R is a software for data analysis and statistical computing.
This software is used for effective data handling and output storage is possible.
It is capable of graphical display.
It is a free software.

Questions :-

Q. i) $2^2 + \sqrt{25} + 35$
ii) $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$
iii) $\sqrt{76 + 4 \times 2 + 9 \div 5}$
iv) $42 + |-10| + 7^2 + 3 \times 9$

Solutions :-

i) $2^2 + \sqrt{25} + 35$
Command : 2^2 + sqrt(25) + 35
output : 44

ii) $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$
Command : 2 * 5 * 3 + 62/5 + sqrt(49)
output : 49.4

iii) $\sqrt{76 + 4 \times 2 + 9 \div 5}$
Command : sqrt (76 + 4 * 2 + 9/5)
output : 9.262829

iv) $42 + |-10| + 7^2 + 3 \times 9$
Command : 42 + abs(-10) + 7^2 + 3 * 9
output : 128

v) $x = 20$; $y = 30$
Find $x+y$, $x^2 + y^2$, $\sqrt{y^3 - x^3}$, $|x - y|$
Command : x + y
output : 50
command : x^2 + y^2
output : 1300
command : sqrt (y^3 - x^3)
output : 137.8405
command : abs (x - y)
output : 10

vi) c(2, 3, 4, 5)^2
output : 4 9 16 25

vii) c(4, 5, 6, 8) * 3
output : 9 12 15 18 27

viii) c(2, 3, 5, 7) * c(-2, -3, -5, -4)
output : -4, -9, -25, -28

ix) c(2, 3, 5, 7) * c(8, 9)
output : 16 27 40 63

x) c(2, 3, 5, 7) * c(1, 2, 8)
output : warning Message : obj length is not a multiple of shorter obj length.

(xi) c(1, 2, 3, 4, 5) ^ c(2, 3, 4)
output : Warning message

(xii) c(1, 2, 3, 4, 5, 6) ^ c(2, 3, 4)
output : 1 8 81 16 125 1296

(xiii) Find the sum, product, maximum, minimum of the values:
5, 8, 6, 7, 9, 10, 15, 5

Command : x = c(5, 8, 6, 7, 9, 10, 15, 5)
len (x)
sum (x)
prod (x)
max (x)
min (x)

output :
8
65
11340000
15
5

(xiv) $\begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$

Command : x ← matrix (nrow = 4, ncol = 2, data = c(1, 2, 3, 4, 5, 6, 7, 8))

output :- [, 1] [, 2]
1 5
2 6
3 7
4 8

xv) $X = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ $Y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ 10 & 6 & 12 \end{bmatrix}$

16.

Find $x+y$, $x \times y$, $2x + 3y$

command: x <- matrix(nrow=3, ncol=3, data=c(1,2,3,4,5,6,7,8,9))
y <- matrix (nrow=3, ncol=3, data=c(2,-2,10,4,8,6,10,-11,12))
x
y

output: $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$

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

command: x+y
output:

3	8	17
0	13	-3
13	12	21

command: x*y
output:

2	16	40
-4	40	-88
30	36	108

command: 2*x + 3*y
output:

8	20	44
-2	34	-17
36	30	54

(xvi) Command: x = c(2, 4, 6, 1, 3, 5, 7, 18, 16, 14, 17, 19, 3, 3, 2, 5, 0, 15, 9, 18, 10, 12)
length(x)
a = table(x)
transform(a)

output:
> 23
> x freq

x	freq
0	1
1	1
2	2
3	3
4	1
5	2
6	1
7	1
8	1
9	1
10	1
12	1
14	2
15	1
16	1
17	1
18	2
19	1

Command: seq (0, 20, 5)
h = cut(x, breaks, right = FALSE)

18

c = table (c)
transform c

output:

b	freq
[0, 5)	8
[5, 10)	5
[10, 15)	4
[15, 20)	6

AM

PRACTICAL No. 2:

32

Problems on p.d.f and c.d.f

1) Can the following be p.d.f ?

(i) $f(x) = \begin{cases} 2-x & ; 1 \le x \le 2 \\ 0 & o.w \end{cases}$

Sol:- $\int f(x)\, dx = 1$

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

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

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

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

$= 2 - 1.5$

$= 0.5$

$\neq 1$ $\therefore f(x)$ is not a p.d.f.

(ii) $f(x) = \begin{cases} 3x^2 & ; 0 < x < 1 \\ 0 & o.w \end{cases}$

Sol:- $\int f(x) ~~= 3x^2~~ dx = 1$

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

$= \int_0^1 \left[\frac{3x^3}{3}\right]_0^1$ $\quad \therefore x^n = \frac{x^{n+1}}{n+1}$

$= \frac{3 \cdot (1)^3}{3} - \frac{3(0)^3}{3}$

$= \frac{3}{3}{}^1 - 0$

$= 1 - 0$

$= 1$

$\therefore f(x)$ is a p.d.f

(iii) $f(x) = \begin{cases} \frac{3x}{2}\left(1 - \frac{x}{2}\right) & ; 0 \le x \le 2 \\ 0 & o.w \end{cases}$

Sol:- $\int f(x)\,dx = 1$

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

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

$= \left[\frac{3x^2}{4}\,dx - \frac{3x^3}{12}\,dx\right]_0^2$

$= \left[\frac{3x^2}{4}\right]_0^2 - \left[\frac{3x^3}{12}\right]_0^2$

$= \left[\frac{3(2)^2}{4}\right] - \left[\frac{3(2)^3}{12}\right]$

$= \frac{12}{4} - \frac{24}{12}$

$= 3 - 2$

$= 1$

$\therefore f(x)$ is a p.d.f

Q.2) Can the following be p.m.f?

(i)

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

Sol:- Since, one probability is negative, $P(x)$ is not a p.m.f

(ii)

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

Sol^n Since $P(x) \ge 0 \; \forall x$ and $\Sigma P(x) = 1$

(iii)

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

Sol^n Since, $\Sigma P(x) = 1.2$, therefore it is not a p.m.f

$\therefore P(x)$ is not p.m.f.

Q.3) Find $P(x \le 2)$, $P(2 \le x < 4)$; 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

~~$P(x \le 2)$~~ Since, $P(x) \ge 0 \; \forall x$ and $\Sigma p(x) = 1$

$\therefore p(x)$ is a p.m.f.

(i) $P(x \le 2)$

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

$= 0.4$

(ii) $P(2 \le x < 4)$

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

$= 0.4$

(iii) P(atleast 4)

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

$= 0.1 + 0.2 + 0.1 = 0.4$

(iv) $P(3 < x < 6)$

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

$= 0.1 + 0.2$

$= 0.3$

33

Q.4) Find c.d.f

(i)

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

Sol^n:-

~~cdm~~ $F(x) = 0$ if $x < 0$
$= 0.1$ if $0 \le x < 1$
$= 0.2$ if $1 \le x < 2$
$= 0.4$ if $2 \le x < 3$
$= 0.6$ if $3 \le x\ 4$
$= 0.7$ if $4 \le x, 5$
$= 0.9$ if $5 \le x\ 6$
$= 01.0$ if $\cancel{6 \le}\ x \ge 6$

(ii)

x	10	12	14	16	18
$p(x)$	0.2	0.35	0.15	0.2	0.1

Sol^n:-

$f(x) =$ 0.20 if $x < 10$
0.55 if $10 \le \cancel{x} < 12$
0.70 if $12 \le x < 14$
0.90 if $14 \le x < 16$
1.00 if $x \ge 18$

AIM:- Probability and Binomial Distribution

Q.1) Find the CDF of the following PDF and draw the graph.

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

Sol^n:-

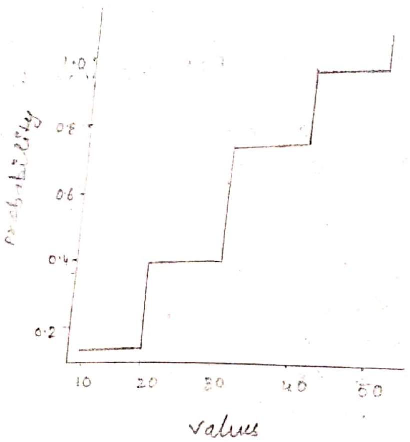
$f(x) = 0$ if $x < 10$
$= 0.15$ $10 \le x < 20$
$= 0.40$ $20 \le x < 30$
$= 0.70$ $30 \le x < 40$
$= 0.90$ $40 \le x < 50$
$= 1.00$ $x \ge 50$

Commands :-

```
> x = c (10, 20, 30, 40, 50)
> x
> [1] 10 20 30 40 50
> Prob = c(0.15, 0.25, 0.3, 0.2, 0.1)
> Prob
> [1] 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 :-

Graph of C.D.F



* Binomial Distribution

Q.1) Suppose there are 12 MCQ's in a test, each question have 5 questions and only one of them is correct. Find the probability of having :- 1) Five correct answer
2) Atmost four correct answer

Soln:-

It is given that $n = 12$, $p = 1/5$, $q = 4/5$, $\cancel{x = 5}$

x = Total number of correct answer

$x \sim B(n, p)$

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

Command :-

```
> dbinom (5, 12, 1/5)
```



```
>[1] 0.05315022

2) >pbinom (4, 12, 1/5)
>[1] 0.9274445
```

Q.2) There are 10 members in a committee the probability of any member attending a meeting is 0.9 find the probability :-
1) Seven members attended.
2) Atleast 5 member attended.
3) Atmost 6 member attended

Sol:

Here, $n = 10$, $p = 0.9$, $q = 0.1$

1) x = Total numbers of ~~all~~ member attended

$x \sim B(n, p)$

$n = 10$, $p = 0.9$, $q = 0.1$

commands :-

```
i) > dbinomial (7, 10, 0.9)
   > [1] 0.057395563

ii) > 1- pbinom (5, 10, 0.9)
    > [1] 0.9983651

iii) > pbinom (6, 10, 0.9)
     > [1] 0.0127952
```

Q. 3) Find the c.d.f and draw the graph

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

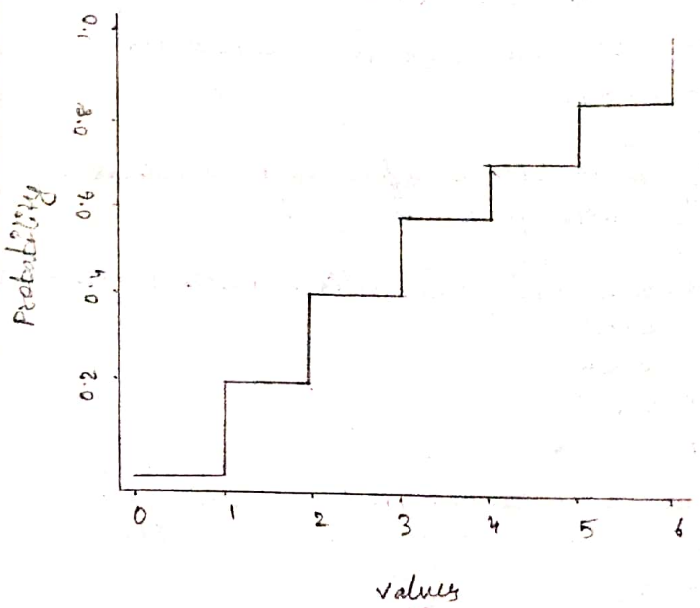
Sol^n

$F(x) = $	0		if $x < 0$
~~F(x)~~ =	0.1	if $x < 0$	$0 \le x < 1$
	0.2	$0 \le x < 1$	$1 \le x < 2$
	0.4	$1 \le x < 2$	$2 \le x < 3$
	0.6	$2 \le x < 3$	$3 \le x < 4$
	0.7	$3 \le x < 4$	$4 \le x < 5$
	0.9	$4 \le x < 5$	$5 \le x < 6$
	1.0	~~$5 \le x < 6$~~	$x \geq 6$

Commands :-

```
x = c(0, 1, 2, 3, 4, 5, 6)
[1] 0 1 2 3 4 5 6

prob = c(0.1, 0.1, 0.2, 0.2, 0.1, 0.2, 0.1)
prob
[1] 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= "probabil
      main = "Graph of C.D.F ", "s")
```


Graph of C.D.F



AM
11-12-19

Practical No. 4

1) Find the ~~complete~~ complete binomial distribution
$n = 5$ & $p = 0.1$

2) Find the probability of exactly 10 success in 100 trial with $p = 0.1$

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

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

5) A student writes 5 MCQ. Each question has 4 options out of which 1 is correct. Calculate the probability for atleast 3 correct answers.

6) X follows binomial distribution $n = 10$, $p = 0.4$, plot the graph of p.m.f & c.d.f.

Note :- 1) $P(x = x)$, dbinom (x, n, p)
2) $P(x \le x)$, pbinom (x, n, p)
3) $P(x > x)$, 1- pbinom (x, n, p)
4) To find the value of x, for which the probability is given as p1. The command will be :- qbinom (p1, n, p)

Solutions :-

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. $n = 100$, $p = 0.1$, $x = 10$
> dbinom (10, 100, 0.1)
[1] 0.1318653

3. i) $n = 12$, $p = 0.25$, $x = 5$
> dbinom (5, 12, 0.25)
[1] 0.1032414

ii) $n = 12$, $p = 0.25$, $x = 5$
> pbinom (5, 12, 0.25)
[1] 0.9455978

iii) $n = 12$, $p = 0.25$, $x = 7$
> 1- pbinom (7, 12, 0.25)
[1] 0.00278151

iv) $n = 12$, $p = 0.25$, $x = 6$
> dbinom (6, 12, 0.25)
[1] 0.04014945

38

4) $n = 10$, $p = 0.15$, $x = 0$

```
> dbinom(0, 10, 0.15)
[1] 0.1968744
```

2) $n = 20$, $p = 0.15$, $x = 3$

```
> 1- pbinom(3, 20, 0.15)
[1] 0.3522748
```

5) $n = 5$, $p = 1/4$, $x = 2$

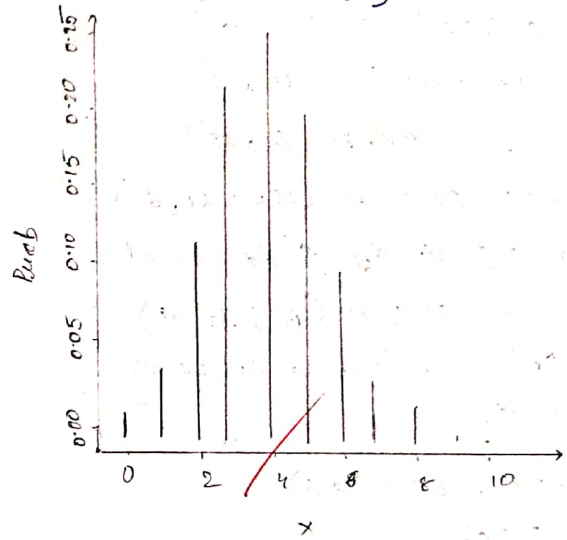
```
> 1- pbinom(2, 5, 1/4)
[1] 0.1035156
```

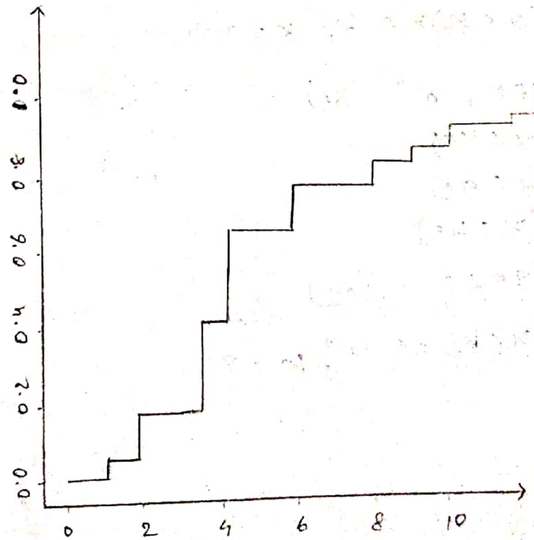
6)

```
n = 10
p = 0.4
x = 0:n
prob = dbinom(x, n, p)
~~prob~~
cumprob = pbinom(x, n, p)
g = data.frame("xvalues" = x, "Probability" = prob)
g
```

	xvalues	Probability
1	0	0.0060466176
2	1	0.0403107840
3	2	0.1209323520
4	3	0.2149908480
5	4	0.2508226560
6	5	0.2006581248
7	6	0.1114767360
8	7	0.0424673280
9	8	0.0106168320
10	9	0.0015728640
11	10	0.0001048576

```
plot(x, prob, "h")
plot(x, cumprob, "s")
```

→ 

→ 

PRACTICAL NO. 5

Normal Distribution

1. $P[X = x]$ = dnorm (x, μ, σ)
2. $P(x \le x)$ = 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)
5. To find the value of k so that:
   $P(x \le k) = P_1$; qnorm (p_1, μ, σ)
6. To generate 'n' random numbers:
   rnorm (n, μ, σ)

Q.1) $X \sim N \ (\mu = 50, \ \sigma^2 = 100)$
Find: 1) $P(x \le 40)$
2) $P(42 \le x \le 60)$
3) $P(x > 55)$
4) $P(x \le k) = 0.7$, $k = ?$

Q.2) $X \sim N \ (\mu = 100, \ \sigma^2 = 36)$
1) $P(x \le 110)$
2) $P(x \le 95)$
3) $P(x > 45)$
4) $P(95 \le x \le 105)$
5) $P(x \le k) = 0.4$; $k = ?$

Q.3) Generate 10 random number from normal distribution with mean $(\mu) = 60$, standard deviation $(\sigma) = 0.5$. Also calculate the sample mean, median, variable and standard distribution

Q.4) Draw the graph of standard normal distribution.

Solution :-
1) 1) $P(X \le 40)$
> a = pnorm (40, 50, 10)
> cat ("P(x <= 40) is : ", a)
[1] P(x <= 40) is : 0.1586553

2) c = pnorm (60, 50, 10) - pnorm (42, 50, 10)
> cat ("P(42 < x <= 60) is : ", c)
[1] 0.6294893

3) $P(x > 55)$
> b = 1 - pnorm (55, 50, 10)
> cat ("P(x>55) is : ", b)
[1] P(x>55) is : 0.3085375

4) $P(x \le k) = 0.7$, $k = ?$
> d = qnorm (0.7, 50, 10)
> cat ("P(x <= k) = 0.7, K is : ", d)
[1] P(x <= k) = 0.7, K is : 55.2440

2) i) P(x ≤ 110)
```
> a = pnorm (110, 100, 6)
> cat ("P(x <=110) is : ", a)
[1] P(x <= 110) is : 0.9522096
```

2) P(x ≤ 95)
```
> b= pnorm (95, 100, 6)
> cat ("P(x <= 95) is : ", b)
[1] P(x<=95) is : 0.2023284
```

3) P (x>115)
```
> c= 2 - pnorm (115, 100, 6)
> cat("P(x>115) is : ", c)
[1] P(x>115) is : 0.006209665
```

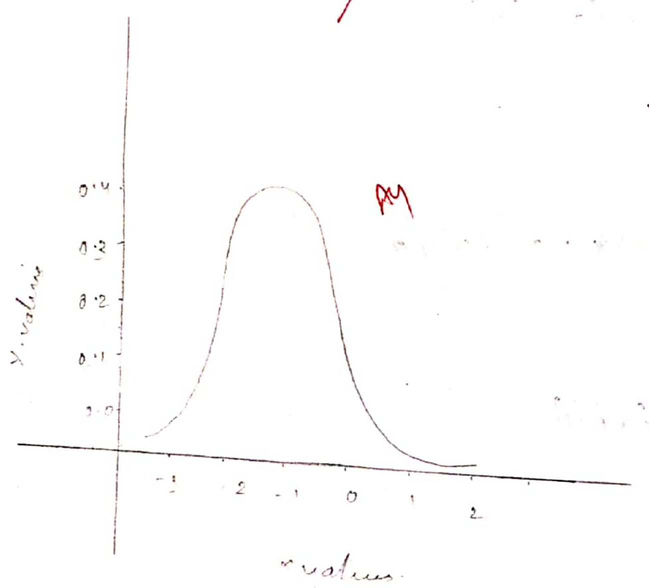
4) P(95<= x <= 105)
```
> d= pnorm (95, 100, 6)- pnorm (105, 100, 6)
> cat ("P (95 <= x <= 105) is : ", d)
[1] P(95 <= X<=105) is : -0.5953432
```

5) P(x≤k) = 0.4, k is 9
```
e = qnorm(0.4, 100, 6)
> cat ("P(x <= k) = 0.4, k is : ", e)
[1] P(X<=k) = 0.4 , K is : 98.47992
```


3)
```
> x = rnorm (10, 60, 5)
> x
[1] 62.4310   59.9626   63.72094   61.2825   61.9102
    51.8524   66.9836   62.5194    61.3294   62.4803

> am= mean(x)
> am
[1] 60.44768

> median = median (x)
> median
[1] 61.6198

> n= 10
> variable= (n-1)* var (x)/n
> variable
[1] 20.37836

> sd= sqrt (variable)
> sd
[1] 4.5142
```

4) Draw the graph of standard normal distribution

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



AY

PRACTICAL No. 6

AIM :- ~~###~~ Z distribution

Mean Test :

Q.1) Test the hypothesis (H_0):

i) H_0 : $\mu = 10$ against H_1 : $\mu \neq 10$

A sample of size 400 is selected which gives a mean 10.2 and standard deviation 2.25.

Test ~~they~~ the hypothesis at 5% level of significance.

Sol^n :-

```
mo (mean of population) = 10
mx (mean of sample) = 10.2
sd = 2.25
n (sample size) = 400
zcal = (mx - mo) / (sd / (sqrt (n)))
cat = ("Zcal is : ", zcal)
>> Zcal is 1.77778
pvalue = 2 * (1 - pnorm (abs (zcal)))
pvalue
>> 0.07544
```

, Thus hypothesis accepted

(NOTE :- If result of tested hypothesis is >0.05, then assumed hypothesis of H_0 : $\mu = 10$ is accepted as verified (correct)

Q.2) Test the hypothesis (H_0):
$H_0: \mu = 75$ against $H_0: \mu \neq 75$
A sample of size 100 is selected and the sample mean is 80, with sd of 3. Test the hypothesis at 5% of significance.

Soln:-

mo = 75
mx = 80
Sd = 3
n = 100
Zcal = (mx - mo) / (sd /(Sqrt(n)))
cat = ("Zcal : ", zcal)
pvalue = 2* (1 - pnorm(abs(Zcal)))

OUTPUT :-
>> Zcal : 16.667
>> pvalue is : 0

Thus, hypothesis rejected.

Q.3) Test the hypothesis (H_0):
$H_0: \mu = 25$ against $H_0: \mu \neq 25$
A sample of 30 is rejected. Test the hypothesis at 5% level of significance.

Sample(x) = 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

Soln:-
mx = mean(x)
>> mx = 26.066
n = length(n)
variance = (n-1) * var(x) / n
sd = Sqrt (variance)
>> sd = 7.2798
mo = 25
zcal = (mx - mo) / (sd / sqrt (n))
>> zcal = 0.8025
pvalue = 2* (1- pnorm (abs(zcal)))
>> pvalue = 0.4223

Thus, hypothesis accepted

* Proportion Test:

Q.1) Experience has shown that 20% students of a college smoke. A sample of 400 students reveal that out of 400 only 50 smoke. Test the hypothesis that the experience keeps the correct proportion or not.

Soln:-
pp = 0.2 (20%)
q = 1- pp
Sq = 50/400
n = 400
zcal = (sp-pp) /(sqrt(pp*q/n)
>> Zcal = -3.75
pvalue = ~~0.00017~~ 2* (1- pnorm (abs (zcal)))
>> pvalue = 0.00017

Thus, hypothesis is rejected.

Q.2) Test the hypothesis : $H_0 : \mu = 0.5$ against $P \neq 0.5$
A sample of 200 is selected where Sp = 0.56 and PP = 0.5 , level of significance is 1%

Solⁿ :-

pp = 0.5

q = 1 - pp

sp = 0.56

n = 200

zcal = (sp - pp) / (sqrt (pp * q/n)

» zcal = 1.6970

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

» pvalue = 0.08968

Thus, hypothesis is accepted.

PRACTICAL - 07

AIM :- Large Sample test

Q.1) A study of noise level in two hospital is calculated below test the hypothesis that the noise in two hospital are same or not.

	Hospital A	Hospital B
No. of sample obj	84	34
Mean	61	59
Standard Deviation	7	8

Sol: H_0 : The noise levels are same

n_1 = 84

n_2 = 34

mx = 61

my = 59

sdx = 7

sdy = 8

```
> z = (mx - my) / sqrt ((sdx^2 / n1) + (sdy^2 / n2))
> z
[1] ~~1.5004~~ 1.2736
> cat ("z calculate is :", z)
> z calculate is = 1.2736
> pvalue = 2 * (1 - pnorm (abs (z)))   > [1] 0.20550
```

Since pvalue > 0.05, we ~~reject~~ accept H_0 at 5% level of significance.

Q.2) Two random sample of size 1000 & 2000 are drawn from two population with the means 67.5 & 68 respectively and with the same S.D of 2.5; test the hypothesis that the mean of two population are equal.

Sol^n

H0: Two 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))
>z
>-5.163978
>cat ("Z calculate is = ", z)
>Zcalculate is -5.163978
>pvalue = 2*(1-pnorm(abs(z)))
>pvalue
>2.417564e-07
```

Since, pvalue < 0.05, we reject the H0 at 5% level of significance.

Q.3) In F.Y. BSC 20% of a random sample of 400 students had defective eye sight in S.Y.BSC 15.5% of 500 sample had a same defect is the difference of proportion is same?

Sol^n:- H0: The proportion of the population are equal.

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

Since, pvalue > 0.05, we accept H0 at 5% level of significance.

Q.4) From each of the box of the apples, a sample size of 200 is collected. It is found that there are 44 bad apple in the 1st sample and 30 bad apple in second apple. Test the hypothesis that the two boxes are equivalent in terms of no. bad apples.

Soln:- Ho :- Two boxes are equal

```
n1 = 200
n2 = 200
p1 = 44/200
p2 = 30/200
p = (n1*p1+n2*p2) /(n1+n2)
p
[1] 0.185
q = 1-p
q
[1] 0.815
z = (p1-p2) /sqrt(p*q*(1/n1 + 1/n2))
z
[1] 1.802741
pvalue = 2*(1-pnorm(abs(z)))
pvalue
[1] 0.07142888
```

Since, pvalue > 0.05, we accept the Ho at 5% level of significance.

Q.5) In MA class out of a sample of 60 mean height is 63.5 inch, with a standard deviation of 2.5. In a M.Com class of of 50 students mean height is 69.5 inch, with a standard deviation of 2.5, test the hypothesis of that the mean of MA & M.com class are same.

Soln

Ho :- Height of two classes are same.

```
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))
z
[1] -12.53359
cat("Z calculate is = ", z)
z calculate is = -12.53359
pvalue = 2*(1- pnorm (abs(z)))
pvalue
[1] 0
```

Since, pvalue < 0.5, we ~~accept~~ reject the Ho at 5% level of significance.

AM
5.2.20

PRACTICAL No. 8

AIM :- Small Sample Test

Q.1) The flower stems are selected and the hights are found to be 63, 63, 68, 69, 71, 71, 72 cm test the hypothesis that the mean hight is 66 cm or not at 1% level of significance.

Sol:→ H_0: Mean = 66 cm

```
> 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 confidence interval:

64.66479   71.62092

Sample estimates:

mean of x

68.14286

Since, pvalue is < 0.01, we reject H_0 at 1% level of significance.

Q.2) Two random sample that drawn from two different population

Sample1 :- 8, 10, 12, 11, 16, 15, 18, 7

Sample2 :- 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.

Sol:- H_0: 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)
```

Welcome Two sample t-test

data: x and y

t = -0.36247, df = 13.887, p-value = 0.7225

alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval:

-5.199712   3.649719

sample estimates:

mean of x   mean of y

12.125   12.878

Since, pvalue is > the 0.05, we accept the H_0 at 5% level of significance.

Q.3) Followings are the weights of 10 people before after a diet program. Test the hypothesis that the diet program is effective or not.

Ans- Before (100, 125, 95, 96, 98, 112, 115, 104, 109, 110)
After (95, 80, 95, 98, 90, 100, 110, 85, 100, 101)

Sol:- Ho: The diet program is not affective.

```
> before = c(100, 125, 95, 96, 98, 112, 115, 104, 109, 110)
> after = c(95, 80, 95, 98, 90, 100, 110, 85, 100, 101)
> t.test (before, after, paired = T, alternative = "less")
```

Paired t-test

data: before and after
t = 2.6089, df = 9, p-value = 0.9858
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf 18.22908
Sample estimates:
mean of the difference
11

Since, pvalue is > 0.05, we accept the Ho at 5% level of significance.

Q.4) The marks before and after a training program.
Before = (20, 25, 32, 28, 27, 36, 35, 25).
After = (30, 35, 32, 37, 37, 40, 40, 23)
test the Ho that the training program is effective or not.

Sol:- Ho:- The training program is not affecting

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

paired t-test

data: a and b
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

Since, pvalue > 0.05, we accept the Ho at 5% level of significance.

Q.5) Two random sample were drawn from two normal population at 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 wether the population have same variance at 5% percent level of significance

Soln:-
Ho:- The variances of two populations are equal.

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

        F test to compare two variances

data:  A and B
F = 0.70686, num df = 8, denom df = 10,
p-value = 0.6359
alternative hypothesis: true ratio of variances
is not equal to 1
95 percent confidence interval:
0.1833662  3.0360393
sample estimates:
ratio of variances
        0.7068567
```

Since, pvalue > 0.05, we accept the Ho at 5% level of significance

Q.6) The arithmetic mean of sample of 100 obs. is 52 if the S.D is 7, test the hypothesis that the population mean is 55 or not at 5% of LOS

48

Sol:- Ho: Population mean = 55

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

Since, pvalue < 0.05, we reject the Ho at 5% level of significance.

AM
12

Pratical No:- 9

AIM :- Chi Square Distribution and ANOVA

Q.1) Use the following data to test whether the cleanliness of a home depends upon the child condition or not :

		Condition of Home clean	Dirty
condn of child	clean	70	50
	Fairly clean	80	20
	Dirty	35	45

Sol^n :- Ho :- 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

∴ Since, pvalue < 0.05 we reject Ho, at 5% level of significance.

Q.2) Data below shows relations between the performance between maths & computer of CS students.

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

Sol:

H0 :- Performance between two subjects are independant

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

m = 3

n = 3

y = matrix (x, nrow = m, ncol = 3)

y

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

pv = chisq.test (y)

pv

Pearson's chi-squared test

data = y

x-squared = 145.78, df = 4, pvalue < 2.2e

Since, pvalue < 0.05, we reject H0 at 0.05 i.e 5% level of significance

Q.3) Perform ANOVA for the following data.

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

Sol: H0: The means of variety of A, B, C, D are equal.

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)

"values" "ind"

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

oneway analysis of means

data: values and ind

F = 11.735, num df = 3, p-value = 0.00183

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		

∵ pvalue < 0.05, we reject hypothesis H0, at 5% level of significance.

Q.4) Perform ANOVA

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

Sol^n: H0 : Means of types A, B, C, D are equal

x1 = c(6, 7, 8)
x2 = c(4, 6, 5)
x3 = c(8, 6, 10)
x4 = c(6, 9, 9)

d = stack (list (b1 = x1, b2 = x2, b3 = x3, b4 = x4))
names (d)
"values" "ind"
oneway.test (values ~ ind, data = d, var.equal)

one way analysis of means
data: values and ind

F = 2.6667, num df = 3, denom df = 8
p-value = 0.1189

∴ p-value > 0.05, we accept the hypothesis at 5% level of significance.
anova = aov (values ~ ind, data = d)
summary (anova)

Q.5)

x = read.csv ("c:/users/admin/Desktop/marks.csv")

x

	stats	calc
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

Practical No. 10

AIM:- Non- Parametric

Q.1) Following are the amounts of sulphur oxide emitted by a factory.
12, 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 hypo that the population median is 21.5 against the alternative that it is less than 21.5

H0 : population median is 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)
m = 21.5
sp = length (x [x > m])
sn = length (x [x < m])
n = sp + sn
pv = pbinom (sp, n, 0.5)
pv = 0.4119

Since, p-value > 0.05, we accept the hypothesis at 5% level of significance.

NOTE : If the alternative is greater than median pv = pbinom (sn, n, 0.5)]

Q.2) For the observations: 12, 19, 31, 28, 43, 40, 55, 49, 70, 63
Apply sign test, to test hypothesis that population median is 25, against the alternative hypothesis that is greater than 25.

H0 : Median is 25
H1 : Median is greater than 25
x = c(12, 19, 31, 28, 48, 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 = 0.0546

Since, pvalue is > 0.05, we accept the hypothesis H0 at 5% level of significance.

Q.3] From the following data: 60, 65, 63, 89, 61, 71, 58, 51, 48, 66

Test the hypothesis using wilcoxon's signed rank for testing the hypothesis that the median is 60, against the alternative it is greater than 60.

H_0: median is 60
H_1: median is greater than 60

x = c(60, 65, 63, 89, 61, 71, 58, 51, 48, 66).
m = 60

wilcoxon signed rank Test with continuity correction

data: x
V = 29, p-value = 0.2386
alternative hypothesis: true location is greater than 60

Since, p-value > 0.05, we accept the hypothesis H_0 at 5% level of significance

[NOTE: If the alternative is less, wilcox.test (x, alter = "less", mu = 60)]

Q.4) Test the hypothesis that median is 12, against the alternative that it is less than 12, using wilcoxon's signed test values: 12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20

H_0: median is 12
H_1: median is less than 12

x = c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20)
m = 12
wilcox.test (x, alter = "less", mu = 12)

wilcoxon ~~to~~ signed rank test with continuity correction

data: x
V = 25, p-value = 0.2521
alternative hypothesis: true location less than 12

Since, p-value > 0.05, we accept hypothesis H_0 at 5% level of significance.