PRACTICAL - 1

Basics of R software.

1). R is a software for statistical analysis and data computing.

2). It is an effective data handling software and out-come storage is possible.

3). It is capable of graphical display.

4). It is a free software.

Solve the followings

1). $4+6+8 \div 2-5$
```
> 4+6+8/2-5
[1] 9
```

2). $2^2+|-3|+\sqrt{45}$
```
> 2^2+abs(-3)+sqrt(45)
[1] 13.7082
```

3). $5^3+7\times5\times8+46/5$
```
> 5^3+7*5*8+46/5
[1] 414.2
```

4). $\sqrt{4^2+5\times3+7/6}$
```
(4^2+5*3+7/6)
```

Q5). round off

$46 \div 7 + 9 \times 8$

```
> round (46/7+9*8)
[1] 79.
```

Q2.

```
> 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 9162185
```

Q3.

```
> x=20   >y=30   >z=2
> x^2+y^3+z
[1] 27402
> sqrt(x^2+y)
[1] 20.73644
> x^2+y^2
[1] 1300
```

Q4.

```
> 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
```

Scanned with CamScanner

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, ncol=3, data=c(4,7,9,-2,
                                          0,-5,
                                          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

> 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
```

Scanned with CamScanner

Marks of statistics of C.S Batch A. 58

$x = c($ 59, 20, 35, 24, 46, 56, 55, 45, 27, 22, 47, 58, 54, 40, 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

Check whether the following is pmf or not:-

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

Since $P(2) = -0.5$, it can't be probability pms. function $P(x)$ is greater than equal to zero.

It cannot be pmf as pmf $\Sigma P(x) = 1$

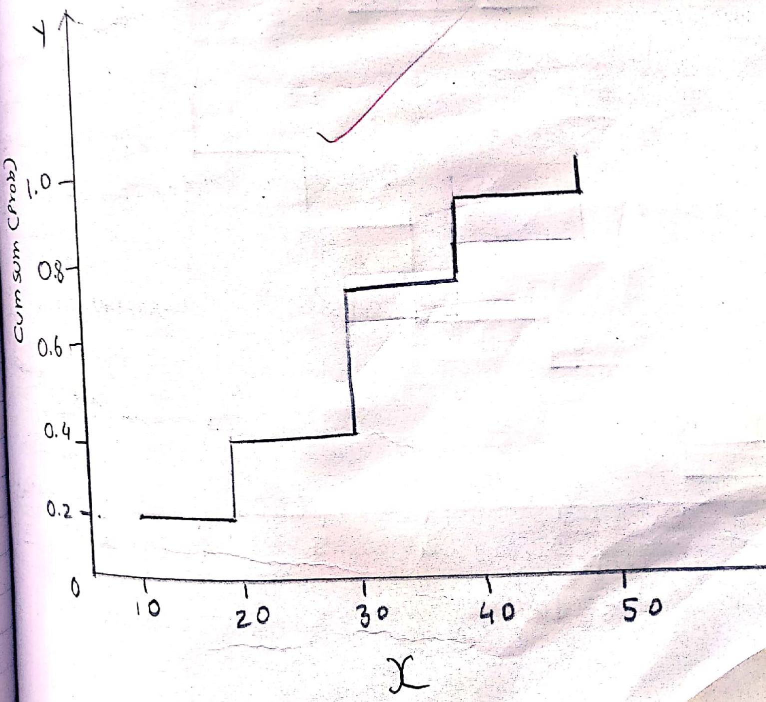
1). Find C.d.f for the following P.M.F and Sketch the graph.

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	$10 \le x < 20$	
0.4	$20 \le x < 30$	
0.75	$30 \le x < 40$	
0.950	$40 \le x < 50$	
1.00	~~$50 \le x < 60$~~	$x \ge 50$

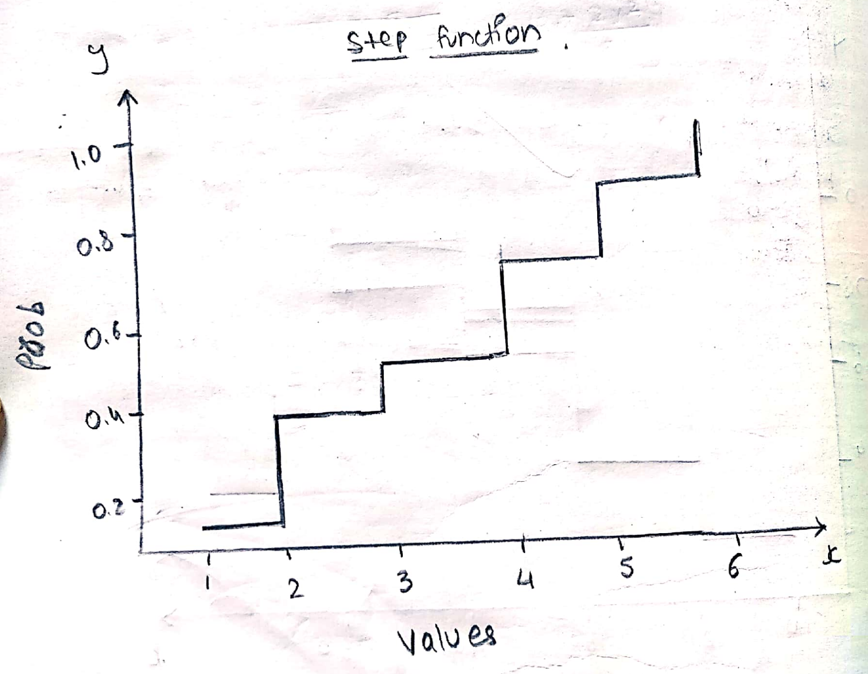
CODE :-

```
> prob = c(0.2, 0.2, 0.35, 0.15, 0.1)
> sum(prob)
[1] 1.0
> cumsum(prob)
[1] 0.2 0.4 0.75 0.950 1.00
> x = c(10, 20, 30, 40, 50)
> plot(x, cumsum(prob), "s")
```



CODE:-

```
> x = c(1, 2, 3, 4, 5, 6)
> prob = c(0.15, 0.25, 0.1, 0.2, 0.2, 0.1)
> cumsum (prob)
> plot (x, cumsum (prob), "s", main="step function",
        xlab = "values", ylab ="prob", col="green")
```



x.P(x)

x	Prob
1	0.15
2	0.25
3	0.1
4	0.2
5	0.2
6	0.1

53

CODE :-

Q1].
```
> dbinom (10, 100, 0.1)
[1] 0.1318653.
```

Q2]. 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.
```

Q3].
```
> dbinom (0:5, 5, 0.1)
[1]. 0.59049  0.32805  0.07290  0.00810  0.00045
     0.00001
```

Q4]. 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] 0.00278151
```

iv).
```
> dbinom (6, 12, 0.25)
[1] 0.04014945
```

(Practical -3).

Formula :- (command).

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

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

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

(iv). If x is unknown and $p_1 = P(X \leq x) =$ qbin

Q1]. Find the probability of exactly 10 success in trials with $p = 0.1$

Q2]. Suppose there are 12 MCQ. ~~sentence~~ Each question has 5 options out of which correct. Find the probability of having :-

i). Exactly 4 correct answers $P(X=4)$.

ii). At-most 4 correct answers $P(X \leq 4)$

(iii). More than 5 correct answers. P

Q3]. Find the complete distribution when n=

Q4]. $n=12$, $p=0.25$. Find the following proba

i). $P(X=5)$

ii). $P(X \leq 5)$

iii). $P(X>7)$

iv). $P(5<x<7)$

Q5]. Probability of a salesman making sa is 0.15. Find the probability of

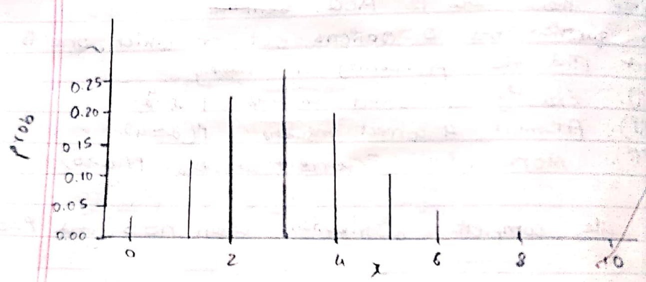
i). No sales out of 10 customer

Q6] A Salesman has a 20% probability of making a sale to a customer, out of 30 customers what minimum no. of sales he can make with 88% probability.

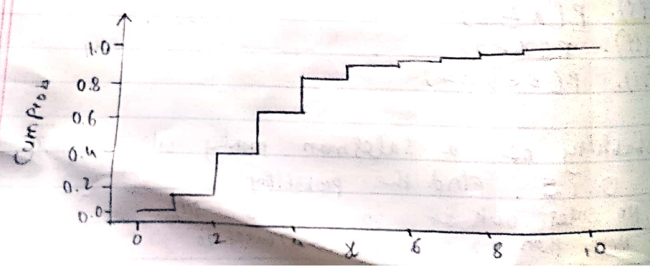
Q7]. X follows binomial distribution with n=10, p=0.3, Plot the graph of pmf. and c.d.f

GRAPH.

P.M.f.



c.d.f.



Q5). i). >dbinom (0, 10, 0.15)

[1] 0.1968744.

ii). >1- pbinom (3, 20, 0.15)

[1] 0.3522748..

Q6]. >qbinom (0.88, 30, 0.2)

>[1] 9.

Q7]. >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	0.0282475249
2	1	0.1210608210
3	2	0.2334744405
4	3	0.2668279320
5	4	0.2001209490
6	5	0.1029193452
7	6	0.0367569090
8	7	0.0090016920
9	8	0.001446[illegible]
10	9	0.0001[illegible]
11	10	0.[illegible]

(Q).
```
> 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] 10.299518   13.439425   11.550703   8.767144
    12.190495
> cat("P(X=5) =", P1)
> cat("P(10 ≤ X ≤ 13) =", P2)
> cat("P(X>14) =", P3)
```

PM

PRACTICAL-4

* commands :-

1) $P(X = x) = \text{dnorm}(x, \mu, \sigma)$
2) $P(X \leq x) = \text{pnorm}(x, \mu, \sigma)$
3) $P(X > x) = 1 - \text{pnorm}(x, \mu, \sigma)$

~~To gen~~

4) To generate random numbers from a distributio (in random numbers). The R code is

rnorm(n, μ, σ)

(Q). A random variable X follows normal distri with mean = 12 = μ and sd = σ = 3
Find:
1). $P(X \leq 15)$
2). $P(10 \leq X \leq 13)$
3). $P(X > 14)$
4). generate 5 observations (random numb

(Q). X follows normal distribution with $\mu = 10$, $\sigma =$
find
i). $P(X \leq 7)$
ii). $P(5 < X < 12)$
iii). $P(X > 12)$
iv). Generate 10 observations.
v). Find k such that probability $P(X < k) = 0.4$

29

(Q3). Generate 5 random nos from a N.D with ~~mean~~ $= 15$ (μ) and s.d $= 4$ (σ). find sample mean, median, s.d, and print it.

(Q4). x follows :-
$X \sim N(30, 100)$
$\therefore \mu = 30$, $\sigma = 10$.

1). $P(X \leq 40)$
2). $P(X > 35)$
3). $P(25 < x < 35)$
4). Find x such that $P(x < k) = 0.6$

(Q5). Plot the standard normal graph.

(Q).
```
> p1 = pnorm (7,10,2)
> p1
[1] 0.0668072
> p2 = pnorm (12,10,2) - pnorm (5,10,2)
> p2
[1] 0.8351351
> p3 = pnorm (12,10,2)
> p3
[1] 0.1586553
> p4 = rnorm (10,10,2)
> p4
[1] 9.592804  [2] 10.139277  12.28  11.15417
    9.066229  10.757278  11.991077  9.742067
    10.300246  9.346953
> p5 = qnorm (0.4,10,2)
> p5
[1] 9.493306
```

(Q).
```
> x = rnorm (5,15,4)
> am = mean (x)
[1] 17.52547
> median = median(x)
[1] 17.02539
> variance n = 5
> varience = (n-1)*(var(x)/n)
> varience
[1]  2.916316
> sd = sqrt (varience)
[1]  5.567117
```

PRACTICAL – 5.

*TOPIC :- NORMAL & t-Test

Q1]. Test the hypothesis $H_0 : \mu = 15$
$H_1 : \mu \neq 15$.

- A Random sample of size 400 is drawn and it is calculated as sample Mean = 14 and s.d = 3. Test the Hypothesis at 5% level of significance.

Formula :- $Z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

In R software :- mo=15 , mx=14 , sd= 3 , n=400
zcal = (mx-mo)/(sd/(sqrt(n)))

Pvalue=(1 - pnorm (~~...~~ abs (zcal)))*2.

⇒ ∴ since Pvalue is ~~more~~ less than 0.05, we aren't accepting $H_0 : \mu = 15$//.

```
1].
> mo=15
> mx=14
> sd=3
> n=100
> zcal = (mx-mo)/(sd/(sqrt(n)))
> zcal
[1] -6.666667.
> cat (" Calculated value of z is =", zcal)
Calculated value of z is = -6.666667.
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 2.616796e-11
```

2]. CODE :

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


Q2]. Test the Hypothesis $H_0: \mu = 10$
$H_1: \mu \neq 10$

A random sample of size 400 is drawn with sample mean 10.2 and s.d = 2.25. Test the Hypothesis at 5% level of significance.

⇒ ∴ since Pvalue is more than 0.05, we accept $H_0: \mu = 10$

[Q3]. Test the Hypothes H_0: ~~the~~ Proportion of smokers in our college is 0.2. The sample is Collected and sample proportion = 0.125. Test the Hypothesis at 5% level of significance (sample size = 400).

⇒ ∴ Since the value is Less than 0.05, we reject H_0: proportion = 0.2.

CODE :-

```
> n=400
> P=0.2
> p=0.125
> Q=1-P
> zcal=(p-P) / (sqrt(P*Q/n))
> zcal
[1] -3.75
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.0001768346
```

[Q4]. Last year → the farmers lost 20% of their crop. A random sample of 60 fields are called and it is found that 9 fields crops are insect polluted. Test the Hypothesis at 1% level of significance.

H_0: $P = 0.2$

H_1: $p = 9/60$

⇒ ∴ Since pvalue is greater than 0.01, we ~~reject~~ accept the pvalue $H_0 : P = 0.2$//.

CODE :-

```
> n=60
> P=0.2
> p=0.15
> Q=1-p
> zcal= (p-P)/(sqrt(P*Q/n))
> pvalue= 2*(1-pnorm(abs(zcal)))
> zcal
[1] -0.9682458.
> pvalue
[1] 0.3329216.
```

[Q3]. Test the Hypothes Ho: ~~the~~
Proportion of smokers in our college is 0.2
The sample is collected and sample proportion = 0.125
Test the Hypothesis at 5% level of significance
(sample size = 400).

⇒ ∴ Since the value is Less than 0.05, we reject
Ho: proportion = 0.2.

CODE :-

```
>n= 400
> P=0.2
> p=0.125
> Q=1-P
> zcal=(p-P) / (sqrt(P* Q/n))
> zcal
[1] -3.75
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.0001768346
```

[Q4]. Last year → the farmers lost 20% of their crop.
A 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 significance.
Ho: P=0.2
H1: $p = 9/60$

⇒ ∴ Since Pvalue is greater than 0.01, we ~~reje~~
accept the Pvalue Ho: P= 0.2//

CODE :-

```
>n=60
> P=0.2
> p=0.15
> Q=1-p
> zcal= (p-P) / (sqrt(P* Q/n))
> Pvalue= 2* (1-pnorm (abs (zcal)))
> zcal
[1] -0.9682458.
> Pvalue
[1] 0.3329216.
```

PRACTICAL-06

Aim: large sample test

[Q1]. Let the population mean [the amount spent by customer in a restaurant] in 250, A sample of 100 customers selected. the sample mean is calculated as 275 and S.D is 30, Test the hypothesis that population mean is 250 or not. The level of significance is 5%.

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

∴ since the pvalue is less than 0.05, we reject the pvalue $H_0: \mu_1 = \mu_2$

[Q2]. In a random sample of 1000 students it is found that 750 use blue pen. Test the hypothesis that population proportion is 0.8 at 1% level of significance.

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

since the pvalue (7.72268e-05) is ~~we re~~ less than 0.01, we reject the pvalue $H_0: \mu_1 = \mu_2$

CODE :-

```
> mx = 275
> sd = 30
> n = 100
> mo = 250
> zcal = (mx-mo)/(sd/(sqrt(n)))
> zcal
[1] 8.333333.
> pvalue = (1 - pnorm(abs(zcal)))*2
> pvalue
[1] 0.
> cat ("zcal is :", zcal)
zcal is : 8.33333
> cat (" pvalue is : ", pvalue)
pvalue is : 0.
```

2). CODE :-

```
> n=1000
> P=0.8
> Q=1-P
> p=750/100
> zcal = (p-P)/(sqrt(P*Q/n))
> zcal
[1] -3.952847
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 7.72268e-05
> cat (" zcal is :"
zcal is
```

55

CODE :-

```
> 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 = (1-pnorm(abs(zcal)))*2
> pvalue
[1] 2.417564e-07
> cat ("zcal is:", zcal)
zcal is: -5.163978
> cat ("pvalue is:", pvalue)
pvalue is: 2.417564e-07
```


(3). Two random samples of size 1000 and 2000 are drawn from two population with the same S.D i.e. 2.5. The sample means are 67.5 and 68 respectively. Test the hypothesis $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ at 5% level of significance.

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$.

$n_1 = 1000$

$n_2 = 2000$

$mx_1 = 67.5$

$mx_2 = 68$

$sd1 = 2.5$

$sd2 = 2.5$

$\therefore$ since pvalue is less than 0.05, we reject the pvalue $H_0: \mu_1 = \mu_2$

(Q4). A study of noise level in 2 Hospitals is given below. Test the claim that the 2 hospitals have same level of noise at 1% level of significance.

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

	Hospital A	Hospital B
Size	84	34
Mean	61.2	59.4
S.D	7.9	7.5

∴ Since the pvalue is greater than 0.01, we accept the pvalue $H_0: \mu_1 = \mu_2$.

(Q5). In a sample of 600 students in a college, 400 use blue ink in another college from a sample of 900 students, 450 use blue ink. Test the hypothesis that the proportion of students using blue ink in 2 colleges are equal or not at 1% level of significance.

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$.

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

Since pvalue is less than 0.01, we reject the pvalue $H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

```
>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=(1-pnorm(abs(zcal)))*2
>pvalue
[1] 0.2450211
>cat("zcal is:", zcal)
zcal is: 1.162528
>cat("Pvalue is:", pvalue)
Pvalue is: 0.2450211
```

CODE :-

```
>n1=600
>n2=900
>P1=400/600
>P2=450/900
>P=(n1*P1+n2*P2)/(n1+n2)
>q=1-P
>zcal=(P1-P2)/sqrt(P*q*(1/n1+1/n2))
>pvalue=(1-pnorm(abs(zcal)))*2
>zcal
[1] 6.381534
>pvalue
[1] 1.753222e-10
>cat("zcal is:", zcal)
zcal is: 6.381534
>cat("Pvalue is:", pvalue)
Pvalue is:
```

```
> p
[1] 0.5666667
> cat(" p is:", p)
p is : 0.5666667.
```

CODE :-

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

$n_1 = 200$
$n_2 = 200$
$p_1 = 44/200$
$p_2 = \frac{30}{200}$

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

∴ since Pvalue is ~~less~~ greater than 0.05, we ~~reject~~ accept the Pvalue :
$H_0: P_1 = P_2$.

PM
27-01-20

PRACTICAL-7.

Aim: Small sample space.

1. The marks of 10 students are given by 63, 63, 66, 67, 68, 69, 70, 70, 71, 72. Test the hypothesis that the sample comes from the population with the average 66.

$H_0: \mu = 66$

Ans:- Since p-value is less than 0.05, we reject the hypothesis $H_0: \mu = 66$ at 5% level of significance.

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

Group 1 – 18, 22, 21, 17, 20, 17, 23, 20, 22, 21.

Group 2 – 16, 20, 14, 21, 20, 18, 13, 15, 17, 21.

H_0: there is no difference between the two groups.

Ans:- Since p-value is less than 0.05, we reject the hypothesis H_0: no difference at 5% level of significance

```
> x=c(63,63,66,67,68,69,70,70,71,72)
> t.test(x)
```

One sample t-test

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

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

reject Ho.

CODE :-

```
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 = 0.03798
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
0.1628205 5.0371795
sample estimates:
mean of x mean of y
20.1 17.5

Ans:-

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

4. Two medicines are applied to groups of patient respectively

Group 1 – 10, 12, 13, 11, 14.
Group 2 – 8, 9, 12, 14, 15, 10, 9.
Is there any significant difference between the two medicines?

Ans → Since the pvalue is greater than 0.05, we accept the hypothesis.

Ho: no significant difference between two medicines.

5. The following are the weights 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.

Ans → Since the pvalue is greater than 0.05, we accept the hypothesis.

Ho: no significant difference.

Ho: not effective

```
> x=c(10, 12, 13, 11, 14)
> y=c( 8, 9, 12, 14, 15, 10, 9).
> t.test (x, y)
        Welch Two sample t-test
data:  x  and  y
t= 0.80384,  df = 9.7594,  p-value =0.4406.
~~alternative hypothesis: true difference in mean is not~~
95 percent confidence interval:
    ~~-Inf 29.0295~~    -1.781171   3.781171
~~sample estimates:~~
mean of x      mean of y   } sample estimates.
  12              11
> if (pvalue > 0.05) { cat("accept Ho") }else{ cat("reject
      accept Ho.
```

5. Code:–

```
> b=c(120, 125, 115, 130, 123, 119)
> a=c(100, 114, 95, 90, 115, 99)
> t.test (b, a, paired=T, alternative="less")
        Paired t-test
data:  b and a
t= 4.3458, df =5, p value =0.9963
alternative hypothesis: true difference in means is less
95 percent confidence interval:
      - Inf    29.0295
sample estimates:
mean of the differences
         19.83333
> if (pvalue > 0.05) { cat ("accept Ho") } else { cat (" reject
accept Ho.
```

6.7.10

78

1]. CODE :-

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

2]. CODE :-

```
> n=700
> p=350/700
> P= 0.5
> Q=1-P
> zcal=(p-P)/(sqrt(P*Q/n))
> zcal
[1] 0 .
> Pvalue= 2*(1-pnorm(abs(zcal)))
> Pvalue
[1] 1.
```


PRACTICAL-8

Aim :- Large Sample Test and Small Sample Test .

(i) **Ans** -> we reject the Pvalue at $H_0 : \mu = 55$.
~~$H_1 : \mu \neq 55$~~.

(ii). **Ans** -> we accept the pvalue at $H_0 : P = 1/2$ at 1% level of significance.

[Q3]. ~~Ans → We accept reject the the Pvalue~~

[Q3]. Ans → We reject the Pvalue at $H_0: \mu = 0.05$ at 5% level of significance.

[Q4]. Ans → We reject the Pvalue at $H_0: \mu = 100$ at 5% level of significance

CODE :-

```
> n1 = 1000
> n2 = 1500
> p1 = 0.02
> p2 = 0.01
> p = (n1 * p1 + n2 * p2) / (n1 + n2)
> q = 1 - p
> zcal = (p1 - p2) / sqrt (p * q * (1/n1 + 1/n2))
> zcal
[1] 2.084842
> pvalue = (1 - pnorm (abs (zcal))) * 2
> pvalue
[1] 0.03708364
> p
[1] 0.014
```

CODE :-

```
> mx = 99
> mo = 100
> sd = 8
> n = 400
> zcal = (mx - mo) / (sd / (sqrt(n)))
> pvalue = (1 - pnorm (abs(zcal))) * 2
> zcal
[1] -2.5
> pvalue
[1] 0.01241933
```

CODE :-

```
> 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 ~~of~~ confidence interval:
  64.66479   71.62092
Sample estimates:
mean of x
68.14286
```

CODE :-

```
> x= c(66,67,75,76,82,84,88,90,92)
> y=c(69,66,74,78,82,85,87,92,93,95,97)
> var.test (x,y)

        F test to compare two variences

data:  x and y
F = 0.70686 , num df = 8 , denom df = 10, p-value = 0.6359
alternative hypothesis: true ratio of variences is not equal to 1
95 percent confidence interval:
  0.1833662  3.0360393
Sample estimates:
ratio of variences
    0.7068567.
```

Ans → we reject the p-value at ·1% Level of significance.

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

Ans → we accept the p-value at 5% level of significance.

$H_0: 61 = 62$ against $H_0: 61 \neq 62$

(Q7). $H_0: \mu_1 = 1200$ against $H_1: \mu_1 \neq 1200$

Ans →. We reject the P-value at $H_0: \mu_1 = 1200$

(Q8). Ans → We accept the Pvalue at 1% level of significance

$H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

... 1150

... 1200

... 12.5

zcal = (mx - mo)/ (sd /(sqrt (n)))

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

zcal

-4

Pvalue

6.334268e-05

n1 = 200

n2 = 300

p1 = 244/200

p2 = 56/300

p = (n1 * p1 + n2 * p2) /(n1+n2)

q = 1-p

zcal = (p1-p2)/sqrt(p * q * (1/n1+1/n2))

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

zcal

0.9128709

Pvalue

0.3613106

1. code:

```
> 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, pvalue = 2.698e-06
```

PRACTICAL - 9

Chi square Test and ANOVA

Q. Use the following data to find whether the condition of the Home and the condition of the child are independent or not.

		Condition of Home	
		Clean	Dirty
Condition of child	Clean		
	Fairly clean		
	Dirty		

H0: Condition of home and child are independent

Since Pvalue is less than 0.05, we reject at 5% level of significance.

2
(Q). Test the hypothesis that the vaccination and the disease are independent or not.

		Vaccine	
		Affective	Not Affective
Disease	Aff.	70	46
	N. Aff.	35	37

Ans ⇒ Disease and the vaccine are independent.

⇒. Since p-value is greater than 0.05, we accept Ho at 5% level of significance

CODE :-

```
> 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' continuity correction

data: y

X-squared = 2.0275, df = 1, p-value = 0.1545

3). CODE :-

```
>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))
>names(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, denom df = 9, p-value =0.00183.
> anova = aov (values ~ ind, data =d)
> 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
---
signif. codes:   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Perform an ANOVA on the following data.

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.

Since the pvalue is less than 0.05 we reject Ho at 5% level of significance.

Scanned with CamScanner

4.

(Q). The following data gives the life of the tyre of 4 brands

Type	Life
A	20, 23, 18, 17, 18, 22, 24
B	19, 15, 17, 20, 16, 17
C	21, 19 22, 17, 20
D	15, 14, 16, 18, 14, 16

Ans → . Ho: the avg. life of A, B, C, D are equal

⇒ . Since the p-value is less than 0.05, we reject Ho at 5% level of significance

5.

(Q).

```
x1 = c(20, ... 18, 22, 24)
x2 = c(19, 15, 17, 20, 16, 17)
x3 = c(21, 19, 22, 17, 20)
x4 = c(15, 14, 16, 18, 14, 16)
d = stack(list(b1=x1, b2=x2, b3=x3, b4=x4))
names(d)
"values" "ind"
oneway.test(values ~ ind, data = d, var.equal = T)
    one-way analysis of means
data: values and ind
F: 6.8445, num df = 3, denom df = 20, p-value = ...
anova = aov(values ~ ind, data = d)
summary(anova)
```

5) code:-

```
> x=read.csv("C:/Users/Administrator/Desktop/maths.csv")
> print(x)
>am=mean(x$statistics)
>am
[1] 37
> n=length(x$statistics)
>n
[1] 10
> sd=sqrt((n-1)*var(x$statistics)/n)
> sd
[1] 12.64911
> am=mean(x$maths)
>am
[1] 39.4
> n=length(x$maths)
>n
[1] 10
> sd =sqrt((n-1)* var(x$maths)/n)
> sd
[1] 15.2
> cor(x$statistics, x$maths)
[1] 0.830618.
```

5 (10). To import excel file in R software.

	statistics	maths
1	40	60
2	45	48
3	42	47
4	15	20
5	37	25
6	36	27
7	49	57
8	59	58
9	20	25
10	27	27

AY
20/2/2

58

PRACTICAL-10

Aim: Non-Parametric test.

Q1. Following are the amounts of sulphur oxide emitted by some industry in 20 days. Applied sign test to test the hypothesis that population median is 21.5 at 5% LOS.

x = 17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26.

→ H0: population median is 21.5.
H_1: Less than 21.5

⇒ Since the P-value is greater than 0.05, we accept ~~the~~ H0 at 5% LOS.

* NOTE:- If the alternative (H_1) is me ≠ z or me < z the the used formula

else, if me > z
pv = pbinom (sn, n, 0.5).

CODE:-

```
> x = c(....)
> me = 21.5
> sp = length(x[x > me])
> sn = length(x[x < me])
> n = sp + sn
> pv = pbinom(sp, n, 0.5)
> pv
[1] 0.4119015
> n
[1] 20.
```

28 CODE :-

```
> x = c(...)
> me = 625
> sp = length(x[x>me])
> sn = length(x[x<me])
> n = sp+sn
> pv = pbinom(sn,n,0.5)
> n
[1] 10
> pv
[1] 0.0546875
```

[Q2]. ~~Write the~~ following is the data ~~of~~ 10 apply sign test to test the hypothesis population median is 625 against the it is more than 625.

$x =$ 612, 619, 631, 628, 643, 640, 655, 6
663.

⇒ Ho: Population mean is 625
H₁: greater than 625

⇒ Since the pvalue is greater than 0.05, ...
Ho at 5% LOS.

tha R CODE :-

```
> x = c( . . . . )
> wilcox.test (x, alter ="less" , mu=12)

data:  x
V=66 ,  p-value = 0.9986
alternative hypothesis : true location is less than 12
```


[Q4]. Using W.S.R.T. Test the population median is of less than 12.

x= 15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 2

→ Ho: population median is 12.
H1: greater than 12

→ since p-value is greater than 0.05, we a
Ho at 5% ~~de~~ ~~LOS~~ LOS.

Q5) The weight of a students before and after they stopped smoking are given below using W.S.R.T. Test that there is no significant change.

Weights before	After
65	72
75	74
75	72
62	66
72	73

→ H_0: before and after there is no change.
H_1: there is change.

⇒ Since the p-value is greater than 0.05, we accept H_0 at 5% ~~level of~~ LOS.

#5. CODE :-

```
> x=c(65,75,75,62,72)
> y=c(72,74,72,66,73)
> d = x-y
> wilcox.test (d, alter ="two.sided", mu=0)
```

data : d
V= 4.5, p-value = 0.4982
alternative hypothesis: true location is not equal to 0

A/1
3.7.21

Q3). The following are the values of a sample. test the hypothesis that the population median is 60 against the alternative it is more than 60 at 5%. LOS using Wilcoxon signed Rank Test.

x = 63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 69, 48, 66, 72, 63, 87, 69.

→ H_0: population me = 60.
H_1: greater than 60

⇒ Since p-value is ~~gre~~ less than 0.05, we reject H_0 at 5% LOS.

* NOTE :- If alternative is less alter = "less".
If alternative is not equal to, then alter = " two sided".

#3 ...

```
> x = c(...)
> wilcox.test (x, alter = "greater", mu = 60)
```

Wilcoxon signed rank test with continuity

data : x

V = 145, p-value = 0.02298.

alternative hypothesis : true location is greater tha