Practical - 1

Aim - Basic of R Software

1) R is a software for statistical analysis and data
2) It is effective data handling software and outcome storage.
3) Capable of graphical display.
4) It is free.

1) 4+6+ 8÷ 2-5

```
> 4+6+ 8 ÷2-5
[1] 9
```

2) $2^2+1-3+\sqrt{5}$

```
> 2^2 + abs (-3) + sqrt (45)
[1] 13.7082
```

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

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

10

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

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

[1] 5.671567

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

[1] 4 6 10 14

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

[1] 4 9 10 21

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

[1] 4 36 512 9162 125.

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

[1] 1.50 0.40 1.75 1.00

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

x^2 + y^3 + z

[1] 27402

> sqrt (x*z+y)

[1] 20.73644

> x^2+y^2

[1] 1300

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

5) y <- matrix (nrow = 3, ncol = 3, data = c (10,12,15,-5,-4,-6)

y <- [,1] [,2] [,3]

[1,] 10 -7 13

[2,] 12 -4 16

[3,] 15 -11 8

50

3) Marks of stastics of CS Batch B

x = c (58, 20, 35, 24, 46, 46, 56, 55, 45, 27, 22, 47, 58, 54, 40, 50, 3
29, 35, 39)

) x = c(data)

mark = seq (20, 60, 5)

a = cut (x, breaks, right = FALSE)

b = table (a)

c = transform (b)

c

a	freq
[20, 25]	3
[25, 30]	2
[30, 35]	1
[35, 40)	4
[40, 45)	1
[45, 50]	3
[50, 55]	2
[55, 60]	4

Q4

Practical 2

→ Probability Distribution

1) Check is following pmf or not

x	P(x)
0	0.1
1	0.2
2	0.3
3	0.4
4	0.3
5	0.5

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

$\therefore 0.1 + 0.2 - 0.5 + 0.4 + 0.3 + 0.5$

$= 1.0$

$\therefore P(x) = 0.5$ & cannot be probability mass function

2)

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

1) $P(x) \geq 0$

$P(x) = 1$

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

$= 1$

$\therefore$ It is pmf

Code

```
Prob = c(0.2, 0.2, 0.3, 0.15, 0.1)
sum(Prob)
1
```

2) Find cdf for following pmf and sketch graph.

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

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

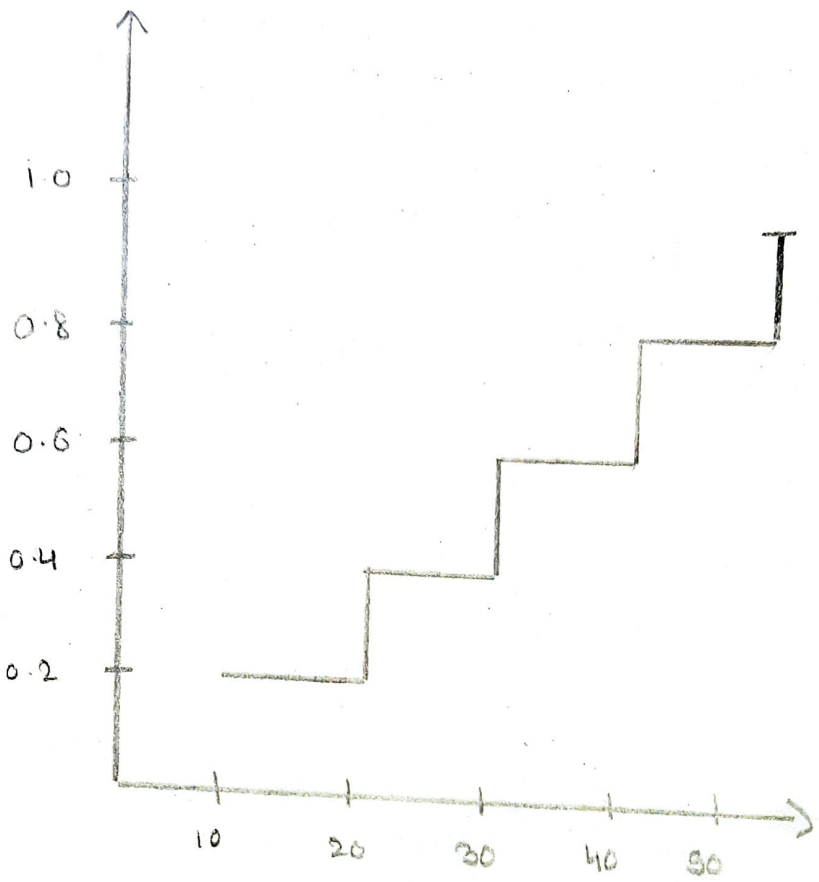
$0.2 \quad 10 \leq x < 20$

$0.4 \quad 20 \leq x < 30$

$0.75 \quad 30 \leq x < 40$

$0.90 \quad 40 \leq x < 50$

$1.0 \quad x \geq 50$



x = c (10, 20, 30, 40, 50)

plot = (x, cumsum (Prob, "s"))

Q2) Find

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

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

0.15 $1 \leq x < 3$

0.40 $2 \leq x < 3$

0.50 $3 \leq x < 4$

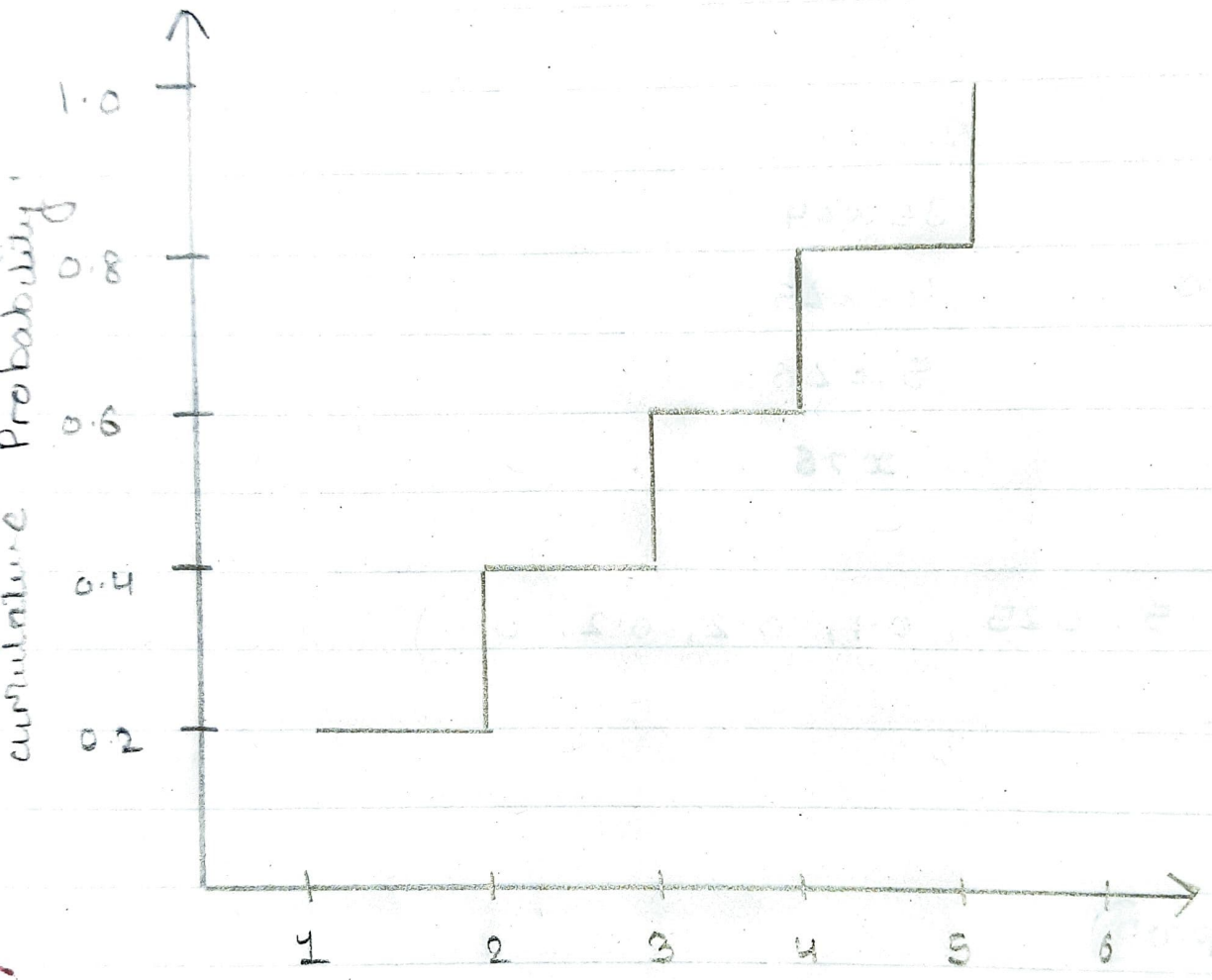
0.70 $4 \leq x < 5$

0.90 $5 x < 6$

1.00 $x \geq 6$

```
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 , 0.50 , 0.10 , 0.90 , 1.00
> x=c (1,2,3,4,5,6)
plot (x, cumsum (prob) , "s" .xlab = "value" , ylab = "cumsum
    probability" , main = " CDF graph " , col = "brown")
```


cumulative Probability
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
values.

Pratical 3

→ Binomial Distribution

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

P(x≤x) = pbinom (x, n, p)

If x is unknown

p1 = P(x≤x) = qbinom (p, n, p)

1) Find probability of exactly 10 success in 100 trials with p=0-1

2) Suppose there are 12 req. each question has 5 option out of which 1 is correct
Find probability of having exactly 4 correct answer.

3) Find complete distribution when n = 5 and p=0-1

---

ans)

1) > x = dbinom (10, 100, 0·1)
> x
[1] 0·1318653

2) dbinom (4, 12, 0-2)
0·1328756
pbinom (4, 12, 0-2)
0·4274445
1-pbinom (5, 12, 0·2)
0·01940528·

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

0 - 0.59049

1 - 0.32805

2 - 0.07290

3 - 0.0810

4 - 0.00045

5 - 0.00001

4) dbinom (5, 12, 0.25)

0.1032414

pbinom (5, 12, 0.25)

0.9455978

1 - pbinom (7, 12, 0.25)

0.00278151

dbinom (6, 12, 0.25)

0.04014945

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

7) n = 10

p = 0.3

x = 0:n

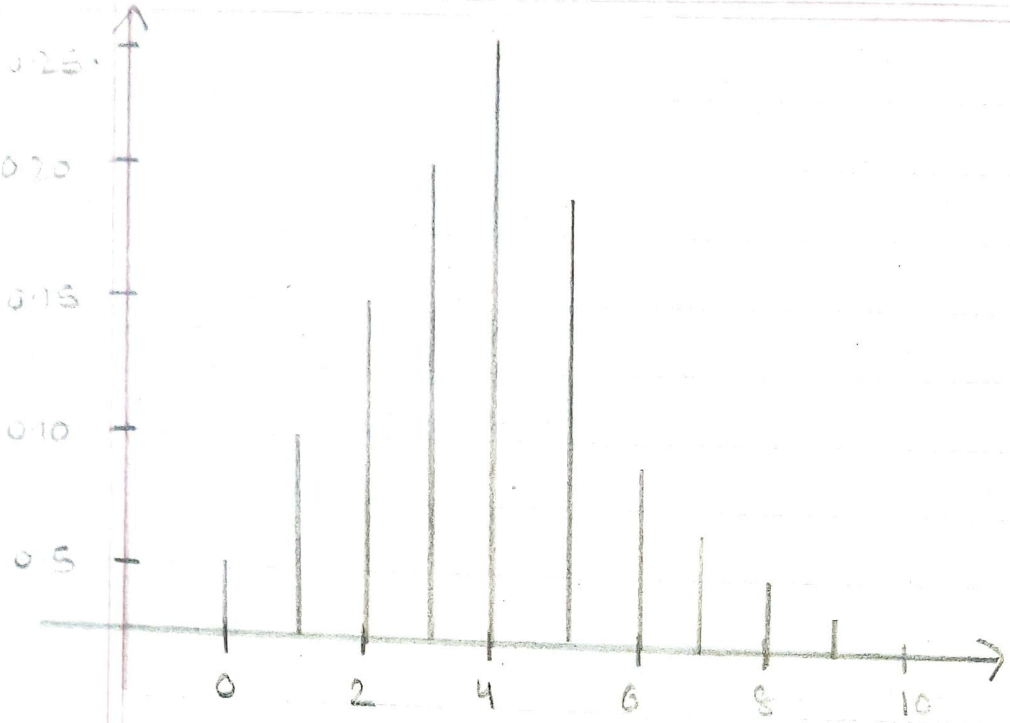
prob = dbinom (x, n, p)

cumprob = pbinom (x, n, p)

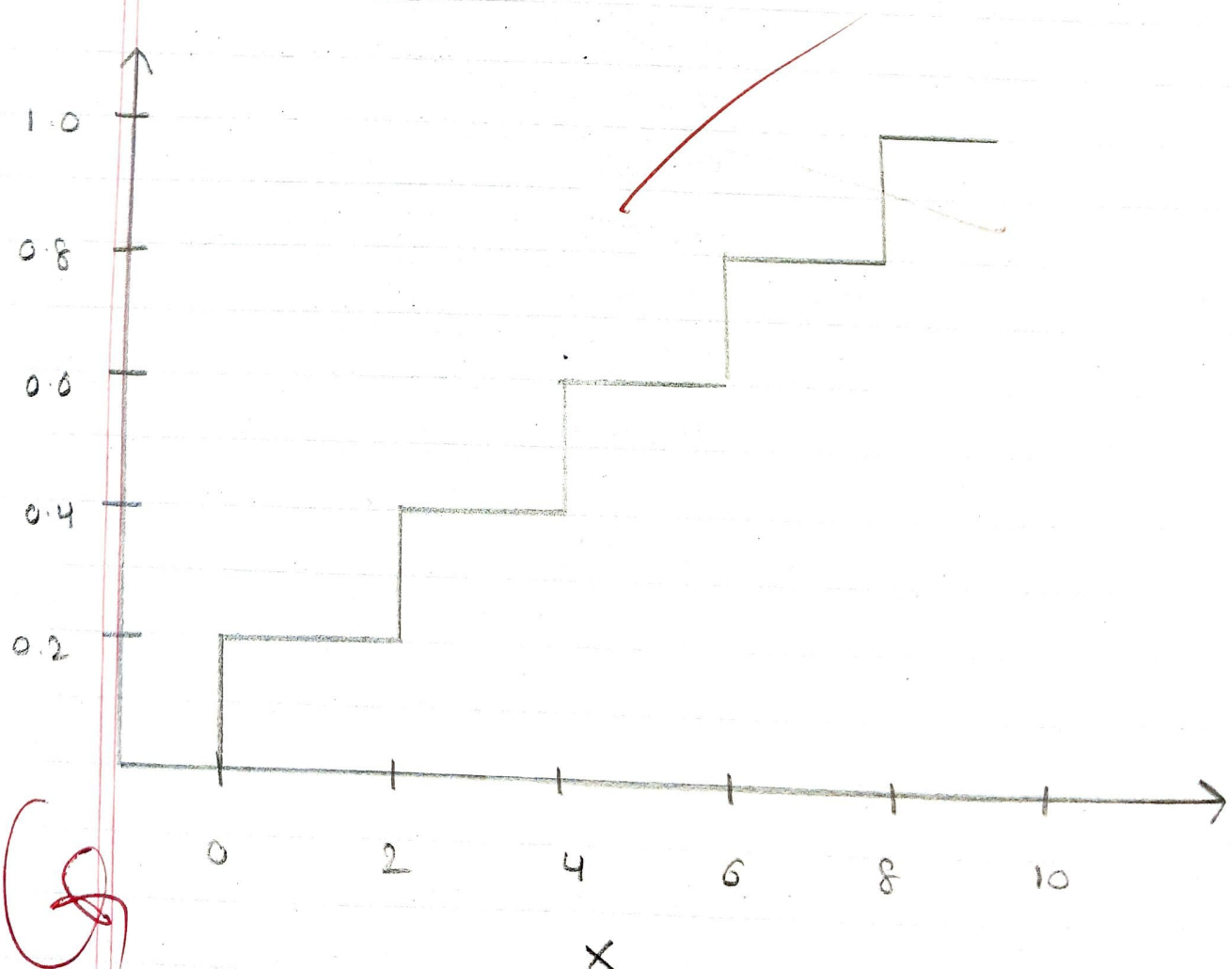
d = data.frame ("x values", = x, "probability" = prob).

print (d)

	x values	probability
2	0	0.0282
3	1	0.1210
4	2	0.2234
5	3	0.2668
6	4	0.2001
7	5	0.1029
8	6	0.0367
9	7	0.0090
10	8	0.0014
11	9	0.001
	10	0.0000



plot (x, cw2prob = "s")



Practical 4

Normal Distribution

1) P(x = x) = dbinom(x, u, σ)

P(x ≤ x) = pnorm(x, u, σ)

P(x > x) = 1 - pnorm(x, u, σ)

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

1) A random var x follows nor. distribution with mean = u = 12 and SD = σ = 3.

Find ① P(x ≤ 15)

P(10 ≤ x ≤ 13)

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(13, 12, 3) - pnorm(10, 12, 3)
p2
[1] 0.3780661
```

cat (p(10 < x < 13) = ", p2)

p3 = 1 - pnorm (14, 12, 3)

p3

[1] 0.252495

cat ("P (x > 14) = ", p3)

P (x > 14) = 0.6252495

p4 = pnorm (5, 12, 3)

p4

[1] 15.254723 16.548505 11.280515

6.419944 12.272460.

2) $x \sim N(30, 100)$, $\sigma = 10$

1) P(x ≤ 40)

P(x > 35)

P(25 < x < 35)

Find k such that P(x < k) = 0.6

F = pnorm (0.6, 30, 10)

F

0.8413447

F1 = 1 - pnorm (35, 30, 10)

F1

0.3085375.

F2 = pnorm (55, 30, 10) - pnorm (0.6, 30, 10)

F2

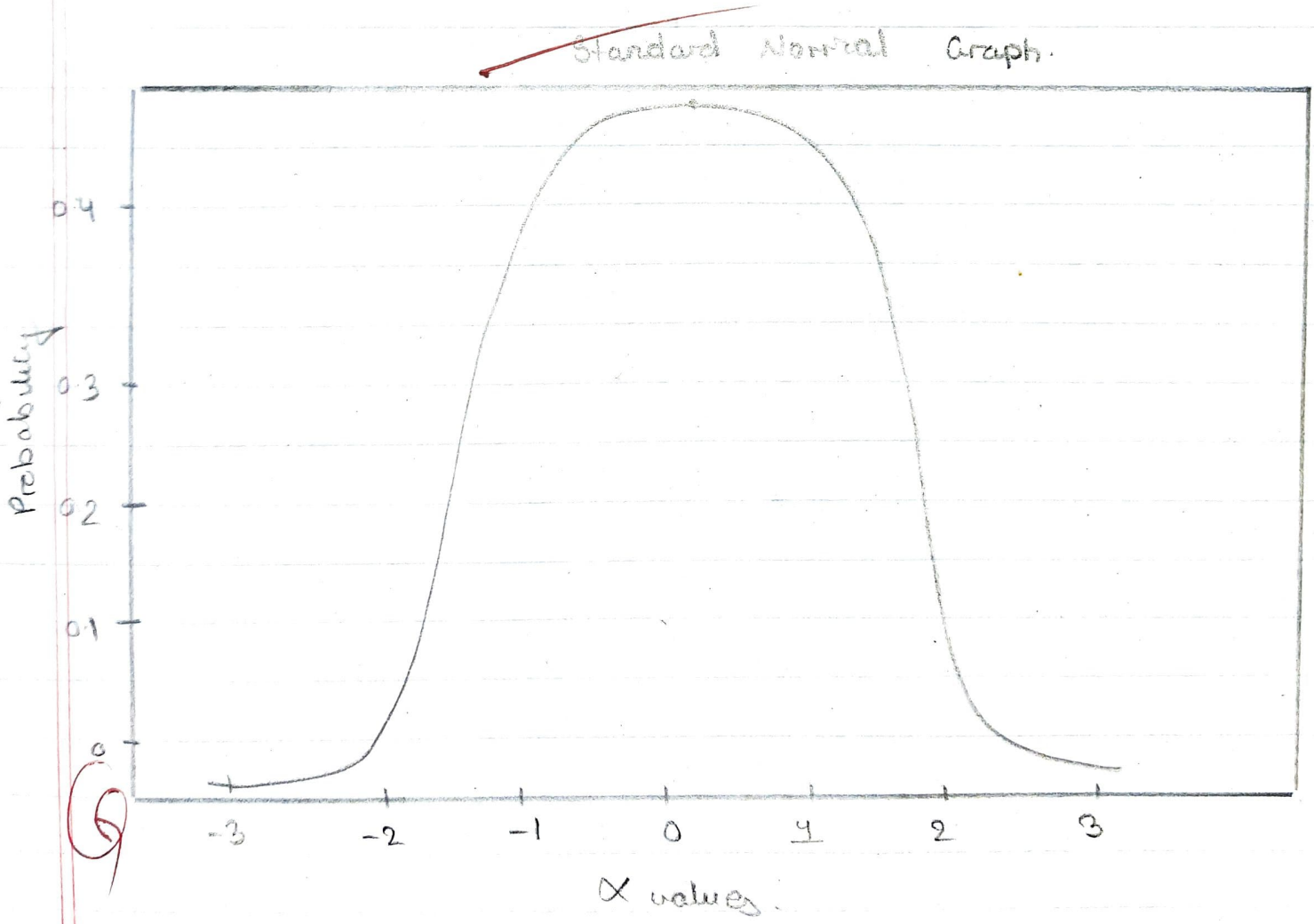
0.3829249.

```
FB = qnorm (0.6, 30, 10)
FB
32.53347
```

5) Plot standard deviation 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 5

Normal and t-test

0 H_0 : $\mu = 15$ μ_1 : $\mu \neq 15$

Test the hypothesis

Random sample of size 400 is drawn and it is calculated.

The sample mean is 14.

SD is 3.

Test hypothesis at 5% level of significance

code :

```
m0 = 15
mx = 14
n = 400
sd = 3
zcal = (mx - m0) (sd / (sqrt (n)))
zcal
-6.666667
cat ("calculate value of z is = ", zcal)
calculated value of z is = -6.666667
pvalue = 2 * (1 - pnorm (abs (zcal)))
pvalue
2.616716e-11
```

value < 0.05 we will ~~get~~ reject value of $H_0 = \mu = 15$.

(2) $H_0: \mu = 12.5$ from following sample at 5% level of significance.

```
x = c(12.25, 11.97, 12.15, 12.08, 12.31, 12.28, 12.31, 12.28, 11.94, 11.89, 12.16
      12.04)
n = length(x)
n
10
mx = mean(x)
mx
12.07
var = (n-1) * var(x)/n
0.019521
sd = sqrt(var)
sd
0.1397176
m0 = 12.5
t = (mx-m0) / (sd/sqrt(n))
t
-8.894909
pvalue = 2 * (1-pnorm(abs(t)))
pvalue
0
```

∴ The value is less than 0.05 the value is accepted.

⑤ Test hypothesis $H_0 = \mu = 10$ against $H_1 : \mu \neq 10$. A random sample size of 400 is drawn with sample mean = 10.2 and sd = 2.25.

```
m0 = 10
n=400
xbar = 10.2
sd = 2.25
zcal = (xbar-m0) / (sd / (sqrt (n)))
zcal
 1.77778
pvalue = 2* (1-pnorm (abs (zcal)))
pvalue
0.07544036
```

Value is greater than 0.05

It is accepted.

11

Last year farmer lost 20% of crop. A random sample of 60 fields are collected and found a field crop are insect polluted

Test hypothesis at 1% significance

$p = 0.2$

$p = 9160$

$n = 60$

zcal (p-p) / (sqrt (p * (1-p)/n))

zcal

0.9687458

pvalue = 2 * (1-pnorm (abs (zcal)))

pvalue

0.3329216

Value is accepted.