PRACTICAL NO 1

Topic: Basics of R software

(i) R is a software for data analysis and statistical computin

(ii) This Software is used for effective data handling and out storage is possible.

(iii) It is capable of graphical display

(iv) It is a free software

Q.1]

(i) $2^2 + \sqrt{25} + 35$

```
> (2**2) + sqrt(25) + 35
[1] 44
```

(ii) $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$

```
> (2*5*3) + (62/5) + sqrt(49)
[1] 49.4
```

(iii) $\sqrt{76 + 4 \times 2 + 9 \div 5}$

```
> sqrt(76 + (4*2) + (9/5))
[1] 9.262829
```

(iv) $42 + |-10| + 7^2 + 3 \times 9$

```
> 42 + abs(-10) + (7**2) + (3*9)
[1] 128
```

Q.2] x = 20
y = 30
find $x+y$, x^2+y^2, $\sqrt{y^3-x^3}$, $|x-y|$

```
> x+y
[1] 50
> (x**2)+(y**2)
[1] 1300
> sqrt((y**3)-(x**3))
[1] 137.8405
> abs(x-y)
[1] 10
```

Q.3]

```
> c(2,3,4,5)^2
[1] 4 9 16 25

> c(4,5,6,8)*3
[1] 12 15 18 24

> c(2,3,5,7)*c(-2,-3,-5,-4)
[1] -4 -9 -25 -28

> c(2,3,5,7)*c(8,9)
[1] 16 27 40 63

> c(1,2,3,4,5,6)*(2,3)
[1] 1 8 9 64 25 216
```

Q.4] Find the sum, product, maximum, minimum of the values.
5, 8, 6, 7, 9, 10, 15, 5

```
> x = c(5, 8, 6, 7, 9, 10, 15, 5)
> length(x)
[1] 8
> sum(x)
[1] 65
> prod(x)
[1] 11340000
> max(x)
[1] 15
> min(x)
[1] 5
```

Q.5]
```
> x <- matrix(nrow=4, ncol=2, data=c(1, 2, 3, 4, 5, 6,
8))

> x
[1]      [,1]  [,2]
    [1,]   1     5
    [2,]   2     6
    [3,]   3     7
    [4,]   4     8
```

Q.6)

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

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

```
> x <- matrix (nrow = 3, ncol = 3, data = c(1,2,3,4,6,
9))
> y <- matrix (nrow = 3, ncol = 3, data = c( [illegible] ,4,8,6,
10,-11, 12))

> x + y
     [1]  [2]  [3]
[1]   3    8   17
[2]   0   13   -3
[3]  13   12   21

> x * y
     [1]  [2]  [3]
[1]   2   16   70
[2]  -4   40  -88
[3]  30   36  108
```

```
> x*2 + y*3
     [1]  [2]  [3]
[1]    8   20   44
[2]   -2   34  -17
[3]   36   30   54

> x = c(2,4,6,1,3,5,7,18,16,14,17,19,3,3,2,5,
  0,15,9,14,18,10,12)
> length(x)
[1] 23
> a = table(x)
> a
x
0 1 2 3 4 5 6 7 9 10 12 14 15 16 17 18 19
1 1 2 3 1 2 1 1 1 1  1  2  1  1  1  2  1

> transform(a)

[1]  x   Freq
     0   1
     1   1
     2   2
     3   3
     4   1
     5   2
     6   1
     7   1
     9   1
     10  1
```

12	1
14	2
15	1
16	1
17	1
18	2
19	1

```
> breaks = seq (0, 20, 5)
> S = cut ( x, breaks, right = FALSE)
> c = table (S)
> transform (c)
```

```
[1]          S       Freq
      1   [0, 5)      8
      2   [5, 10)     5
      3   [10, 15)    4
      4   [15, 20)    6
```

M

PRACTICAL NO 2

Q.1] Can the following be p.d.f?

(i) $f(x) = \begin{cases} 2-x \ ; & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$

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

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

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

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

$\neq 1$

$\Rightarrow$ Not a p.d.f

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

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

$= \int_0^1 \frac{\cancel{3}x^3}{\cancel{3}}$

$= x^3\Big|_0^1$

$= 1$

$\Rightarrow$ It is a p.d.f

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

$$\Rightarrow \int_0^2 \left(\frac{3x}{2} - \frac{3x^2}{4}\right) dx$$

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

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

$$= \frac{3}{4}\left[x^2\right]_0^2 - \frac{1}{4}\left[x^3\right]_0^2$$

$$= \frac{3}{\cancel{4}}\left[\cancel{4}\right] - \frac{1}{\cancel{4}}\left[\overset{2}{\cancel{8}}\right]$$

$= 1$

$=$ It 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

Since, 1 of the probability is negative, it 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

Since $p(x) \geq 0 \quad \forall x$
and $\sum p(x) = 1$
$\therefore$ It is a p.m.f

(iii)

x	10	20	30	40	50
$p(x)$	0.2	0.3	0.2	0.6	0.2

Since $p(x) \geq 0 \quad \forall x$
and $\sum p(x) \neq 1$
$\therefore$ It is not a p.m.f

Q.3] Find $P(X \leq 2)$, $P(2 \leq x < 4)$, $P(\text{atleast } 4)$, $P(3 < x < 6)$

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

$\Rightarrow$ $P(X \leq 2)$

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

$= 0.1 + 0.1 + 0.2$

$= 0.4$

$P(2 \leq x < 4)$

$= P(2) + P(3)$

$= 0.2 + 0.2$

$= 0.4$

$P(\text{atleast } 4)$

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

$= 0.1 + 0.2 + 0.1$

$= 0.4$

$P(3 < x < 6)$

$= P(4) + P(5)$

$= 0.1 + 0.2$

$= 0.3$

PRACTICAL NO 3

Probability Distribution

1] find c.d.s of the following p.m.f and draw the graph.

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

$f(x) = 0$ if $x < 10$
$= 0.15$ $10 \leq x \leq 20$
$= 0.40$ $20 \leq x < 30$
$= 0.70$ $30 \leq x < 40$
$= 0.90$ $40 \leq x < 50$
$= 1.0$ $x \geq 50$

```
> x = c(10, 20, 30, 40, 50)
> prob = c(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")
```

Binomial Distribution

2] Suppose there are 12 mcq's in a test each question 3 option and only one of them is correct. Find the probability of having in 5 correct answer (2) Atmost 4 correct answers.

→ Given that

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

X = Total no of correct answers

$X \sim B(n, p)$

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

```
> n=12; p=1/5; q=4/5; x=5
> sum (n,p,q,x)
[1] 18

> dbinom (5,2,1/5)
[1] 0.05315022

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

3] There are 10 members in a committee, the probability of any member attending a meeting is 0.9. Find the probability

(i) 7 members attended (ii) Atleast 3 members attended
(iii) Atmost 6 members attended.

→ Given that

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

x = Total no of member attended

$x \sim B(n, p)$

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

```
> n=10; p=0.9; q=0.1
> dbinom (7,10,0.9)
[1] 0.0573956З
```

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

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

4] 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

→ (x) = c(1, 2, 3, 4, 5, 6)

```
> prob = c(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 = "probabilit
main = "graph of c.d.f", "s")
```

8-1-20

PRACTICAL NO 4

Binomial Distribution

1] Find the complete binomial distribution where $n = 5$ and $p = 0.1$

2] Find the probability of exactly 10 success in 100 trials where $p = 0.1$

3] X follows binomial distribution with $n = 12$, $p = 0.25$.
Find (i) $P(X = 5)$
(ii) $P(X \leq 5)$
(iii) $P(X > 7)$
(iv) $P(5 < X < 7)$

4] The probability of a salesman makes a sale to a customer is 0.15. Find the probability
(i) No sale for 10 customer
(ii) More than 3 sale in 20 customer

5] A student writes 5 mcq's. Each question has 4 options out of which one is correct. Calculate the probability for atleast 3 correct answers.

6] X follows binomial distribution with $n = 10$, $p = 0.4$
Plot the graph of p.m.f and c.d.f

* **Note:**

$P(X = x) = \text{dbinom}(x, n, p)$
$P(X \leq x) = \text{pbinom}(x, n, p)$
$P(X > x) = 1 - \text{pbinom}(x, n, p)$

To find the value of X for which the probability given as p_1, command is qbinom.

⇒ qbinom (p_1, n, p)

⇒ **Answers:**

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] x = 10 , n = 100 , p = 0.1

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

Scanned with CamScanner

* **Note:**

$P(X = x) =$ dbinom (x, n, p)
$P(X \leq x) =$ pbinom (x, n, p)
$P(X > x) = 1 -$ pbinom (x, n, p)

To find the value of X for which the probability is given as p_1, command is qbinom.

⇒ qbinom (p_1, n, p)

⇒ **Answers:**

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] $x = 10$, $n = 100$, $p = 0.1$

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

3] $n = 12$, $p = 0.25$

(i) $P(X=5)$ = dbinom (x, n, p)

```
> dbinom (5, 12, 0.25)
[1] 0.1032414
```

(ii) $P(X \leq 5)$ = pbinom (x, n, p)

```
> pbinom (5, 12, 0.25)
[1] 0.9455978
```

(iv) $P(5 < X < 7)$ = dbinom (x, n, p)

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

(iii) $P(X \ngtr 7)$

$= 1 - P(X \leq 7)$

```
> 1 - pbinom (7, 12, 0.25)
[1] 0.00278151
```

4] $p = 0.15$

(i) $n = 10$, $p = 0.15$, $x = 0$

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

(ii) $p = 0.15$, $n = 20$

$P(X>3) = 1 - P(X \leq 3)$

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

5] $n = 5$, $p = 1/4$

$P(X \geq 3)$
$= 1 - P(X \leq 2)$

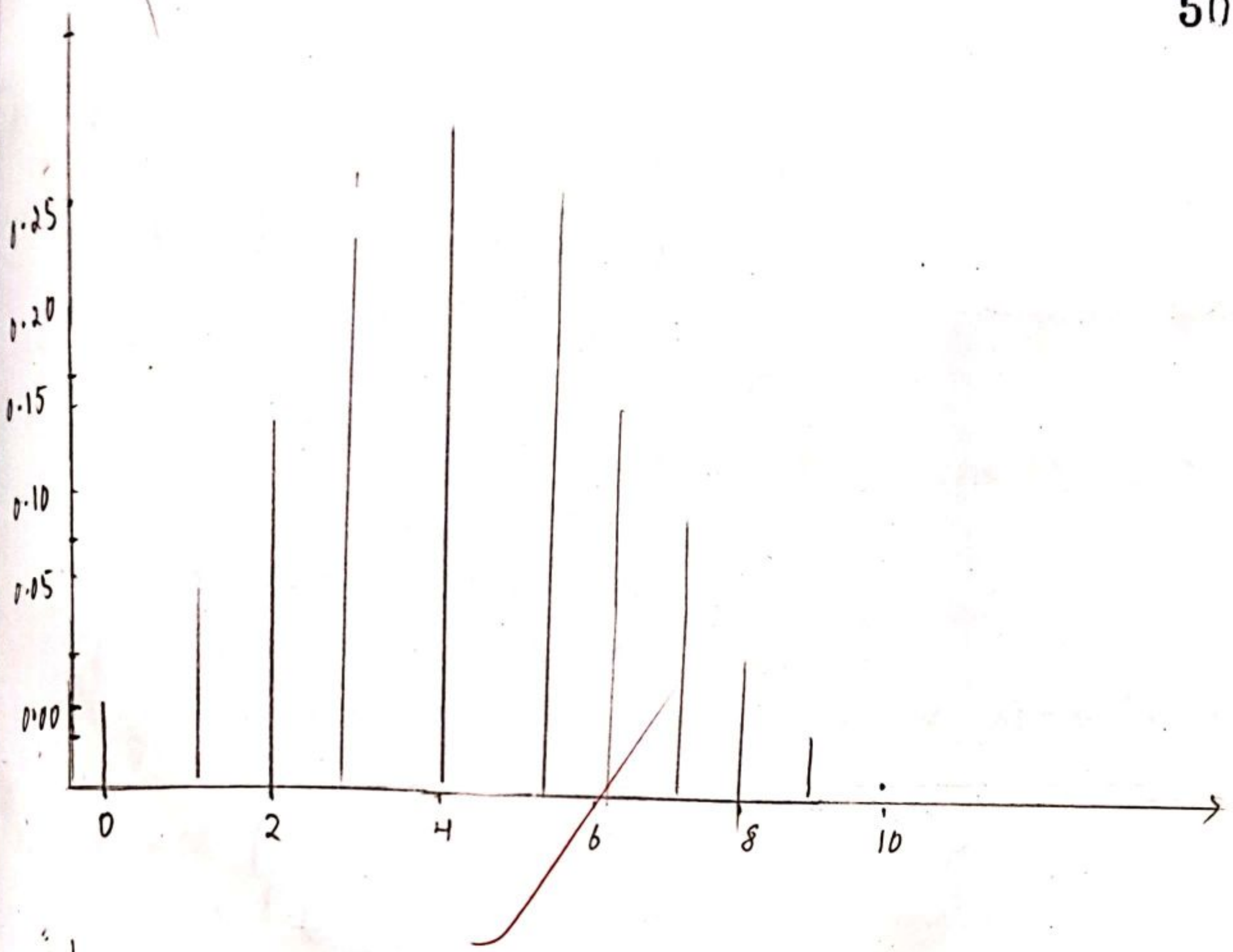
```
> 1- pbinom (3,5,1/4)
[1] 0.015625
```

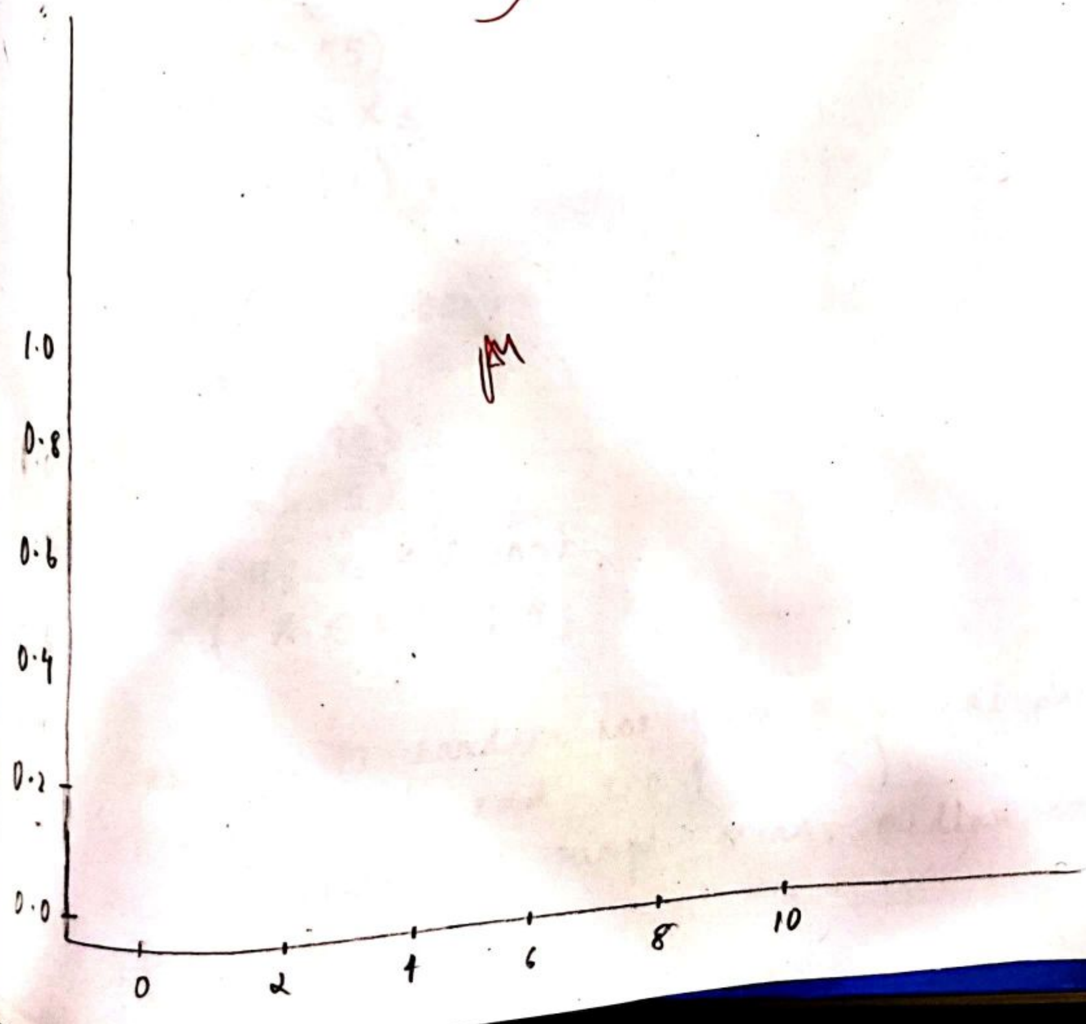
6] $n = 10$, $p = 0.4$

```
x = 0:n
prob = dbinom (x,n,p)
cumprob = pbinom (x,n,p)
d = dataframe ("x values" = x , "probability" = prob)
print(d)
```

```
> prob = dbinom (0:10, 10, 0.4)
> prob
> cumprob = pbinom (0:10, 10, 0.4)
> cumprob
> d = data.frame ("x values" = 0:10, "probability" = 0.4)
> print (d)
> plot (0:10, prob, "h")
> plot (0:10, cumprob, "s")
```

Scanned with CamScanner


0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10


1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10

PRACTICAL NO 5

* NOTE :

$P(X = x) = \text{dnorm}(x, \mu, \sigma)$
$P(X \leq x) = \text{pnorm}(x, \mu, \sigma)$
$P(X > x) = 1 - \text{pnorm}(x, \mu, \sigma)$
$P(x_1 < X < x_2) = \text{pnorm}(x_2, \mu, \sigma) - \text{pnorm}(x_1, \mu, \sigma)$

To find the value of k so that
$P(X \leq K) = P_1$; $\text{qnorm}(p_1, \mu, \sigma)$
To generate n random no the command is $\text{rnorm}(n, \mu, \sigma)$

1] $X \sim N(\mu = 50, \sigma^2 = 100)$
find (i) $P(x \leq 40)$
(ii) $P(X > 55)$
(iii) $P(42 \leq X \leq 60)$
(iv) $P(X \leq k) = 0.7$; $k = ?$

2] $X \sim N(\mu = 100, \sigma^2 = 36)$
(i) $P(X \leq 110)$
(ii) $P(X \leq 95)$
(iii) $P(x > 115)$
(iv) $P(95 \leq X \leq 105)$
(v) $P(X \leq k) = 0.4$; $k = ?$

3] Generate 10 random nos from a normal distribution with mean ($\mu = 60$) and S.D ($\sigma = 5$).
Also calculate the sample mean, median, variance and standard deviation.

5

4] Draw the graph of standard normal distribution.

Answers:

1] (i)>a = pnorm (40, 50, 10)
cat ("P (X ≤ 40) is = ", a)

[1] 0.1586553
P (X < 40) is 0.1586553>
(ii) > b = 1-pnorm (55, 50, 10)
cat ("P (X > 55) is = ", b)
P (X > 55) is 0.03085375 >
[1] 0.3085375

(iii) > c = pnorm (60, 50, 10) - pnorm (42, 50, 10)
~~cat~~.
[1] 0.6294893
P (X < k) = 0.7, k is 55.24401>
(iv) > d = qnorm (0.7, 50, 10)
[1] 55.24401

Scanned with CamScanner

2] (i) > a = pnorm (110, 100, 6)

[1] 0.9522096

(ii) > b = pnorm (95, 100, 6)
[1] 0.2023284

(iii) > c = 1 - pnorm (115, 100, 6)
[1] 0.006209665

(iv) > d = pnorm (105, 100, 6) - pnorm (95, 100, 6)
[1] 0.5953432

(v) > e = qnorm (0.4, 100, 6)
[1] 98.47992

3] n = 10, μ = 60, σ = 5
> x = rnorm (10, 60, 5)

[1] 64.30801 56.28556 57.12444 62.66236 57.07212 52.90556 63.04319
59.37558 59.78091 56.45464

> am = mean (x)
> am
[1] 58.30124

> me = median (x)
> me
[1] 58.22385

```
> n=10
> variance = (n-1)* var(x)/n
> variance
[1] 17.05339

> sd = sqrt(variance)
> sd
[1] 4.129575
```

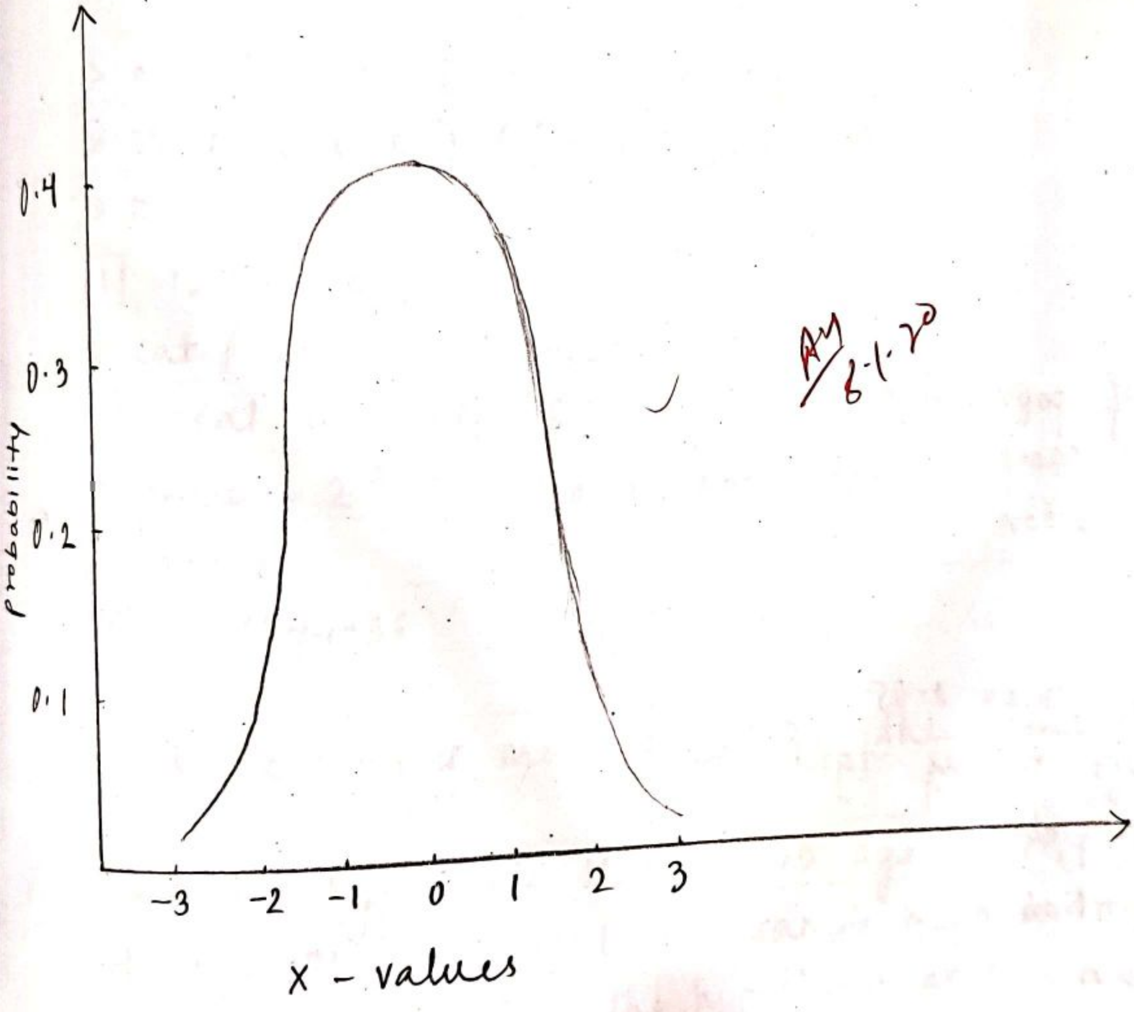
4.]
```
> x = seq(-3, 3, by = 0.1)
> x
[1] -3.0 -2.9 -2.8 -2.7 -2.6 -2.5 -2.4 -2.3 -2.2 -2.1 -2.0
-1.9 -1.8 -1.7 -1.6 -1.5 -1.4 -1.3 -1.2 -1.1 -1.0 -0.9
-0.8 -0.7 -0.6 -0.5 -0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3
0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7
1.8 1.9 2.0 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0

> y = dnorm(x)

> plot(x, y, xlab="x values", ylab="probability", main="standard
normal graph")
```

probability 0.4 0.3 0.2 0.1

STANDARD NORMAL GRAPH



AY 8-1-20

Mo = mean of population
mn = " " sample
sd = std deviation
n = total

if ans is more than 0.05
then accept Ho as right ans
or else it is false
P = popln proportion
p = sample proportion

PRACTICAL NO 6

Topic: Z distribution

1] Test the hypothesis (H_0) $H_0 : \mu = 10$ against $H_1 : \mu \neq 10$. A sample of size 400 is selected which gives the mean 10.2 and standard deviation 2.25. Test the hypothesis at 5% level of significance

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

∵ The answer is more than 0.05, ~~the value of H_0 is correct.~~ H_0 is accepted.

2] Test the hypothesis $H_0 : \mu = 75$ against $H_1 : \mu \neq 75$. A sample of size 100 is selected and sample mean is 80 with standard deviation of 3. Test the hypothesis at 5% level of significance.

```
> mo = 75; mx = 80; sd = 3; n = 100
> zcal = (mx-mo)/(sd/(sqrt(n)))
> zcal
[1] 16.66667
> cat ("zcal is =", zcal)
[1] zcal is = 16.6667
```

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

Since, the value of ~~H~~ is less than 0.05, the value H₀ is incorrect.

3] Test the hypothesis $H_0 : \mu = 25$ against $H_1 : \mu \neq 25$ at 5% l. of significance. Following sample of 30 is selected.
S: 20, 24, 27, 35, 30, 46, 26, 27, 10, 20, 30, 37, 35, 21, 22, 24, 25, 26, 27, 28, 29, 30, 39, 27, 15, 19, 22, 20, 18

```
> x = c(          )
> mx = mean (x)
> mx
[1] 26.06667
> n = length (x)
> n
[1] 30
> Variance = (n-1)* var(x)/n
> variance
[1] 52.99556
> sd = sqrt (variance)
> sd
[1] 7.279805
> mo = 25; mx = 26.07; sd = 7.28; n = 30
```

```
> zcal = (mx - mo)/(sd/(sqrt(n)))
> zcal
[1] 0.8050318
> cat("zcal is =", zcal)
[1] zcal is = 0.8050318
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.4208013
```

∵ The value is ~~less~~ more than 0.05, the value of H_0 is correct

4) 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 gives the correct proportion or not

```
> P = 0.2
> Q = 1-P
> Q
[1] 0.8
> p = 50/400
> p
[1] 0.125
> n = 400
> Zcal = (p-P)/(sqrt(P*Q/n))
> zcal
[1] -3.75
```

∵ The value is less than 0.05, the value of H_0 is incorrect

Scanned with CamScanner

57

5] Test the hypothesis $H_0 : P = 0.5$ against $H_1 : P \neq 0.5$. A sam of 200 is selected and the sample proportion is calculated $p = 0.56$. Test ~~the~~ the hypothesis at 2% level of significance

```
> P=0.5
> Q=1-P
> Q=0.5
> p=0.56
> n=200
> zcal = (p-P)/(sqrt(P*Q/n))
[1] 1.697056
> pvalue = 2*(1-pnorm(abs(zcal)))
[1] 0.089688
```

∵ The value is more than 0.05, the value of Ho is correct

22-4-20

PRACTICAL NO 7

Topic : Large Sample Tests

1] A study of noise level in two hospital is calculated below. Test the hypothesis that the noise level into hospital. Test the hypothesis that the noise level in 2 hospitals are same or not

~~Hos A~~	Hos A	Hos B
No of sample obs	84	34
Mean	61	59
S.D	7	8

Ho : The noise levels are same

```
> n1 = 84
> n2 = 34
> mx = 61
> my = 59
> sdx = 7
> Sdy = 8
> z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))
  z
[1] 1.273682
> cat("z calculated is = ", 2)
[1] z calculated is = 2
> pvalue = 2* (1- pnorm (abs (z)))
[1] ~~0.0458~~ 0.1291
```

Since pvalue is 0.1291, we reject Ho at 5% level of significance

2] Two random samples of size 1000 & 2000 are drawn from populations with a means 67.5 and 68 respectively and with the same S.D of 2.5. Test the hypothesis of the mean of 2 populations are equal.

Ho = 2 populations 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
[1] -5.163978
> cat ("z calculated is =", z)
[1] z calculated is = 2
> pvalue=2* (1-pnorm (abs(2)))
>pvalue
[1] 2.417564e-07
```

Since pvalue is 2.4175e-07, we reject Ho at 5% level of significance.

3] In a first year class (FYBsc), 20% of a random sample of 400 students had defective eyesight. In SY class, 15.5% of 500 sample had same defect. Is the difference of proportion is same?

Ho = 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
> q
[1] 0.825
> z = (p1-p2)/ sqrt (p*q*(1/n1+1/n2))
> z
[1] 1.76547
> cat ("z calculated is =", z)
[1] z calculated is 1.76547
> pvalue = 2* (1-pnorm(abs(z)))
> pvalue
[1] 0.07748487
```

Since pvalue is 0.077, we accept ~~reject~~ Ho at 5% level of significance.

:a

4.] From each of the box of the apples, a sample size of 200 collected. It is found that, there are 44 bad apples in the 1st sample and 30 bad apples in the 2nd sample. Test the hypothesis that 2 boxes are equivalent in terms of no. bad apples.

H_0 = The proportion of boxes is same.

```
> 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] 1.8027
> cat("z calculated is = ", z)
z calculated is 1.8027
> pvalue = 2 * (1- pnorm (abs(z)))
> pvalue
[1] 0.0714
```

Since pvalue is 0.0714, we ~~reject~~ accept at 5% level of signifi

5] In a MA class, out of a sample of 60, Mean height is 63.5 inch with a S.D 2.5. In a MCom class, out of 50 students, mean height 69.5 inches with a S.D of 2.5. Test the hypothesis that the mean of M.A and MCom class are same.

H0 = Mean height of MA and MCom students are equal.

```
>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] -12.5335
>cat("z calculated is =",z)
z calculated is = -12.5335
>pvalue = 2*(1-pnorm(abs(z)))
>pvalue
[1] 0.04550
```

Since pvalue is 0.04550, we reject at 5% level of significance

AM
5.2.20

PRACTICAL NO 8

Topic: Small Sample Test

1.] The tens are selected and height are found to be 63, 68, 69, 71, 64, 70, 72 cms. Test hypothesis that mean height 66 cm or not at 1.

H_0 : Mean = 66 cms

```
> mean = 66
> x = c (63,63,68,69,71,71,72)
> t. test(x)
  One Sample t. test
 data : x
t = 47.94 , df = 6 , pvalue = 5.22e-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
```

∴ pvalue < 0.01 is rejected on h_0 in 1% Level of Significance.

2] Two random sample was drawn from two different popula

Sample 1 : 8, 10, 12, 11, 16, 15, 18, 7

Sample 2 : 20, 15, 18, 9, 8, 10, 11, 12

Test the hypothesis that there is no difference, between the popul mean at 5% loss.

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)
```

Welch Two Sample t. test

data : x & y

t = -0.36247 , df = 13.837 , P-value = 0.7225

Alternative hypothesis : True difference in mean is not equal

-5.192719 3.692719

Sample estimates

Mean of x Mean of y

12.125 12.875

P-value < 0.01 is accepted in H_0 on 1% Level of Significance

3] following are the weights of 10 People

Before = (100, 125, 95, 96, 98, 102, 115, 104, 109, 110)

After = (95, 80, 95, 98, 90, 100, 110, 85, 100, 101)

H_0 = The Diet Program is not affective

```
> x = c ( 100, 125, 95, 96 , 98, 102, 115 , 104, 109, 110)
> y = c (95, 80, 95, 98, 90, 100, 110, 85, 100, 101 )
> t. test ( x, y, Paired = T, alternative = "less")
```

Paired t. test

data: x and y

t = 2.3215 , df = a , pvalue = 0.9773

Alternative hypothesis : True difference in means is less than 0

95% confidence interval

-Inf 17.89635

Sample estimates:

Mean of the differences 10

4] Marks before and after a training program is given below:

Before = 20, 25, 32, 28, 27, 36, 35, 25

After = 30, 35, 32, 37, 37, 40, 40, 23

Test the hypothesis that training program is effective or not.

H_0 = The training program is not effective

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

Paired t-test

Data: x and y

t = -3.3859, df = 7, p-value = 0.9992

Alternative hypothesis: True difference in means is greater than

percent confidence interval.

-8.967399 Inf

Sample estimates:

Mean of the difference -5.75

5] Two random sample were drawn from two normal population &

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 whether the population have same variance at 5% level of

H₀: Variance of the population are equal

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

F test to compare two variances

Data: X and y

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:

6.1833662 3.0360393

Sample estimates:

Ratio of variances 0.7068567

6] The A.P of sample 100 observations is 52 if S.D is 7 test the hypothesis that the population mean 55 or not at 5% of level of significance.

H_0: Population mean = 55

```
> n=100
> mx =52
> mo =55
> sd=7
> zcal=(mx-mo)/(sd/sqrt(n))
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 1.82153e-0.5
```

69

PRACTICAL NO 9

Chi Square and ANOVA

Use the following data to test whether the of the name depen upon the child condition or not.

		Condition of home clean	Dirty
Cond of child	clean	70	50
	fairly dirty	80	20
	dirty	35	45

→ H_0 Condition of the name and child are independent

```
> x = c(70, 80, 35, 50, 20, 45)
> m = 3
> n = 2
> y = Matrix(x, nrow = m, ncol = n)

      [,1]  [,2]
[1,]   70    50
[2,]   80    20
[3,]   35    45

> pvalue = chisq.test(y)

> pvalue
```

Pearson's chi-squared test

data = y

x = squared - 25.646, dr - 2, pvalue = 2.698e-0.6

Pvalue is less than 0.05. we reject H_0 at 5% LOS

2] Table below shows the relation between the performances of mathematics and computer of CS students.

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

H_0: performance between maths and computer are independence

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

m = 3

n = 3

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

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

pvalue. chiq. test (y)
pvalue

data y
x-squared = 145.75, df = 4, p.value < 2.2e-16

3] Varieties Observation

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

Mean A, B, C, D

```
> 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))
> oneway.test(values~ind, data=d, var equal = T)
anova = aov(values~ind, data=d)
F = 11.735, num df = 3, denom.df = 9, p-value = 0.00183
summary(anova)
```

Non-Paramatric Test

following are the amounts of sulphur oxide

Data:

17, 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 hypothesis that population median is 21.5 against the alternative it is less than 21.5

H_0: population median = 21.5

H_1: 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)
> x
[1]
> m = 21.5
> sp = length (x[x>m])
> sn = length (x[x<m])
> n = sp + sn
> n
[1] 20
> pv = pbinom (sp, n, 0.5)
> pv
[1] 0.411
```

NOTE: If the alternative is greater than median,

pv = pbinom (sn, n, 0.5)

2] For the observations
12, 19, 31, 28, 43, 40, 55, 49, 70, 63
Apply sign test to test population median is 25 against the alternative is more than 25.

```
> x = c( 12, 19, 31, 28, 43, 40, 55, 49, 70, 63)
>M = 21.5 25
>sp = length ( x[x > M])
>sn = length ( x[x < M])
>n = sp + sn
> pv = pbinom (sn, n, 0.5)
> pv
[1] 0.054
```

3] 60, 65, 63, 89, 61, 71, 58, 51, 48, 66
Test the hypothesis using wilcoxon sign test.
For testing the hypothesis that the median is 60 against 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
> sp = length ( x[ x > m])
> sn = length (x[x < m])
> n = sp + sn
> pv = pbinom ([illegible], n, 0.5)
[1] 0.253
```

> wilcox.test (x, alter = "greater", MU = 60)
V=29, p-Value = 0.2386
alternative hypothesis: true location is greater than 60

If the alternative is less,
wilcox.test (x, alter = "less", mu = -)

If the alternative is not equal to
wilcox.test (x, alter = "2.sided", MU = 6)

9] Using wilcoxon test, test the hypothesis where the median is 12 against the alternative is less than 12.

Data:

12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20

```
> x = c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20)
> m = 12
> sp = length (x [x > m])
> sn = length (x [x < m])
> n = sn + sp
> pv = pbinom (sn, n, 0.5)
> pv
[1]
> wilcox.test (x, alter = "less", MU = 12)
V = 25, p-value = 0.775
Alternative hypothesis: true location is greater than 12
```