PRACTICAL - I

Basic of R Software.

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

2. It is any ~~objective~~ effective data handling ~~stor~~ software and outcome storage is possible.
3. It is capable of graphical display.
4. It is a free software.

Q1. Solve the following:

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

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

Q2.

1. c(2,3,5,7) * 2
> x = c(2,3,5,7) * 2
> x
~~[1] 4 6 10~~
[1] 4 6 10 14

2. c(2357) * c(2,3)
> [1] 4 9 10 21

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

4. c(2,3,5,7) ^ 2
[1] 4 6 10 14

5. c(4,6,8,9,4,5) ^ c(1,2,3)
[1] 4 12 24 9 8 15

Q3. Problem:
x = 20
y = 30
z = 2

Find:
i) $x^2 + y^3 + z$
ii) $\sqrt{x^2 + y}$
iii) $x^2 + y^2$

> x^2 + y^3 + z
[1] 27402

> sqrt(x^2+y)
[1] 20.73644

> x^2 + y^2
[1] 1300

Q4) Convert into a matrix form
> 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

Q5.) Find x+y and 2x+3y where.

$$x = \begin{bmatrix} 4 & -2 & 6 \\ 7 & 0 & 7 \\ 9 & -5 & 3 \end{bmatrix}, \quad y = \begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 5 \end{bmatrix}$$

```
> 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,-6,7,9,5))
> y
        [,1]  [,2]  [,3]
[1,]    10    -5     7
[2,]    12    -4     9
[3,]    15    -6     5

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


Q5.)

Marks: 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

Am 5.12.19

PRACTICAL-2

i) Check whether the followings 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 the given data is pmf then
$\Sigma P(x) = 1$ and $P(x_i) \geq 0$

Hence,
$\Sigma P(x) = P(1) + P(2) + P(0) + P(3) + P(4) + P(5)$
$= 0.1 + 0.2 + (-0.5) + 0.4 + 0.3 + 0.5$
$\Sigma P(x) = 1$

$\because$ The $\Sigma P(x)$ is equal to 1. but $P(2) = -0.5$
Hence it is not a p.m.f

Source code:

```
prob= c(0.1,
```

ii.)

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

The condition for pmf is $\Sigma P(x) = 1$ and $P(x) \geq 0$
so,
$\Sigma P(x) = P(1) + P(2) + P(3) + P(4) + P(5)$
$= 0.2 + 0.2 + 0.3 + 0.2 + 0.2$
$\Sigma P(x) = 1.1$

$\because$ The $\Sigma P(x) = 1.1$.
$\therefore$ The given data is not a pmf because the $\Sigma P(x) \neq 1$

* Source code:-

```
> prob=c(0.2,0.2,0.3,0.2,0.2)
> sum(prob)
[1] 1.1
```

iii.)

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

The condition for pmf is $\Sigma P(x) = 1$ and $P(x) \geq 0$
so,
$\Sigma P(x) = P(10) + P(20) + P(30) + P(40) + P(50)$
$= 0.2 + 0.2 + 0.35 + 0.15 + 0.1$
$\Sigma P(x) = 1$

$\because$ The $\Sigma P(x) = 1$ Hence it is a pmf.

* Source code

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

iii) x 32
$P(x)$

Q2)
i) Find the c.d.f for the following pmf.
and sketch the graph.

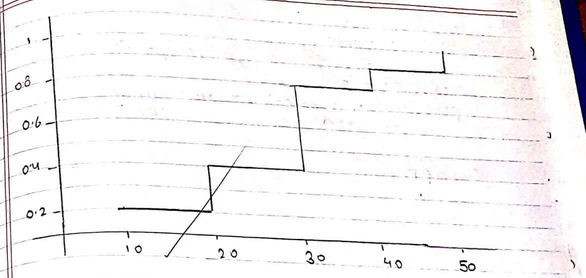
x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1

Sol^n:-
The given data is pmf.
∴ To find c.d.f from pmf we will use the

$F(x) = 0$ $x < 10$
$= 0.2$ $10 \leq x < 20$
$= 0.4$ $20 \leq x < 30$
$= 0.75$ $30 \leq x < 40$
$= 0.90$ $40 \leq x < 50$
$= 1.0$ $x \geq 50$

x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1
cdf	0.2	0.4	0.75	0.95	1.0

source code:-
```
>prob=c(0.2,0.2,0.35,0.15,0.1)
>cumsum(prob)
 [1]   0.20   0.40   0.75   0.90   1.0
>x=c(10,20,30,40,50)
>plot(x,cumsum(prob),"s")
```

iii) Find the c.d.f for the following pmf and sketch the graph.

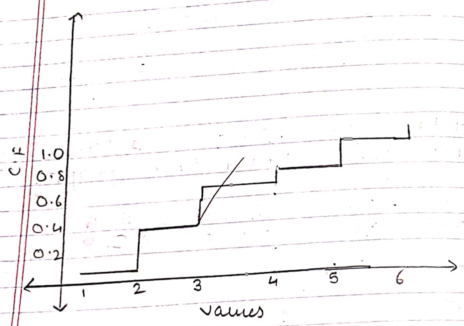
x	1	2	3	4	5	6
$P(x)$	0.15	0.25	0.1	0.2	0.2	0.1

Sol^n:-
$F(x) = 0$ $x < 1$
$= 0.15$ $1 \leq x < 2$
$= 0.40$ $2 \leq x < 3$
$= 0.50$ $3 \leq x < 4$
$= \cancel{0.70}$ 0.70 $4 \leq x < 5$
$= 0.90$ $5 \leq x < 6$
$= 1.00$ $x \geq 6$

Source code:-

```
> x= c (0.15 ,0.25 ,0.1 ,0.2, 0.2, 0.1)
> y= c (1, 2, 3, 4, 5, 6 )
> plot( y, cumsum(x) , xlab="values", ylab="C.F",
  col= "red" , "s" , main= "CDF-GRAPH" )
```

Graph:- CDF-GRAPH



Q3) Check whether the following is pdf or not

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

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

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

$\int_0^1 f(x)$

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

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

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

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

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

$\int_0^1 f(x)$

$= \int_0^1 3x^2$

$= 3\int_0^1 x^2$

$\left[3\frac{x^3}{3}\right]_0^1 \qquad \left[\because x^n = \frac{x^{n+1}}{n+1}\right]$

$= x^3$

$= 1$

$\therefore \because$ The $\int_0^1 f(x) = 1 \quad \therefore$ It is a pdf

Topic :- To study Binomail Distribution

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

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

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

iv) If x is unknown

$P_1 = P(X \le x)$ . . . (given)

qbinom(P_1, n, p)

Q1] Find the probability of exactly 10 success in 100 trials with $p = 0.1$

> dbinom(10, 100, 0.1)

[1] 0.1318653

Q2] Suppose there are 12 MCQ. Each question has 5 options out of which 1 is correct. Find probability of :-

i) exactly 4 correct answers.

ii) Almost 4 correct answers.

iii) More than 5 correct answers.

i)

i) > dbinom(4, 12, 1/5)

[1] 0.1328756

ii) > pbinom(4, 12, 1/5)

[1] 0.927445

iii) > 1 - pbinom(4, 12, 1/5)

Q3.) Find the complete distribution when n=5, p=0.1

```
> dbinom(0:5, 5, 0.1)
[1] 0.59049
    0.32805
    0.7290
    0.00810
    0.00045
    0.00001
```

Q4.) n=12, p=0.25. Find:

i.) P(X=5)   ii.) P(X≤5)   iii.) P(X>7)

iv.) P(5<X<7)

```
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
```


~~Q3.) Find the complete distribution when n=5, p=0.1~~

~~>dbinom(0:5,5,0.1)~~

Q5.) Probability of salesman making a sell to customer 0.15. Find the probability of.

i.) no sells out of 10 customer

```
>d binom(0,10,0.15)
[1] 0.1908744
```

ii.) [1] 0.3522748

Q6.) A salesman has 20% probability of making a sell to a customer out of 30.

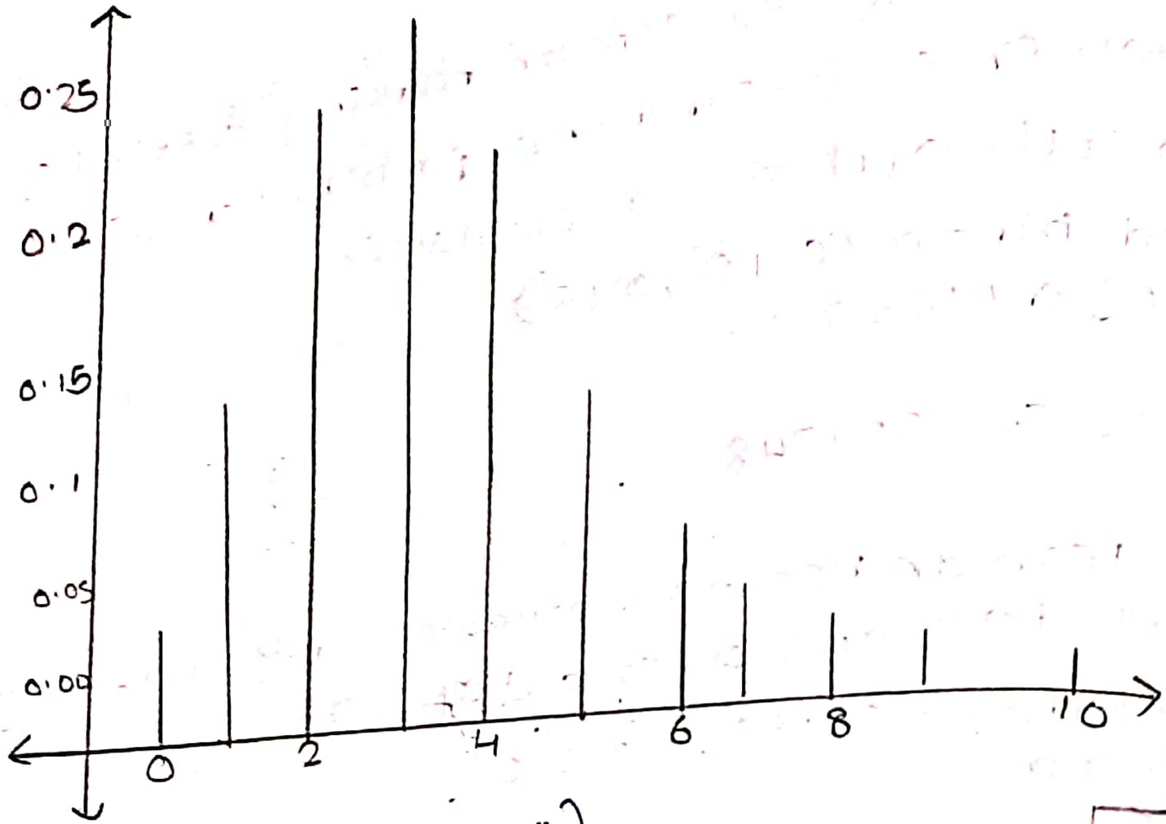
```
> qbinom(0.88, 30, 0.2)
[1] 9
```

Q7.) X followed binomial distribution with n=10, p=0.3.

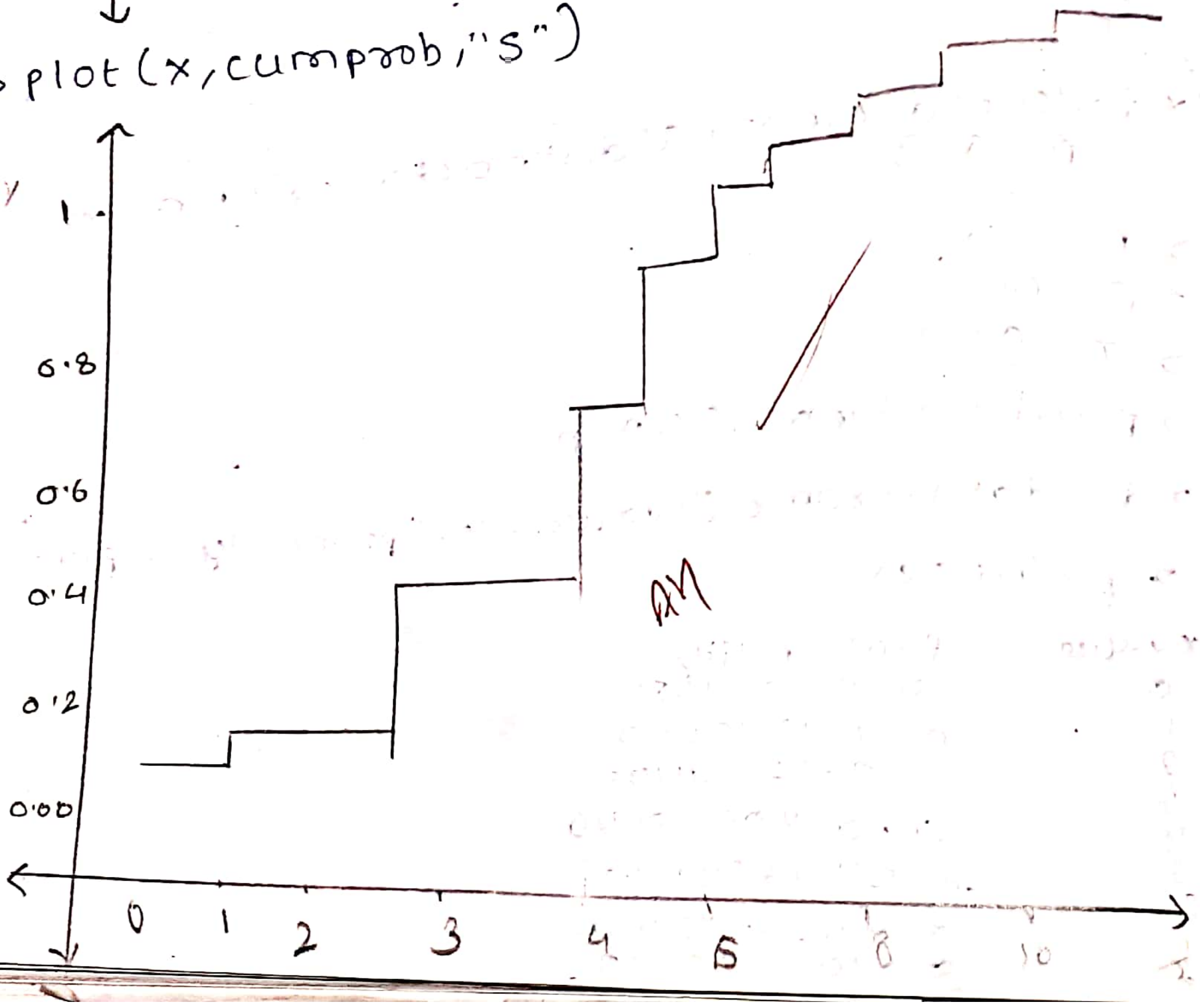
```
> n=10
> p=0.3
> x=0:n
> prob=dbinom(x,n,p)
> d=data.frame("xvalues"=x, "probability"=prob)
> print(d)
```

	xvalues	Probability
1	0	0.02824752
2	1	0.1210608210
3	2	0.233474405
4	3	0.0 200120940
5	4	0.367569090
6	5	
7	6	0.1029193452
8	7	0.1029193452
9	8	

> plot(x, prob, "h")



> plot(x, cumprob, "s")



AM

19

i.) > p1= pnorm (15,12,3)
> p1
> cat ("P(x <=15) =", p1)
[1] 0.8413447
P(X <= 15)= 0.8413447

ii) > p2= pnorm(13,12,3) - pnorm(10,12,3)
> p2
[1] 0.3780661
> cat ("P(10<=x<=13) =", p2)
P(10<=x<=13)=0.3780661

iii > rnorm (5,12,3)
[1] 16.380316 ~~13.5398777~~ 13.050514
4.181289

Q2) X follows normal distribution with $\mu=10$, $\sigma=2$, Find:-

i) > p1= pnorm (7,10,2)
> p1
[1] 0.0668072

ii) > p2 = pnorm (12,10,2)
> p2
[1] 0.8351351

iii) > p3 = 1 - pnorm (12,10,2)
> p3

Q3) Answers:-

```
> x = rnorm(5,15,4)
> x
[1] 14.35444  11.8206  12.44472  15.79432
    16.68769

> am = mean(x)
> cat("Sample mean is = ", am)
Sample Mean is = 14.2208

> me = median(x)
> cat("sample median is", Me)
sample median is = 14.35444

> n = 5
> variance = (n-1) * var(x)/n
> sd = sqrt(variance)
> cat("sample sd is", sd)
  sample is = 1.870638
```

Q5) Plot the Standard Normal graph.

```
> x = seq(-3,3, by=0.1)
> y = dnorm(x)
> plot(x,y,xlab="xvalues", ylab="Probab"
      main="standard Normal Graph")
```

52

Practical-5

Normal and T-Test.

Q1] H0: $\mu = 15$

$H_1: \mu \neq 15$

Random sample of size=400 is drawn and it is calculated the sample mean is 14, S.D is 3. Test the Hypothesis at 5% level of significance.

```
>mo=15
>mx=14
>sd=3
>n=40
>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] 2.616791e-11
```

∴ The pvalue is less than 0.05 then the value of null Hypothesis is rejected.

4) Last years the farmers lost 20% of their crops. A random sample of 60 feilds are collected and it is found that 9 feilds crops are insect polluted. Test the Hypothesis at 1% level of significance.

```
> P=0.2
>p= 9/60
> n=60
> Q=1-P
> zcal= (p-P)/sqrt(P*Q/(n))
> zcal
[1]-0.9682
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
> [1] 0.3329216
```

∵ The value of p is greater than ~~0.05~~ 0.01

∴ The value is accepted

5) Test the Hypothesis from the following sample at $H_0: \mu = 12.5$ 5% level of significance.

```
>x= c(12.25,11.97,12.15,12.08,12.31,12.28,11.94,11.89,12.11,12.
>n= length(x)
>n
>mx= mean(x)
> var= (n-1) * var(x)/n
> sd= sqrt (var)
> sd
> mo=12.5
> t=(mx-mo)/(sd/(sqrt(n)))
>t
[1] -8.894909
>
```



```
> pvalue= 2* (1-pnorm (abs(t)))
> pvalue
[1] 0
```

∵ The pvalue is less than 0.05

∴ It is rejected.

AN 16.1.20

29

23/11/2020

Practical - 6

Topic:: Large sample Test.

i.) Let, the population mean (the amount spent by per customer in a ~~resta~~ hotel) is 250. A sample of 100 customers selected. The sample mean is calculated as 275 and the S.D is 30. Test the hypothesis that the population is 250 or not at 5% level of significance.

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

Ho: $\mu = 275$ against H1: $\mu \neq 275$

Q1] >mo=250
>mx=275
>sd=30
>n=100
>zcal= (mx-mo)/(sd/sqrt(n))
>zcal
[1] 8.3333
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0

∵ pvalue is less than 0.05
∴ It is rejected at 5% level of significance.

66

Q2] Ho: P=0.8 against H1: P≠0.8
>P=0.8
>Q=1-P
>p=750/1000
>n=1000
>zcal= (p-P)/(sqrt(P*Q/n))
>zcal
[1] -3.952847
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 7.72268e-05

∵ pvalue is less than 0.01
∴ It is rejected at 1% level of significance.

iii] Two random sample of size 1000 and 2000 are drawn from two population with the same s.d=2.5, The sample means are 67.5 and 68 respectively. Test the Hypothesis Ho: $\mu_1 = \mu_2$ against H1: $\mu_1 \neq \mu_2$ at 5% level of significance.

Ho

iv) A study of noise level in a two hospitals is given below. Test the claim that the two Hospitals have the same level of noise at 1% level of significance.

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

v) In a sample of 600 students in a college 400 use blue ink, in another college from sample of 900 students 450 use blue ink test the Hypothesis that the proportion of students using blue ink in two colleges are equal or not at 1% level of significance.

iii) $H_0: \mu_1 = \mu_2$ ag $H_1: \mu_1 \neq \mu_2$

```
> n1 = 1000
> n2 = 2000
> mx1 = 67.5
> mx2 = 68
> sd1 = 2.5
> sd2 = 2.5
> zcal = (mx1 - mx2) /sqrt ((sd1^2/n1) + (sd1^2/n2))
> zcal
[1] -5.163978
> pvalue = 2* (1-pnorm(abs (zcal)))
> pvalue
[1] 2.417564e-07
```

∵ pvalue is less than 0.05

∴ It is rejected at 5% level of significance.

iv) $H_0: \mu_1 \neq \mu_2$ ag $H_1: \mu_1 \neq \mu_2$

```
> 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
1.162528
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 0.2450211
```

∵ pvalue is greater that 0.01

∴ It is accepted at 1% level of signifi

v) $H_0: P_1 = P_2$ 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)
> q = 1 - p
> zcal = (p1 - p2) /sqrt( p*q*(1/n1 + 1/n2))
> zcal
[1] 6.381534
> pvalue = 2* (1-pnorm (abs(zcal)))
> pvalue
[1] = 1.735222e-10
```

∵ pvalue is lessthan 0.01

∴ It is rejected at 1% level of significance

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

```
> 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))
> zcal
[1] 1.802741
```

> pvalue = 2* (1-pnorm (abs(zcal))
> pvalue
[1] 0.07142888

∵ pvalue is greater than 0.05
∴ pvalue is accepted at 5% level of significance.

AM
23-1-20

Topic:- Small Sample Test:-

Q1] The marks of 10 students are given by 63, 63, 66, 67, 68, 69, 70, 71, 70, 72. Test the Hypothesis that the sample comes from a a population with avg marks 66.

soln:-
H0: μ = 66

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

data: x
t = 68.319, df = 9, p-value = 1.558e-13
alternative hypothesis: true mean is not equal to a 95% confidence interval 65.65171 70.14829.

sample estimates:
Mean of x 67.9

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

Mean of x Mean of y:
26.1 17.5

70.

> pvalue = 0.03798
> attach {
> if (pvalue > 0.05) {cat("Accept)}
else {cat("Reject")}
> Reject.
Since pvalue is less than 0.05
∴ It is rejected.

Q3) The sales data of six shops before and after a special campaign is effective or not:

Before: 53, 28, 31, 48, 50, 42
After: 58, 29, 30, 55, 56, 45
Test the hypothesis that campains is effective or not.

```
> 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 = 2.7815, df = 5, pvalue = 0.9806
Alternative hypothesis: True difference in means greater than 0.

95 percent confidence interval:
-6.035547 Inf

Sample estimates:
Mean of differences
-3.5
>
>pvalue = 0.9806
>if(pvalue > 0.05){cat("Accept")}
else{cat("Reject")}
Accept.
> since p value is greater than 0.05 we accept hypothesis at 5% level of significance

Q4) Ho: There is no 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, pvalue = 0.9963

Alternative hypothesis: true difference in mean is less than 0
95 percent confidence interval:
-Inf 29.0295
Sample estimates:
Mean of difference
19.8333
>pvalue = 0.9963

>if(p value > 0.05) {cat("Accept")}
else {cat("Reject")}

Accept >

Since the pvalue is greater than 0.05 we accept the Hypothesis at 5% level of significance

AH

Answers:

H0: $\mu = 55$

H1: $\mu \neq 55$

> n = 100

> mx = 52

> mo = 65

> sd = 7

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

> zcal

[1] -4.285714

> cat ("calculated z value is :", zcal)

calculated z value is : -4.285714

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

> pvalue

[1] 1.82153e-05

Since pvalue is less than 0.05.

∴ we Reject the Hypothesis at 5% level of significance.

2) H0: μ

> p = 0.5

> Q = 1 - p

> P = 350/700

> n = 700

> zcal = (p - P) / (sqrt (p*Q/n))

> zcal

[1] 0

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

> pvalue

[1] 1

since pvalue is greater than 0.05
∴ we accept Hypothesis at 1% level of significance.

3) $H_0: P_1 = P_2$ as $H_1: P_1 \neq P_2$

```
> n1=1000
> n2=1500
> p1=0.02
> p2=0.01
> p= (n1* p1 + n2* p2)/(n1+n2)
> p
[1] 0.014
> q=1-p
[1] 0.986
> zcal= (p1-p2)/sqrt(p* q * (1/n1+1/n2))
> zcal
[1] 2.084842
> pvalue =2*(1-pnorm(abs(zcal))
> pvalue
[1] 0.03708364
```

Since pvalue is less than 0.05
∴ It is rejected at 5% level of significance.

$H_0: \mu = 100$

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

since pvalue is less than 0.05
∴ we reject at 5% level of significance

5)
```
> x= c (63,63,68,69,71,71,72)
> t.test (x)
```

One sample t-test

data : x
t= 45, df= 5, pvalue = 1.023e-07
alternative hypothesis : True mean is not equal to 0
95 % confidence interval:
63.64413 71.35587
Sample estimates:
Mean of x
67.5

6) ~~>x:x~~
~~Ho: μ=1200~~

6) Ho: ~~The~~ Population have the same variance
Ho: $\sigma_1 : \sigma_2$

```
>x = c(66,67,75,76,82,84,88,90,92)
>y = c(64,66,67,74,78,82,85,87,92,93,95,97)
>var.test (x,y)
```

F test to compare two variables

data: x and y

F = 0.70686, num df = 8, denom df = 10, pvalue = 0.6359

alternative hypothesis: true ratio of variance is not equal to 1 95% confidence interval:
0.1833662 3.0360393

Sample estimates:
Ratios of variable
0.7068567

7) Ho: μ=1200

```
>mx:1150
>mo:1200
>sd:125
>n:100
>zcal: (mx-mo)/(sd/sqrt(n))
>zcal
[1] -4
>pvalue = 2* (1-pnorm (abs(zcal)))
>pvalue
[1] 6.33424 e-05
```

since pvalue is less than 0.05
∴ It is rejected at 5% level of significance.

8) Ho: $P_1 \neq P_2$

```
>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)))
[1] 0.3613104
```

Since pvalue is less than 005 ∴ It is rejected at 5% level of significance.

AH 17/3/20

PRACTICAL-9

Use the following data to test whether the condition of Home ~~or~~ and condition of ~~a~~ child are independant or not.

condition of Home

	Clean	Dirty
cond. of child clean	70	50
fair	80	20
Dirty	35	45

Soln:-

Ho: condition of Home & child are independant

~~x= scan (70, 80, 35, 50, 20, 45)~~

```
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
xsquared= 25.646 , df= 2,   pvalue - 2.698e-06
```

since pvalue is less than 0.05

∴ It is rejected.

∴ They are not independant

27 Test the Hypothesis that the vaccination and the disease are independant or not:-

		Vaccine	
		Aff	NA
Disease	Aff.	70	46
	NA.	35	37

Soln:-

Ho: vaccination and disease are independant.

```
> x=c (70, 35, 46, 47)
> m= 2
> n= 2
> y = matrix( x, nrow= m, ncol= n)
> y
> pv = chisq.test(y)
> pv
```

Pearson Chi square test

data: y

X-squared = 2.0275, df=1, pvalue= 0.1545

> oneway.test(values~ind, data=d, var.equal=T)

data: values and ind

F= 11.735, num df = 3, denom df = 9, ~~pvalue~~ 0.0
pvalue = 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		

∵ pvalue is less than 0.05
∴ It is accepted.

The following data the life of tyres of 4 brand

Type	Observation
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)

```
> x1 = c (20, 23, 18, 17, 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)
[1]   "values"   "ind"
> oneway.test(values ~ ind, data = d, var.equal = T)

F = 6.8445, num df = 3, denom df = 20, pvalue = 0.0023

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

∵ pvalue is less than 0.05
∴ It is rejected.

Q5) Copy Data from Excel file

```
>
> x = read.csv ("C:/users/admin/Desktop/marks.csv")
> x
```


	Stas	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

```
> am = mean (x $ Stats)
> am
[1] 37
> am = mean (x $ Maths )
> am
[1] 34.4
> m1 = median (x $ Stats)
> m1
[1] 38.5
> m2 = median (x $ Maths)
> m2
[1] 37
> n = length (x $ Stats)
> n
[1] 10
> Sd = sqrt ( (n-1) * var( x$Stats ) /n ))
> Sd
[1] 12.64911
> n = length(x $ Maths)
> Sd = sqrt ((n-1)* var (x$Maths)/n ))
> Sd
[1] 15.2
> cor (x$Stats, x $ Maths)
[1] 0.830618
```

A1 2/12

Topic:: Non-Parametric Test

Q1] Following are the amount of sulphur oxide emitted by industry in 20 days. Apply signtest to test the Hypothesis that median is 21.5 at 5% level of significance.

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

Ho: population median is 21.5

```
> x = c(17. .. . 26)
> length(n)
[1] 20
> me = 21.5
> sp = length(n[x>me])
> sn = length(n[x<me])
> n = sp+sn
> pv = pbinum(sp, n, 0.5)
> pv
[1] 0.4119015
```

∴ pvalue is more than 0.05
∴ we accept the Hypothesis.

Q2.) Following is the data of ten observation apply signtest to test the Hypothesis that the population median is 625. begins the alternative it is more than 625.

612, 619, 631, 628, 643, 640, 655, 649, 670, 663

Ho: population median is 625

```
>x= c(622 . . 633)
>length(x)
[1] 10
> me= 625
>sp= length (x[x>me])
>sn= length (x[x< me])
> pv= pbinorm (sn, n, 0.5)
> pv
[1] 0.0546875
```

∴ Pvalue is more than 0.05 Hence we accept the Hypothesis.

Q5]

The weight of students before and after they stop smoking below. Using WSRT test that there is no significance chance

Before(x)	(y) After
65	72
75	74
75	72
62	66
72	73

H0: There is no change

H1: There is change

```
>x= c(65,75,75,62,72)
>y= c(72,74,72,66,73)
>d= x-y
>d
[1] -7 1 3 -4 -1
>wilcox.test(x,alter"twoside", mu=0)
```

data: x

V= 15 , pvalue= 0.05791

∵ pvalue is greater than 0.05

∴ we accept the Hypothesis.

A.M. 5.7.20