Solve the ~~flo~~ following

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

```
> 4 + 6 + 8/2 - 5
[1] 9
```

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

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

$5^3 + 7 \times 5 \times 8 + 46/5$

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

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

```
> sqrt(4^2 + 5*3 + 7/6)
[1] 5.671567
```

round off
46÷7+9×8

, round (46/7+9*8)

[1] 79

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)

] 4 9 30 14

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

[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,

[1] 4 36 512 9 16 5

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

] 1.50 0.40 1.75 1.00

x=20 > y=30 > z=2

x^2 + y^3 + z

] 27402

sqrt (x^2 + y)

] 20.73644

x^2 + y^2

] 1300

Q4
>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
[3,]	3	7
[4,]	4	8

Q5. find x+y & 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, 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

> 2 * x + 3 * y	[,1]	[,2]	[,3]
[1,]	38	-19	33
[2,]	50	-12	41
[3,]	63	-28	21

Marks of students of CS Batch A

x = c(59, 20, 35, 24, 46, 56, 55, 45, 45, 27, 22, 47, 58, 50, 32, 36, 29, 35, 39)

```
> x = c(data)
> breaks = seq (20, 60, 5)
> a = cut (x, breaks, right = FALSE)
> b = table (a)
> c = transform (b)
> c
```

	a	freq
1	[20, 25)	3
2	[25, 30)	2
3	[30, 35)	1
4	[35, 40)	4
5	[40, 45)	1
6	[45 50)	3
7	[50 55)	2
8	[55 60)	4

Practical 2.

Probability distribution.

Check whether the following are p.m.f 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 $\Sigma P(x) = 1$

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

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

$= 1.0$

$\because P(2) = -0.5$, it can't be a probability mass funct

$\because P(x) \geq 0 \ \forall, x$

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 $\sum P(x) = 1$ so,

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

$0.2 + 0.2 + 0.3 + 0.2 + 0.2$

$= 1.1$

$\therefore$ The given data is not a ~~Point~~ P.m.f because the $P(x) \neq$

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) $\sum P(x) = 1$

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

$0.2 + 0.2 + 0.35 + 0.15 + 0.1$

$= 1$

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

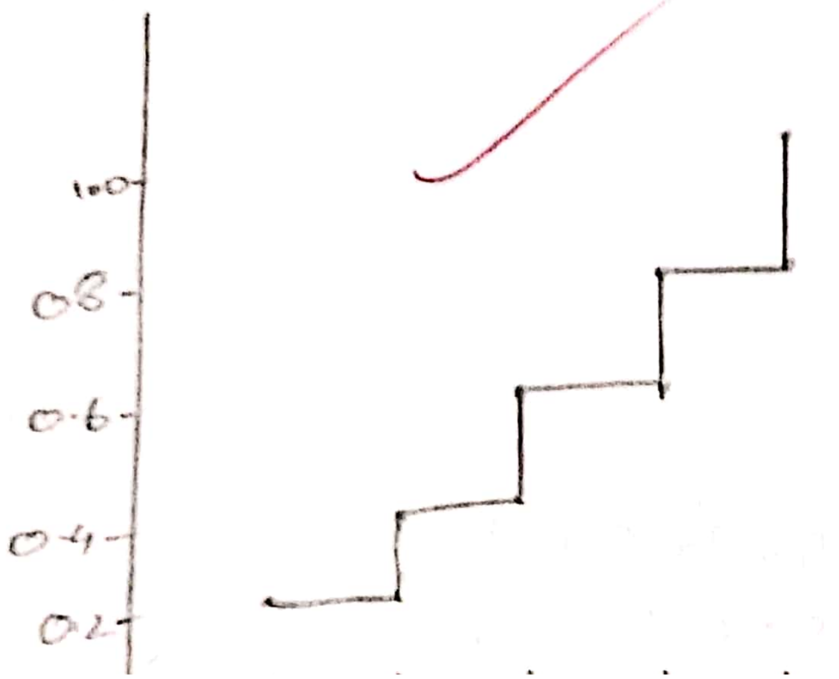
Code:

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

2. 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

$$f(x) = \begin{cases} 0 & x < 10 \\ 0.2 & 10 \le x < 20 \\ 0.4 & 20 \le x < 30 \\ 0.75 & 30 \le x < 40 \\ 0.90 & 40 \le x < 50 \\ 1.0 & x \ge 50 \end{cases}$$



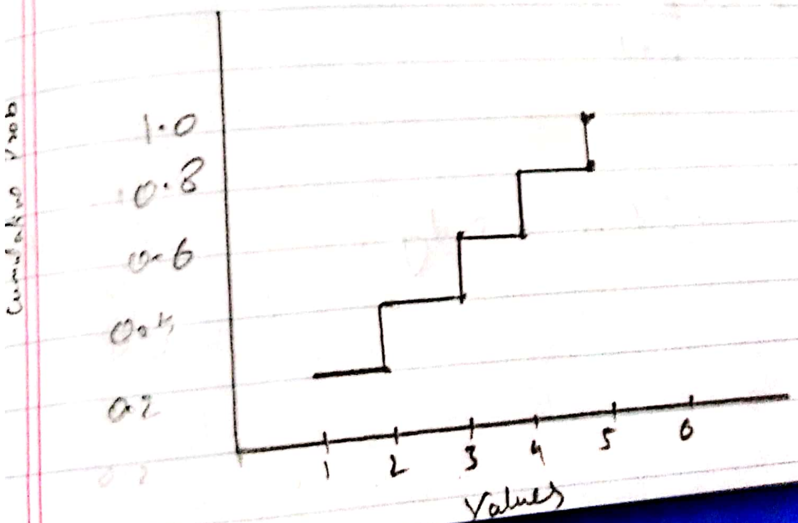
x 1 2 3 4 5 6

p(x) 0.15 0.25 0.1 0.2 0.2 0.1

p(x) = 0	x < 1
0.15	1 ≤ x < 2
0.40	2 ≤ x < 3
0.50	3 ≤ x < 4
0.70	4 ≤ x < 5
0.90	5 ≤ x < 6
1.00	x ≥ 6

```
>Prob = c(0.15, 0.25, 0.1, 0.2, 0.02, 0.1)
>sum(Prob)
[1] 1
>cumsum(Prob)
[1] 0.15, 0.40, 0.50, 0.70, 0.90, 1.00
>x = c(1, 2, 3, 4, 5, 6)
>Plot(x, cumsum(Prob), "s" xlab = "Value", ylab = "Cumul
Probability",
  main = "CDF Graph", col = "brown")
```

CDF Graph



1.8.

3. Check that whether the following is p.d.f or not

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

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

$\int_0^1 f(x)$

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

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

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

$\because$ The $\int_0^1 f(x) \neq 1$ $\therefore$ It is not a pdf

2) $f(x) = 3x^2$; $0 < x < 1$

$\int_0^1 f(x)$

$\int_0^1 3x^2$

$3\int_0^1 x^2$

$\left[3\, x^3/3\right]_0^1$ $\because \int x^n = \dfrac{x^{n+1}}{n+1}$

$= x^3$

$= 1$

$\because$ The $\int_0^1 f(x) = 1$ $\therefore$ It is a p.d.f.

Scanned with CamScanner

> x= dbinom (10, 100, 0.1)
> x
[1] 0.1318653

i) dbinom (4,12,0.2)
[1] 0.1328756
ii) pbinom (4, 12, 0.2)
[1] 0.927445
iii) 1- Pbinom (5, 12, 0.2)
[1] 0.01940528

dbinom (0:5, 5, 0.1)

0 - 0.59049
1 - 0.32805
2 - 0.07290
3 - 0.00810
4 - 0.00045
5 - 0.00001

i) dbinom (5, 12, 0.25)
[1] 0.1032414
2) pbinom (5, 12, 0.25)
3) 1- pbinom (7, 12, 0.25)
[1] 0.00278151
4) dbinom (6, 12, 0.25)
[1] 0.04014945

1.8

Practical - 3

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

$P_1 = P(X \le x)$ = qbinom (P_1, n, p)

1) Find the probability of exactly 10 success in hundred trials
$p = 0.1$

2) Suppose there are 12 mcq, each question has 5 option and 1 is correct find the probability of having exactly 4 correct ans
i) Atmost 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 probability.

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

> dbinom (0, 10, 0.15)

[1] 0.1968744

> pbinom (3, 20, 0.15)

[1] 0.3522748

qbinom (0.88, 30, 0.2)

[1] 9

7) > n = 10

> p = 0.3

> x = 0 : n

> Prob = dbinom (x, n, p)

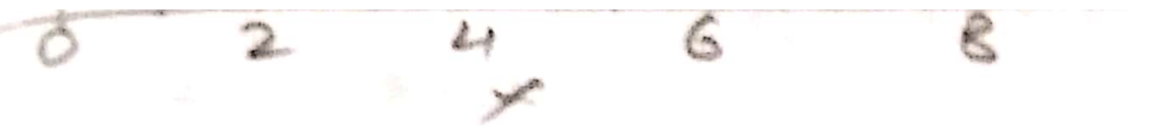
> cumProb = pbinom (x, n, p)

> d = data.frame ("x values" = x", "Probability" = Pro

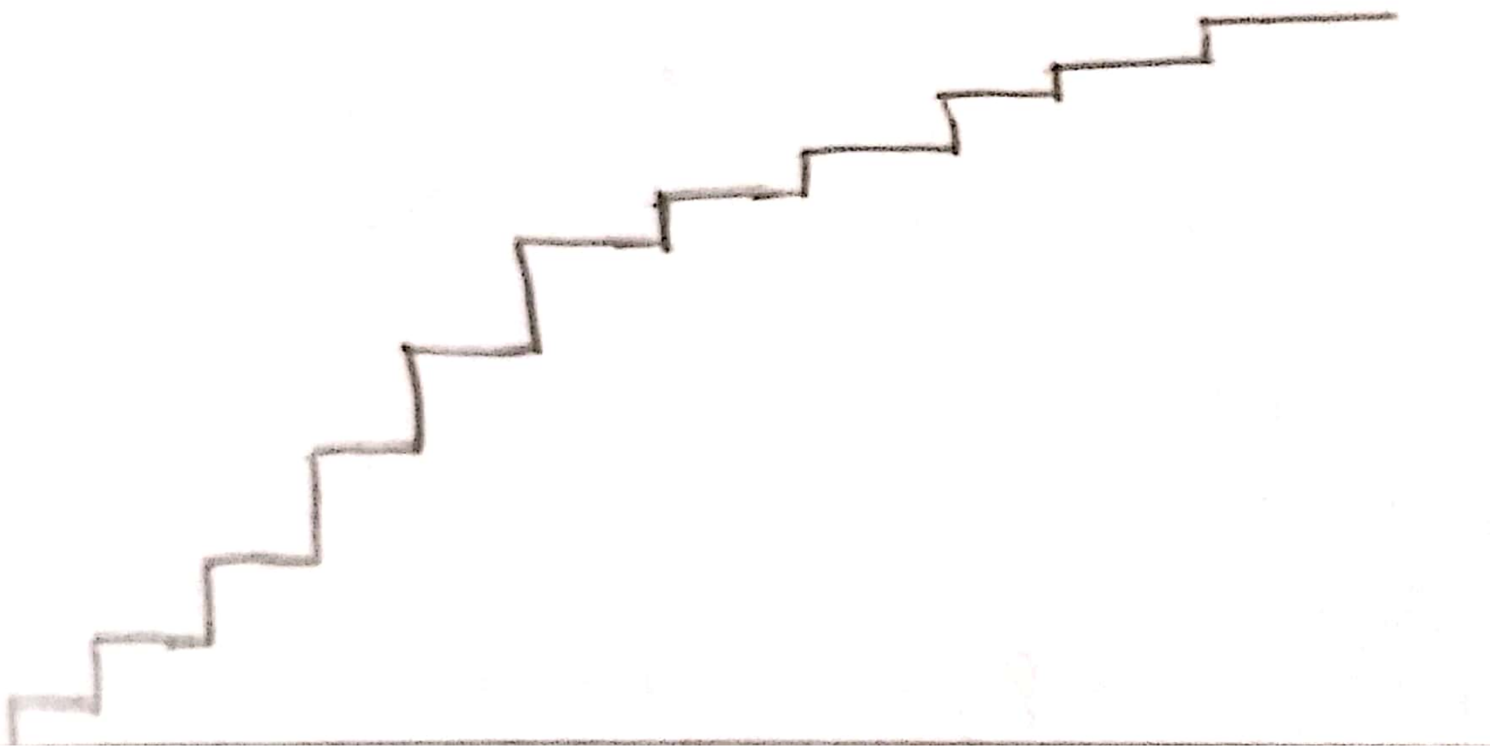
> Print (d)

	x values	Probability
1	0	0.0282
2	1	0.1210
3	2	~~0.1210~~ 0.2334
4	3	~~0.2334~~ 0.2668
5	4	~~0.2668~~ 0.2001
6	5	~~0.2001~~ 0.1029
7	6	~~0.1029~~ 0.0367
8	7	~~0.0367~~ 0.0090
9	8	~~0.0090~~ 0.0014
10	9	~~0.0014~~ 0.0001
11	10	0.0000

Scanned with CamScanner



x, cumprob, "s")



Practical 2)

Normal Distribution.

i) $P(X=x) = dnorm(x, \mu, \sigma)$
ii) $P(X \le x) = Pnorm(x, \mu, \sigma)$
iii) $P(X > x) = 1 - Pnorm(x, \mu, \sigma)$

iv) To generate random number from a normal distribution (n random numbers) the R code is $rnorm(n, \mu, \sigma)$

1. A random variable X follows normal distribution with mean, $\mu = 12$ & S.D. = σ = 3 find i) $P(X \le 15)$ ii) $P(10 \le X \le 13)$ iii) $P(X > 14)$ iv) Generate 5 observation (random numbers)

```
> 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] 0.378066
> cat("P (10 <= x <= 13)=" P2)
P (10 <= x <= 13) = 0.378066
> P3 = 1 - Pnorm (14,12,3)
> P3
[1] 0.2524925
```

> cat("P(X>14):", P3)
P(X>14) : 0.2524925
> P4 = rnorm(5, 12, 3)
> P4
[1] 15.254723 16.545501 11.280515 6.419444
12.272460

2. X follows normal distribution with μ=10, σ=2 find i) P(X ≤ 7) ii) P(5 < X < 12) iii) P(X > 12) iv) Generate 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] 0.1586553
> a4 = rnorm(10,10,2)
> a4
[1] 11.608431 9.920417 12.637741 8.073254
8.721380 9.143726 9.366824 11.707.1
9.537584 10.715006
> a5 = qnorm(0.4, 10, 2)
> a5
[1] 9.493306

```
> sd = sqrt(v)
> sd
[1] 3.331613
> cat ("SD is = ", sd)
S.D = 3.331613
```

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

i) $P(x \le 40)$

ii) $P(x > 35)$

iii) $P(25 < x < 35)$

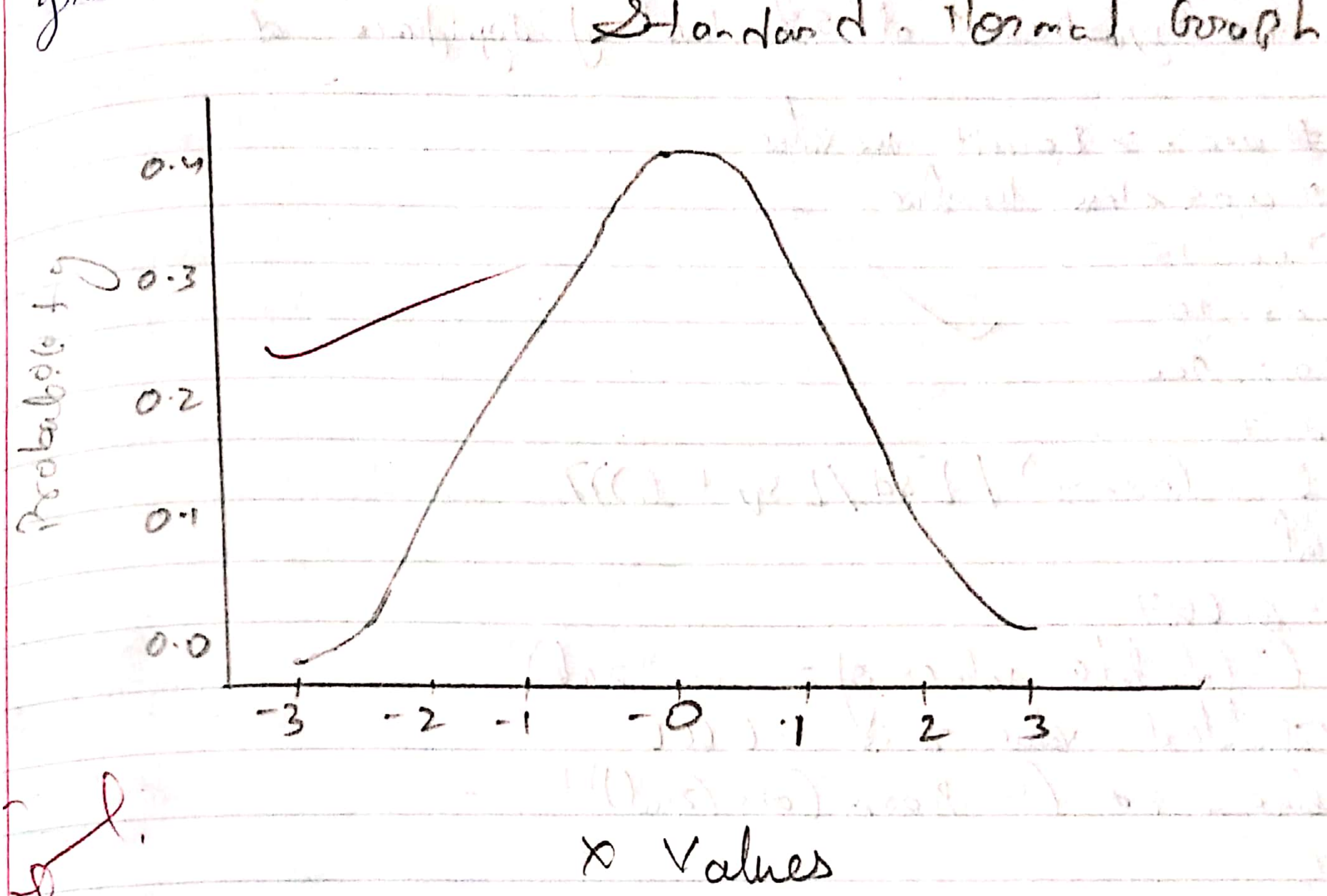
iv) find k such that $P(x < k) = 0.6$

```
> P1 = pnorm(40,30,10)
> P1
[1] 0.8413447
> P2 = 1 - pnorm(35,30,10)
> P2
[1] 0.3085375
> P3 = pnorm(25,30,10) - pnorm(35,30,10)
> P3
[1] -0.3829249
P4 = qnorm(0.6,30,10)
P4
] 32.53347
```

Scanned with CamScanner

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 5

Normal & Z. test.

i) $H_0: \mu = 15$ $H_1: \mu \neq 15$

Test the hypothesis
Random sample of size 400 is drawn and sd is
full calculated. The sample mean is 14 & S.D is 3
the hypothesis at 5% level of significance at

#0.05 > 1 accept the value
0.05 < less the value.

```
>mo = 15
>mx = 14
>n = 400
>sd = 3
>zcal = (mx-mo) / [sd / [sqrt (n))]
>zcal
[1] -6.6667
>cat (" calculate values of z is= "zcal)
    calculated value z is = -6.666
>pvalue = 2* (1- Pnorm (abs (zcal)))
>pvalue
[1] 2.616796
```

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

Test the hypothesis Ho: μ = 10 against H₁: μ ≠ 10
A random sample size of 400 is taken with sample mean = 10.2
& s.d = 2.25 Test the hypothesis of.

> mo = 10
> n = 400
> mx = 10.2
> sd = 2.25
> zcal = (mx - mo) / [sd / (sqrt(n)))
> zcal
[1] 1.7778
> pvalue = 2 * (1 - pnorm (abs (z(a)))
> pvalues
> 0.07544036
∴ The values Pvalue is greater than 0.05. The value is accepted

> Test the hypothesis Ho: proportional of smoker in college is 0 is collected & calculated the sample of as 0.125 test the at 5% level of significants

> 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 is = - 3.75
Pvalue = 2 * (1 - pnorm (abs (zcal)))
Pvalue
) 0.0001768346 (Reject)

4] last year farmen's lost 20% of their farm crops. A r... of 60 fields are collected 80 it is test hypothesis of 1% significance.

```
>P=0.2
>p=9/60
>n=60
>zcal=(p-P)/[sqrt(P*Q/n)]
>zcal
>[1]-0.96824
>Pvalue = 2*(1-pnorm(abs(zcal)))
Pvalue
>[1] 0.332916.
```

~~Z~~ Test

Test the hypothesis Ho : μ = 12.5 from the followed sample 5% level of significance.

```
> x = c(12.25, 11.97, 12.25, 12, 0.8, 12.31, 12.28, 11.94, 11.8
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
> sd = sqrt (variance)
> sd
> [1] 0.1397176
> t = (mx - mo) / (sd / sqrt (n))
t
[1] -8.894909
> pvalue = 2 * (1 - pnorm (abs (t)))
> pvalue
[1] 0
```

∴ This value is less than 0.05 the value accepted

Practical 6

① Let the population mean (the amount spent per customer restaurant) is 250. A sample of 100 customers selected sample mean is customer as 275 & sd = 30 test the that the population mean is 250 or not on 5% significant.

② In a random sample of 100 student it is found that 750 we value is blue test the hypothesis that population to proportion is 0.8 at 1% level of sign

-> ① > m0 = 250
> mx = 275
> sd = 30
> n = 100
zcal = (mx - m0) / (sd / (sqrt (n)))
> cat ("calculated value of is z:", zcal)
[1] calculated value of is z is = 8.333
> pvalue = 2 * (1 - pnorm (abs (zcal)))
pvalue
[1] 0
The value is less than 0.05 we will reject of H0: μ = 250

p=0.8
q=1-p
p=750/1000
n=1000
zcal=(p-P)/(sqrt(P*Q/n))
print("calculation value is z is:", zcal)
[1] calculated value of z: -3.95527
pvalue=2*(1-pnorm(abs(zcal)))
pvalue
[1] 7.77628-32
∴ The value is less 0.01 we reject it

3) A study of noise level in 2 hospital given below list 2 hospital have same level of noise at 1% level significance

Hos. A	Hos. B
84	34
61.2	59.4
7.9	7.5

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

28

>Pvalue = 2* [1- pnorm (abs (zcal)))
Pvalue
Pvalue
[1] 0.2450211
The value is greater than 0.01 we accept the value.

5) In a sample of 600 student in a college 400 used a blue ink. In another college from a sample of 900 student 450 use blue ink. Test the hypothesis that the proportion of students using blue ink in the two college to have equal or not at 1% level of significant

As $H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

>n1 = 600
>n2 = 900
>P1 = 400/600
>P2 = 450/900
>P = (n1 * P1 + n2 * P2) (n1+n2)
>P
[1] 0.4333
>Zcal = (P1-P2) /sqrt(p*q*(1/n1 + 1/n2))
>Zcal
[1] 6.381534
>Pvalue = 2* (1- Pnorm (abs (zcal)))
pvalue
[1] 1.7532 e-10
∴ Pvalue is less than 0.01 the value is rejected

Practical 7

Topic : Small Sample test

The marks of 10 students are given 63, 63, 66, 67, 68, 69, 71, 72 test the hypothesis that the sample comes from the population with average 66.

Ho : μ = 66
x = c(66, 63, 66, 67, 68, 69, 70, 70, 71)
t.test (x)
one sample
data : x
t = 68.319, df = 9, p-value = 1.558e-13
alternative hypothesis
True mean is not equal to 0
95 percent confidence interval
65.65171 70.14829
sample estimates
mean of x
67.9

The p-value is less than 0.05 we reject the hypoth
of significance.

2. Two groups of student scored 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 group.

```
> 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)
   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
confidence interval :
   0.1628205      5.0371795
 Sample estimates :
  mean of x  mean of y
     20.1        17.5
> p.value = 0.03798
> if (pvalue > 0.05) {cat ("accept ho") }
  else {cat ("reject ho") }
reject Ho.
```

The sales data of 36 shops before & after a special campaign are given below.
before : 53, 18, 31, 48, 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 & 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, p value = 0.9806
alternative hypothesis
True difference in means is greater than 0 95 percent confidence
-6.035547 inf
sample estimates
mean of the difference
-3.5

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

mean of the difference
(9.83333

P.value is greater than 0.05 we accept the hyp
at 5% level of significance.

2 medicine are applied to two group of period respectively

gp1: 10, 12, 13, 11, 14

gp2: 8, 9, 12, 14, 15, 10, 9

Is there any significance difference b/w 2 medicine.

H0: There is no significance difference.

> x = c(gp1)

> y = c(gp2)

> 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 10.3333

∴ Pvalue is greater than 0.05 we accept the H0 at 5% level of significance.

>zcal
[1] 0
> pvalue = 2 * (1-pnorm(abs(zcal)))
> pvalue
[1] 1
As pvalue is greater than 0.05 we acce
H_0 at 1% level of significance.

$H_0: \mu = 66$ against $H_1: \mu \neq 66$

>x = c(63, 63, 69, 77, 71, 72)

>t.test(x)

>[1]

data = x

t = 47.99 df = 6 p-value = 5.522e-09

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval

64.6647 71.67092

sample estimates:

mean of x

68.14286

Since pvalue is less than 0.05 we reject H_0 at 1% level

$H_0: \sigma_1 = \sigma_2$ 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, 87, 92, 93, 95, 97)

>var.test(x, y)

F test to compare two variance

data: x & y

F = 0.7888 num df = 7 , denom. df = 910

pvalue = 0.7737

alternative hypothesis: The ratio of ~~var~~ value is not

95 percent confidence interval

0.199509 3.75188

118

Soln: Sample estimate
ratio of variance
0.2880255

Pvalue is greater than 0.05 we accept H_0 at 5% significance.

b) $H_0: P_1 = P_2$ 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
> q
[1] 0.8
> Zcal = (P1-P2)/sqrt(P*q*(1/n1 + 1/n2))
> Zcal
[1] 0.3613104
```

∴ Pvalue is greater than 0.05 we accept ~~the~~ H_0 at 1% significance.

dew
(1.68, 3.35, 2.50, 6.23, 3.24)

Sign test
```
> x <- c(1.68, 3.35, 2.50, 6.23, 3.24)
> n <- length(x); p0 <- 0.5;
> x > 4
[1] FALSE FALSE FALSE TRUE FALSE
> s <- sum(x > 4); s;
x > 4
[1] 1
> binom.test(s, n, p= 0.5, alternative = "greater");
```

Exact binomial test

data: s and n

number of successes = 1, number of trials = 5, p
= 0.9688

alternative hypothesis: true probability of success
than 0.5 95% confidence interval:

0.01020622 1.00000

Sample estimates

Probability of success

~~wilcoxon test~~

wilcoxon test:

① The score of 8 student in reading before & after lesson are as follow. Test whether there is effect of reading.

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

```
> 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");
  data : D
  V = 10.5, P.Values = 0.8722
alternative hypothesis true location is greater
warning message:
In wilcox.test.default (D, alternative = "gr
Cannot compute exact p.value with ties.
```

Mann Whitney - Wilcoxon's test.

The diameter of a ball bearing was measured by 6 inspectors using two different kind of callipers. The result were test
ball bearing for

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

Given

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 & y

w = 24, p value = 0.197

alternative hypothesis: true location shift is gre

Practical 10

1. Use the following data to test whether the condition of home & condition of child are independent or not:

Cond Child	Cond Home clean	dirty
clean	70	50
Fairly clean	80	20
dirty	35	45

H0: condition of Home & child are independent

```
>x = c(70,80,35,50,20,45)
>m=3
>n=2
>y=matrix(x,nrow=m,ncol=n)
>y
      [,1]  [,2]
[1,]   70    50
[2,]   80    20
[3,]   35    45
>pv=chisq.test(y)
>pv
Pearson's chi-squared test
```

data: y
X-squared = 25.646
df = 2
p-value = 2.698e-06
They are dependent
∴ p value is less than 0.05 we reject the hypothesis at 5
level of significance.

2. Test the hypothesis that vaccination & disease are independent or not.

Vaccine	Affected	No affected
Disease		
Affect	70	46
No affected	35	37

Ho: Disease & Vaccine are independent.

```
> x = c(70, 35, 46, 37)
> m = 2
> n = 2
> y = matrix (x, nrow = m, ncol = n)
> y
      [,1]  [,2]
[1,]   70    46
[2,]   35    37
> pv = chisq.test (y)
> pv
```

Pearson's chi-squared test with yates ~~continuty~~ continui

data : y
X-squared = 2.0273
df = 1
p-value = 0.1545

∴ p value is more than 0.05 we accept the hypothesis at 5% level of significance.

3) Perform a ANOVA for the following data.

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

H0: The mean's are equal for A, B, C, D.

```
>x1=c(50,52)
>x2=(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))
>names(d)
[1] "values" "ind"
>oneway.test(values~ind,data=d,var.equal=T)
        one-way analysis of means
data: value and ind
F= 11.735 df = 3, denom df = 9,
pvalue = 0.00183
```

P-value is less than 0.05 we reject the hypothesis.

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

	Df	Sum	mean sq	F value	Pr(>F)
ind	3	71.06	23.688	11.73	0.00183
Residuals	9	18.17	2.019		

Signif codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1