P(x = 5)$ = dbinom (x, n, p)

> dbinom (5, 12, 0.25)

[1] 0.1032414

(2) $P = (x \leq 5)$ = pbinom (x, n, p)

> pbinom (5, 12, 0.25)

[1] 0.9455978

(4) $P(5<X<7)$ = dbiom (x, n, p)

> dbinom (6, 12, 0.25)

[1] 0.04014945

(3) $P(X>7)$ = 1-pbinom (x, n, p)

1-pbinom (7, 12, 0.25)

[1] 0.00278151

4. $p = 0.15$

1. $n = 10$, $p = 0.15$, $x = 0$

> dbinom (0, 10, 0.15)

[1] 0.1968744

2) $p = 0.15$, $n = 20$

$P(X>3) = 1 - P(X \leq 3)$

> 1-pbinom (3, 20, 0.15)

[1] 0.3522748

5 $n = 5$, $p = 1/4$

$P(X \geq 3)$

$= 1 - P(X \leq 2)$

> 1-pbinom (3, 5, 1/4)

[1] 0.015625

6 $n = 10$, $p = 0.4$

```
x = 0:n
> prob = dbinom (x, n, p)
> cumprob = pbinom (0:10, 10, 0.4)
> d = data frame ("x values" = 0:10, "probability" = 0.4)
> print (d)
> plot (0:10, prob, "h")
```

experience as shown that 20% of student of a college smoks. A sample of 400 student reveal that out of 400 only 50 smoke. Test the hypothesis that the experience gives correct proportion are not

```
P=0.2
Q=1-P
Q
p=50/400
0.125
n=400
zcal=(p-P)/(sqrt(P*Q/n))
zcal
[1]-3.75
```

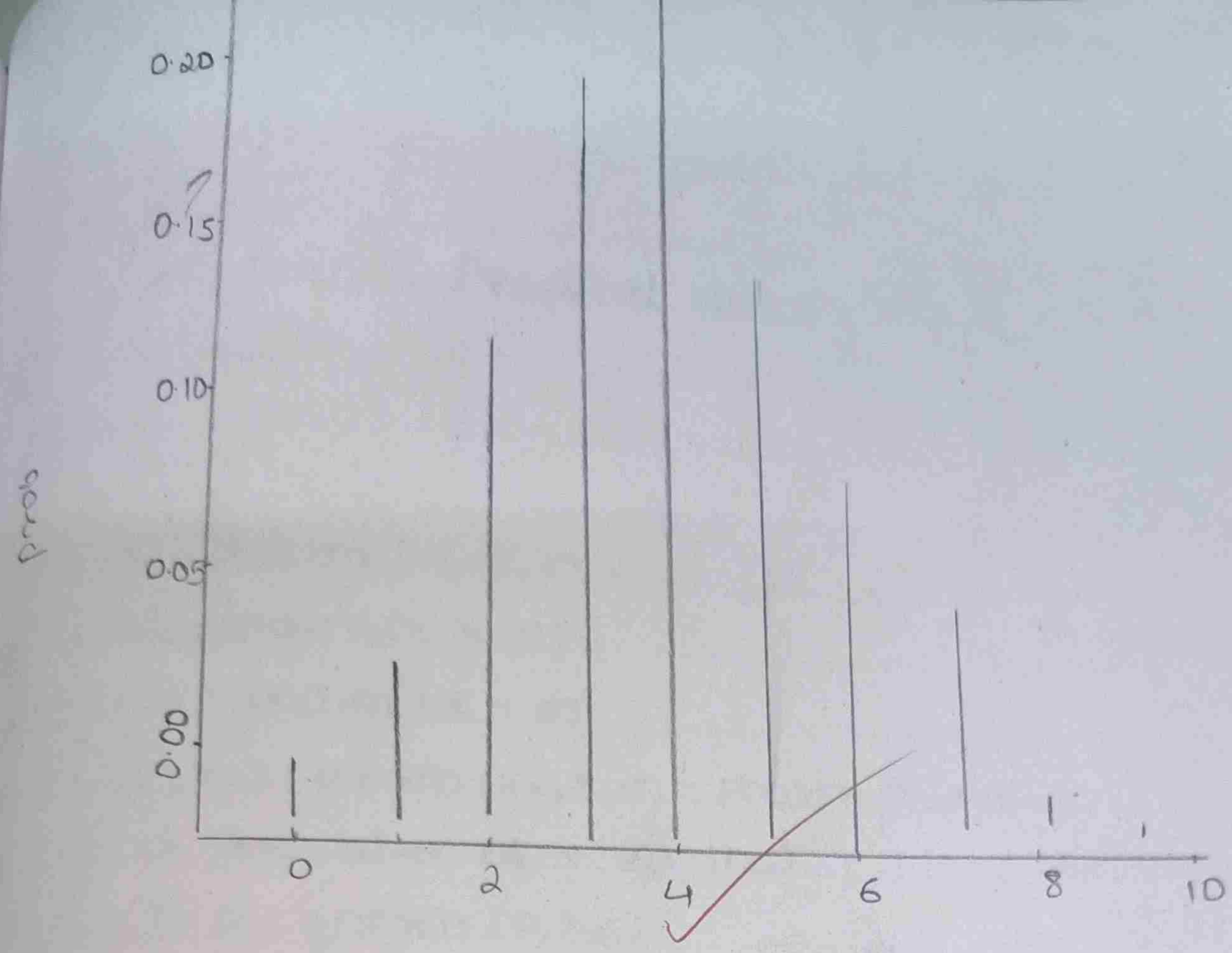
∴ value of is less then 0.5, therefore Ho is ~~incorrect~~ rejected.

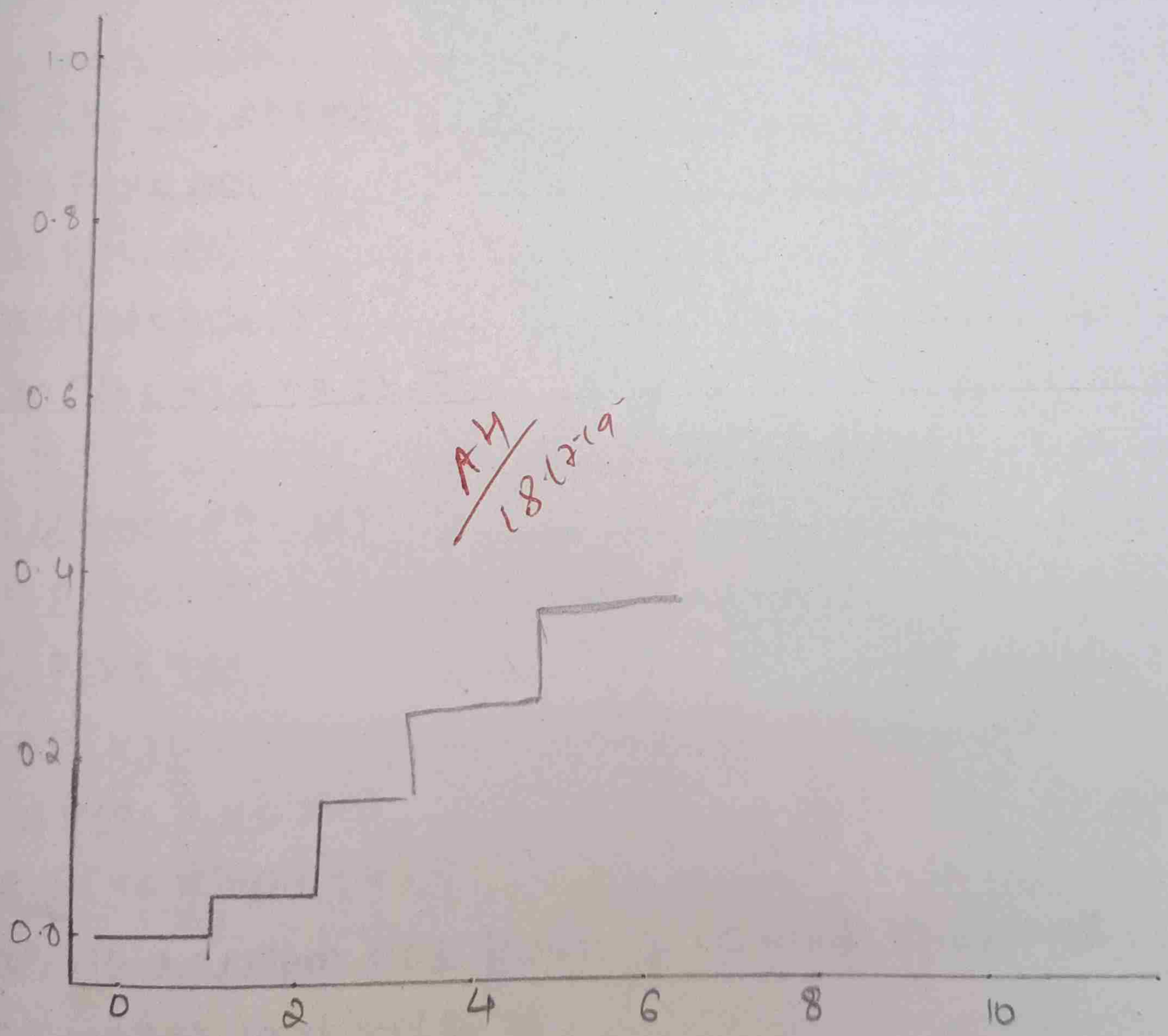
Test the hypothesis Ho: P=0.5 against H1: P≠0.5. A sample of 200 is selected and the sample proportion is calculated p=0.06. Test the hypothesis at 2% level of significance

```
P=0.5
Q=1-P
p=0.56
n=200
zcal=(p-P)/(sqrt(P*Q/n))
[1]1.697056
pvalue=2*(1-pnorm(abs(zcal)))
[1]0.08968602
```

∴ Value is more than 0.5 therefore Ho is ~~right~~ accepted

AM
22-01-20





AH
18-12-19

Practical No:- 5

Note :-

1. $P[X = x]$ = dnorm (x, μ, σ)
2. $P[X \leq x]$ = pnorm (x, μ, σ)
3. $P[X > x]$ = 1 - pnorm (x, μ, σ)
4. $P[x_1 < X < x_2]$ = pnorm (x_2, μ, σ) - pnorm (x, μ, σ)
5. To find the value of k so that
   $P[X \leq K] = p_1$; qnorm (p, μ, σ)
6. Two generaten random no. the command is
   rnorm (n, μ, σ)

1. $X \sim N\ (\mu = 50, \sigma^2 = 100)$

find (1) $P(X \leq 40)$

(2) $P(X > 55)$

(3) $P(42 \leq X \leq 60)$

(4) $P(X \leq K) = 0.7$; $k = 7$

2. $X \sim N\ (\mu = 100, \sigma^2 = 36)$

find (1) $P(X \leq 110)$

(2) $P(X \leq 95)$

(3) $P(X > 115)$

(4) $P(95 \leq X \leq 105)$

(5) $P(X \leq K) = 0.4$; $k = ?$

3. Generate 10 random nos from a normal distribution with mean $(\mu = 60)$ and S.D $(\sigma = 5)$
Also calculate the sample mean, median, variance and standard deviation

4. Draw the graph of standard normal distribution

Solution :-

1. >a = pnorm(40,50,10)
>cat("P(X<=40) is =",a)
[1] 0.1586553

2 >b = 1-pnorm(55,50,10)
[1] 0.3085375

3 >c = pnorm(60,50,10)-pnorm(42,50,10)
[1] 0.6294893

4. >d = qnorm(0.7,50,10)
[1] 55.24401

2.1. >a = pnorm(110,100,6)
>a
[1] 0.9522096
>b = pnorm(95,100,6)
>b
[1] 0.2023284

```
>c=1-pnorm (115,100,6)
>c
[1] 0.00620966 5
>d= pnorm (105,100, 6)- pnorm (95,100,6)
>d
[1] 0.5953432
>e=qnorm (0.4,100,6)
>e
[1] 98.47992
```

n=10, μ=60, σ = 5

```
>x=norm (10, 60, 5)
[1] 64.30801  56.28556  51.12444  62.66236  57.07212  52.90556  63.04319
59.37558  59.78891  56.45464

>am= mean (x)
>am
[1] 58.30124

>me= median (x)
>me
[1] 58.22385
>n=10
>variance = (n-1)*var (x)/n
[1] 17.05339
>sq. sqrt (variance)
[1] 4.129575
```

4. >x = seq(-3, 3, by = 0.1)

>x

[1] -3.0 -2.9 -2.8 -2.7 -2.6 -2.5 -2.4 -2.3 -2.2 -2.1

-2.0 -1.9 -1.8 -1.7 -1.6 -1.5 -1.4 -1.3 -1.2 -1.1

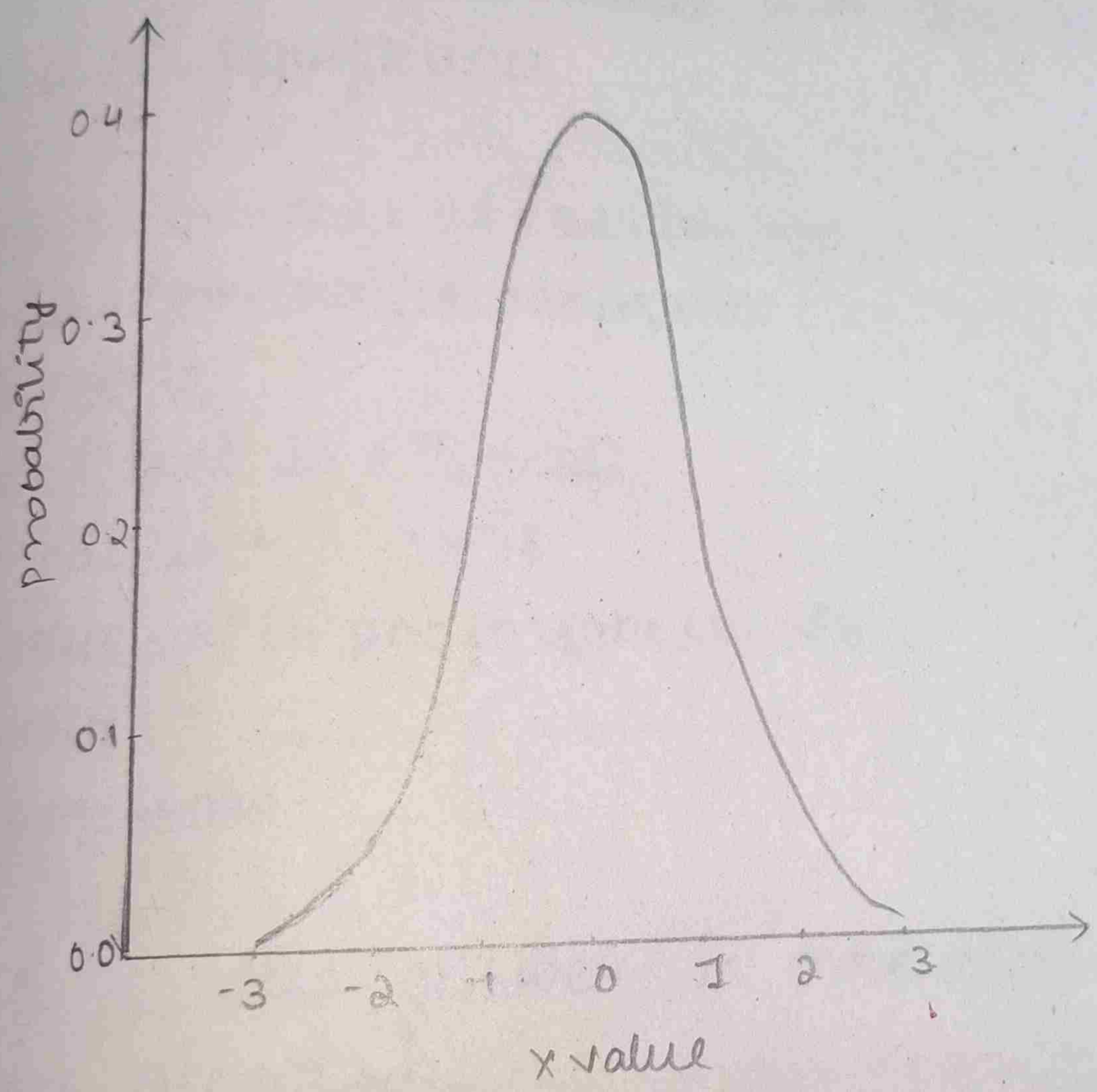
-1.0 -0.2 -0.3 -0.4 -0.5 -0.6 -0.7 -0.8 -0.9 -1.0

>y = dnorm(x)

>y

>plot(x, y, xlab = "x values", ylab = "probability", main = "standard normal graph")

Standard value graph



Ans
8.01.24

Practical No. 6

Topic :- Z distribution

Test the hypothesis (Ho) $H_0: \mu = 10$ against $H_1: \mu \neq 10$ sample of size 400 is selected which gives the mean 10.2 and standard deviation 2.25 Test the hypothesis at 5% level of significance

```
>mo=10; mx=10.2; sd=2.25; n=400
>zcal=(mx-mo)/(sd/(sqrt(n)))
[1]1.777778
>cat("zcal is + ",zcal)
[1]zcal is + 1.777778
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1]0.07544036
```

yes ∵ the answer is more than 0.05, ~~the value~~ of Ho is ~~right~~ accepted

Test the hypothesis $H_0: \mu = 75$ against $H_1: \mu \neq 75$ A sample of size 100 is selected and sample mean is 80 with standard deviation of 3. Test the hypothesis at 5% level of significance

```
>mo=75; mx=80; sd=3; n=100
>zcal=(mx-mo)/(sd/(sqrt(n)))
>zcal
[1]16.66667
```

```
>cat ("zcal is = ", zcal)
>zcal is = 16.66667
>pvalue = 2*(1-pnorm (abs(zcal)))
>pvalue
[1] 0
```

∴ The value is less then 0.5, the value of Ho is ~~incorrect~~ not accepted

3. Test the hypothesis Ho: $\mu_0 = 25$ against $H_1: \mu \neq 25$ at 5% level of significance. Following sample of 30 is selected
20, 24, 27, 35, 30, 46, 26, 27, 10, 20, 30, 37, 35, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 39, 27, 15, 19, 22, 20, 18

```
> x = c(          )
> mx = mean(x)
> n = length(x)
> variance = (n-1)* var (x) /n
> sd = sqrt (variance)
> sd
> mo = 25; mx = 26.07 ; sd = 7.28 ; n = 30
> zcal = (mx - mo)/(sd/(sqrt (n)))
> zcal
> cat ("zcal is = ", zcal)
> zcal is 0.8050318
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 0.4208013
```

The value is less than 0.5, therefore it is ~~incorrect~~ not accep

Practical No. 7

Large Sample tests

1. A study of noise level into hospital is calculated below. Test the hypothesis that the noise level into hospital. Test the hypothesis that the noise level in 2 hospital are same or not

	HOS A	HOS B
No. of sample obs	84	34
Mean	61	59
S.D	7	8

Ho: The noise level are same

n1 = 84

n2 = 34

mx = 61

my = 59

sdx = 7

sdy = 8

```
> z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))
[1] 1.273682
> cat("z calculated is=", z)
> pvalue = 2*(1-pnorm(abs(z)))
[1] 0.1291361
```

Since pvalue is 0.12, we reject Ho at 5% level of significance

Two random sample of size 1000 and 2000 are drawn from 2 population with the means 67.5 and 68 respectively and with same ~~standard dev~~ S.D of 2.5 Test the hypothesis the mean of 2 population are equal

Ho = 2 population are equal

```
n1 = 1000
n2 = 2000
mx = 67.5
my = 68
sdx = 2.5
sdy = 2.5
>z = (mx - my)/ sqrt ((sdx^2/n1) + (sdy^2/n2))
[1] -5.16
>cat ("z calculated is = ",z)
>pvalue = 2*(1-pnorm(abs(z)))
>pvalue
[1] 2.417564e-07
```

Since pvalue is 2.417564e-07, we reject Ho at 5% level of significance

3. In fybsc 20% of a random sample of 400 students had defective eye sight in ~~fy~~ class 15.5% of 500 sample had the same defect is the difference of proportion is same?

Ho = The proportion of population is equal

```
>n1 = 400
>n2 = 500
>p1 = 0.2
>p2 = 0.155
> p = (n1* p1 + n2* p2) / (n1+n2)
[1] 0.175
> q = 1-p
[1] 0.825
> z = (p1-p2) / sqrt (p* q (1/n1-1/n2))
[1] 1.76547
>cat ("z calculated is = ", z)
>pvalue = 2* (1-pnorm (abs (z)))
>pvalue
[1] 0.07748487
```

Since pvalue is 0.7748487, we ~~reject~~ accept Ho at 5% level of significance

4. From each of the box of the apple an sample size of 200 is collected. It is found that there are 44 bad apple in the 1st sample and 30 bad apple in 2 sample test the hypothesis that the 2 boxes are equivalent in terms of number of bad apple

Ho : The proportion of boxes is same

```
>n1 = 200
>n2 = 200
>p1 = 44/200
>p2 = 30/200
>p1 = 0.22
>p2 = 0.15
>p = (n1*p1 + n2*p2)/(n1+n2)
[1] 0.185
> q = 1 - p
[1] 0.815
>z = (p1 - p2)/sqrt(p*q*(1/n1+1/n2))
[1] 1.802741
>cat("z calculated is =", z)
>pvalue = 2*(1-pnorm(abs(z)))
[1] 0.07142888
```

∴ Since pvalue is 0.07142888, we ~~reject~~ accept at 5% level of significant

5. In a MA class out of the sample of 60 mean height is 63.5 inch with a S.D 2.5. In an M.com class out of 50 student mean height 69.5 inch with S.D of 2.5 test the hypothesis that the mean of MA & M.com class are same.

Ho = Mean height of MA and Mcom student are equal

```
>n1 = 60
>n2 = 50
>mx = 63.5
>my = 69.5
>sdx = 2.5
>sdy = 2.5
>z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))
[1] -12.5335
>cat ("z calculated is=", z)
>pvalue = 2*(1-pnorm(abs(z)))
[1] 0.04550026
```

Since pvalue is 0.04550026, we reject at 5% level of significant

Am
5/3/20

Small Sample test

The tens are selected and height are found to be 63, 63, 68, 69, 71, 71, 72 cms Test hypothesis that mean height are 66 cms or not at 1

Ho = Mean = 66 cms

```
>mean = 66
>x = c(63, 63, 68, 69, 71, 71, 72)
>t.test(x)
one sample t.test
data: x
t = 47.94, df = 6, Pvalue = 5.22 e-09
alternative hypothesis mean is not equal to 0
95 percent confidence interval
 64.66479      71.62092

Sample estimates
mean of x
 68.14286
```

pvalue < 0.01 is rejected or in 1% level of significane

2. Two random sample was drawn from two different population

Sample 1 = 8, 10, 12, 11, 16, 15, 18, 7

Sample 2 = 20, 15, 18, 9, 8, 10, 11, 12

test the hypothesis that there is no difference between the two population mean at 5% level of significance

→ Ho there is no difference in the population mean

```
> x = c (8,10,12,11,16,15,18,7)
> y = c (20,15,18,9,8,10,11,12)
> t.test (x, y)
```

Welch two sample t-test

data: x & y

t = -0.36247, df = 13.837 · p-value = 0.7225

alternative hypothesis : true difference in mean is not equal

-5.192719 3.692719

Sample estimates

mean of x mean of y

12.125 12.875

p-value < 0.01 is accepted in Ho on 1% level of significance

3. following are the weight of 10 people

Before = (100, 125, 95, 96, 98, 102, 115, 104, 109, 110)
After = (95, 80, 95, 98, 90, 100, 110, 85, 100, 101)
Ho = The diet program is not affective

```
> x=c(100, 125, 95, 96, 98, 102, 115, 104, 109, 110)
> y=c (95, 80, 95, 98, 90, 100, 110, 85, 100, 101)
> t.test (x, y, paired = T, alternative = "less")
```

Paired t-test

data : x and y
t = 2.3215 , df = 9 , p-value = 0.9773
alternatives hypothesis : true difference in means is less than 0 95 percent confidence interval
-Inf 17.89635
Sample estimates
mean of the differences
10

p-value < 0.01 is accepted in Ho on 1% level of significance

4. Marks before & after a training program is given below

before = 20, 25, 32, 28, 27, 36, 35, 25
after = 30, 35, 32, 37, 37, 40, 40, 23
Test the hypothesis that training program is effective or not

Ho. The training program is not effective

```
> x = c (20,25,32,28,27,36,35,25)
> y = c (30,35,32,37,37,40,40,23)
> t.test (x,y, paired = T, alternatives = "greater"
```

Paired t-test

data: x and y
t = -3.3859, df = 7, p-value = 0.9942
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval
-8.967399 Inf
Sample estimates
mean of the difference
-5.75

pvalue < 0.01 is accepted in Ho at 1% los

5. 2 random sample were drawn from two normals population and the 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% level of significance

Ho : variance of the population are equal

```
> x= c (66,67, 75, 76, 82, 84, 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 = 0.70686, num df = 8, denom df = 10 p-value = 0.6359
95% confidence interval
0.1833662 3.0360393
sample estimates
ratio of variance
0.7068567

Ho: population mean = 55

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

AM
7/6/2

Pearson's Chi-squared test

data: y

x-squared - 25.646, df - 2, p-value = 2.698 e-06

Pvalue is less than 0.05 we reject Ho at 5% LOS

2. Table below show the relation between the performances of mathematics and computer of CS students

		Maths		
		HG	MG	LG
	HG	56	71	12
Comp.	MG	47	163	38
	LG	14	42	85

Ho: performance between maths and computer are independent

x = c (56, 47, 14, 71, 163, 42, 12, 38, 85)

m = 3

n = 3

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

	[, 1]	[, 2]	[, 3]
[1,]	56	71	12
[2,]	47	163	38
[3,]	14	42	85

Practical No 9

Chi square and ANOVA

Use the following data to test whether the condition of the home depands upon the child condition or not

		Cond. of home	
		clean	Dirty
Cond. of child	clean	70	50
	fairly dirty.	80	20
	Dirty	35	45.

→ Ho Condition of the home and child are independent

```
> 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

```
> pvalue = chisq.test (y)
> pvalue
```

pvalue.chiq.test(y)
pvalue

datay
x-squared = 145.78, df = 4, p-value < 2.2e-16

3. Varieties Observation

A	50, 52
B	53, 55, 53,
C	60, 58, 57, 56
D	52, 54, 54, 55

mean 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))
> oneway.test (values~ind, data=d, var.equal=T)
anova = aov (values ~ ind, data = d)
F = 11.735, num df = 3, denom df = 9, p-value = 0.00183
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			

∴ Pvalue is less then 0.05 we reject Ho at 5% LOS

```
> n = sp+sn
> pv = pbinom (sn, n, 0.5)
[1] 0.2539063
v = 29, p-value = 0.2386
```

Note: If the alternative is less we have to
wilcox.test (x, alter = "greater", mu=60)

If the alternative is not equal to
wilcox.test (x, alter = "2.sided", mu=60)

use wilcoxkson sign, test the hypothesis where median is 12 against the alternative is less than 12

12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20

```
x = c(12,13,10,20,15,5,1,7,6,11,9,20)
[1] 12 13 10 20 15 5 1 7 6 11 9 20
> m = 12
> sp = length (x[x>m])
> sn = length (x[x>m])
> n = sp+ sn
> pv = pbinom (sn, n, 0.5)
[1] data:x
v = 25 , pvalue = 0.2521
```

AN
4.3.20

2. For the observation 12, 19, 31, 28, 43, 40, 55, 49, 70, 63 apply sign test to test population medium is 25 against the alternative is more than 25

```
> x = c(12,19,31,28,43,40,55,49,70,63)
[1] 12 19 31 28 43 40 55 49 70 63
> m = 25
> sp = length(x[x > m])
> sn = length(x[x < m])
> n = sp + sn
> pv = pbinom(sn, n, 0.5)
> pv
[1] 0.0546875
```

3. For the following
60, 65, 63, 89, 61, 71, 58, 51, 48, 66. Jest the hypothesis using wilcockson sign test. for tested the hypothesis that median is 60 against the alternative it is greater than 60.

H_0: Median is 60

H_1: It is greater than 60

```
> x = c(60,65,63,89,61,71,58,51,48,66)
[1] 60 65 63 89 61 71 58 51 48 66
> m = 60
> sp = length(x[x > m])
> sn = length(x[x < m])
```

Non-paramatric test

Following are the amounts of sulphur oxide by a factor
17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26
apply sign test to test the hypothesis that the population medium is 21.5 against the alternative is less than 21.5

Ho: population median = 21.5
H1: It is less than 21.5

```
>x = c(17,15,20,29,19,18,22,25,27,9,24,20,17,6,24,14,15,23,24,26)
>x

[1] 17 15 20 29 19 18 22 25 27 9 24 20 17 6 24 14 15 23 24 26
>m = 21.5
>sp = length(x[x>m])
>sn = length(x[x<m])
>n = sp+sn
[1] 20
>pv = pbinom(sp,n,0.5)
[1] 0.4119015
```

If the alternative is greater then medium

```
pv = pbinom(sn,n,0.5)
```

Terms

	Ind	Residuals
sum of square	18	18
Deg of freedom	3	8

Residual standard error : 1.5
Estimated effects may be unbalanced

> summary (anova)

	Df	sum sq	mean sq	Fvalue	pr(>F)
ind	3	18	6.00	2.667	0.119
residuals	8	18	2.25		

Since the pvalue is greater than 0.05 we accept the hypothesis at 5% LOS

Perform's anova on the following table

Types	Observation
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

a = c 6, 7, 8)
b = c (4, 6, 8)
c = c (8, 6, 10)
d = c (6, 9 9
d = stack

Oneway test (values mind, data = d, var. equal = T)

One-way analysis of means
data: values and ind
F = 2.6607, num df = 3, denom df = 8, p-value = 0.1
* anova = aov (value ~ ind, data = d)
> anova

Call
aov (formula = values ~ ind, data = d)