THAKUR COLLEGE OF SCIENCE & COMMERCE

NAAC
Accredited
with Grade "A"
(3rd Cycle)

THAKUR
TRUSTS

ISO
9001 : 2015
Certified

Degree College

Computer Journal

CERTIFICATE

SEMESTER Sem - II UID No. ____________

Class FYBSC (C.S) Roll No. 1810 Year 2019-20

This is to certify that the work entered in this journal is the work of Mst. / ~~Ms.~~ Ayush Kumar Mishra

who has worked for the year 2019-20 in the Computer Laboratory.

Teacher In-Charge

Head of Department

Date : 9.3.20

Examiner

★★ INDEX ★★

Scanned with CamScanner

Practical - 1.

Basics of R - Software

1.] R is a Software of data analysis and statistical computing.

2.] It is a Software by which effective data handling and outcome storage is possible

3.] It is capable of graphical display.

4.] It is a free Software

Q1.] $2^2 + |-5| + 4 \times 5 + 6/5$

```
> 2^2 + abs(-5) + 4*5 + 6/5
[1] 30.2
```

Q2.] $x = 20$

$y = 2x$

$z = 2x + y$

sqrt(z)

```
-> > x = 20
   > y = 2*x
   > z = x + y
   > sqrt(z)
   [1] 7.745967
```

Q3.] $x = 10$

$y = 15$

$z = 5$

c.] $\sqrt{xy2}$

d.] round $\sqrt{xy2}$

>x=10

>y=15

>z=5

>x+y+z

[1]30

>x*y*z

[1](150)

>sqrt(x*y*z)

[1](27.38613)

A Vector is R software is denoted by the syntax

c(2,3,5,7)^2

>x=c(2,3,5,7)

>x^2

[1] 4 9 25 49

c(2,3,5,7) ^c (2,3)

4 27 25 343

c(2, 3, 5, 7, 9) ^c (2,3)

Warning message

c(1,2,3,4,5,6) ^c(2,3,4)

[1] 1 8 9 64 25 216

c(4, 6, 8, 10) * 3

12 18 24 30

c(4,6,8,10) * c(-1,-2,-3,-4)
[1] -4 -12 -24 -40

c[2,3,5,7] * c[-2,-3]

[1] 0.2500000 0.03703703

c(2,3,5,7)+10

12 13 15 17

c(2,3,5,7) + c(-2,-3,-1,0)

c(2,3,5,7)/2

[1] 1.0 1.5 2.5 3.5

find the sum of product maximum and minimum for values of

x = c(2,8,4,9,11,10,7,6)^2

> sum(x)

[1] 455

> Product(x)

[1] 4425974784

> max(x)

[1] 121

> min(x)

[1] 4

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

$$\underset{x}{\begin{bmatrix} 4 & 7 & 4 \\ 5 & 8 & 0 \\ 6 & 9 & 2 \end{bmatrix}} + \underset{y}{\begin{bmatrix} 6 & 11 & 9 \\ 4 & 12 & 7 \\ 5 & 8 & 4 \end{bmatrix}}$$

```
x <- matrix (nrow = 3, ncol = 3, data = c (4,5,6,7,8,
                                              9,4,0,2)
y <- matrix (nrow = 3, ncol = 3, data = c (6,4,5,11,12,8,
                                              9,7,4)
```

\> x

	[,1]	[,2]	[,3]
[1,]	4	7	4
[2,]	5	8	0
[3,]	6	9	2

\> y

	[,1]	[,2]	[,3]
[1,]	6	11	9
[2,]	4	12	2
[3,]	5	8	4

, 6, 7, 8]

\> x * 2

	[,1]	[,2]	[,3]
[1,]	8	14	8
[2,]	10	16	0
[3,]	12	18	4

> y * 2

	[,1]	[,2]	[,3]
[1,]	12	22	18
[2,]	8	24	14
[3,]	10	16	8

> x + y

	[,1]	[,2]	[,3]
[1,]	10	18	13
[2,]	9	20	7
[3,]	11	17	6

> x * y

	[,1]	[,2]	[,3]
[1,]	24	77	36
[2,]	20	96	0
[3,]	30	72	8

RM

Practical - 2
Binomial Distribution

n = total no. of trials
p = p (success)
q = p (failure)
x = no. of success out of n

$p(x) = {}^nC_x \cdot p^x \cdot q^x, \ x = 0, 1 \ldots\ldots n$

$E(x) = np$

$V(x) = npq$

dbinom (x, y, p) — p(x)

n * p — E(x)

n * p * q — V(x)

Pbinom (x, n, p) — $p(x \le x)$, $p(x > x) = 1 -$ pbinom (x+n)

Q1] Toss a coin 30 times with probability (Head = 0.6)
let x be the no. of heads
Find the probability of:
(i) 7 Heads
(ii) 4 Heads
(iii) At most 4 Heads
(iv) At least 6 Heads
(v) No heads
(vi) All heads
(vii) expectation & Variance

```
> n = 10
> p = 0.6
> q = 0.4
> dbinom (7, 10, 0.6)
[1] 0.1449908
> dbinom (4, 10, 0.6)
[1] 0.114767
> pbinom (4, 10, 0.6)
[1] 0.1662386
> 1 - pbinom (6, 10, 0.6)
[1] 0.3822806
> dbinom (0, 10, 0.6)
[1] 0.0001048576
> dbinom (10, 10, 0.6)
[1] 0.006046618
> exp = 10 * 0.6
> exp
[1] 6
> Var = 10 * 0.6 * 0.4
> Var
> 2.4
```

Scanned with CamScanner

] there are 12 mcq's in an english question paper question has 5 answers and only 1 of them (i) correct find the Probability of having
(i) 4 correct answer
(ii) atmost 4 correct answer
(iii) atleast 3 correct answer

n = 12 , p = 1/4

(i) P(X = 4)

> dbinom (4, 12, 1/5)

[1] 0.1228756

(ii) P(X ≤ 4)

> pbinom (4, 12, 1/5)

[1] 0.9274445

(iii) P(X ≥ 3) = P(X > 2)

> 1 - pbinom (2, 12, 1/5)

[1] 0.4416543

Q3] Find complete binomial distribution where n = 5
p = 0.1

> dbinom (0, 5, 0.1)

[1] 0.59049

> dbinom (1, 5, 0.1)

[1] 0.32805

> dbinom (2, 5, 0.1)

[1] 0.729

> dbinom (3, 5, 0.1)

[1] 0.0081

> dbinom (4, 5, 0.1)

> dbinom (5, 5, 0.1)

[1] 1e-05

ii] Find Probability or exactly 10 sucess outcome 100

trials = 0.1

> dbinom (10, 100, 0.1)

[1] 0.1318653

] x follows binomial distribution with n=12,

p=0.25

Find (i) $P(x \leq 5)$ (ii) $P(x > 7)$

(iii) $P(5 < x < 7)$

(i) $P(x \leq 5)$

> pbinom (5, 12, 0.25)

[1] 0.9455978

(ii) 1 - pbinom (7, 12, 0.25)

[1] 0.00278151

(iii) $P(5 < x < 7)$

[1] 0.~~8~~4 014 94 J.

(IV)

) There are 10 members. In a commitee probability of any member attending a meeting 0.9 words Probability that 7 or more member will present in metting

n = 10, x = 7, P = 0.9

> 1 - Pbinom (6, 10, 0.9)

[1] 0.9872048

Q.7] A sellsman has 20% probability sell of a typical day he will meet 50 customer what minimum no of he will made with 88.1 probability

:> n=50, p=0.88

> qbinom (p, n, p)

> qbinom (0.88, 30, 0.2)

[1] 9

Q8] n=10, P=0.6 find the binomial distribution plot the graph of p.m.f & c.d.f.

> n = 10; p = 0.6

> x = 0:n

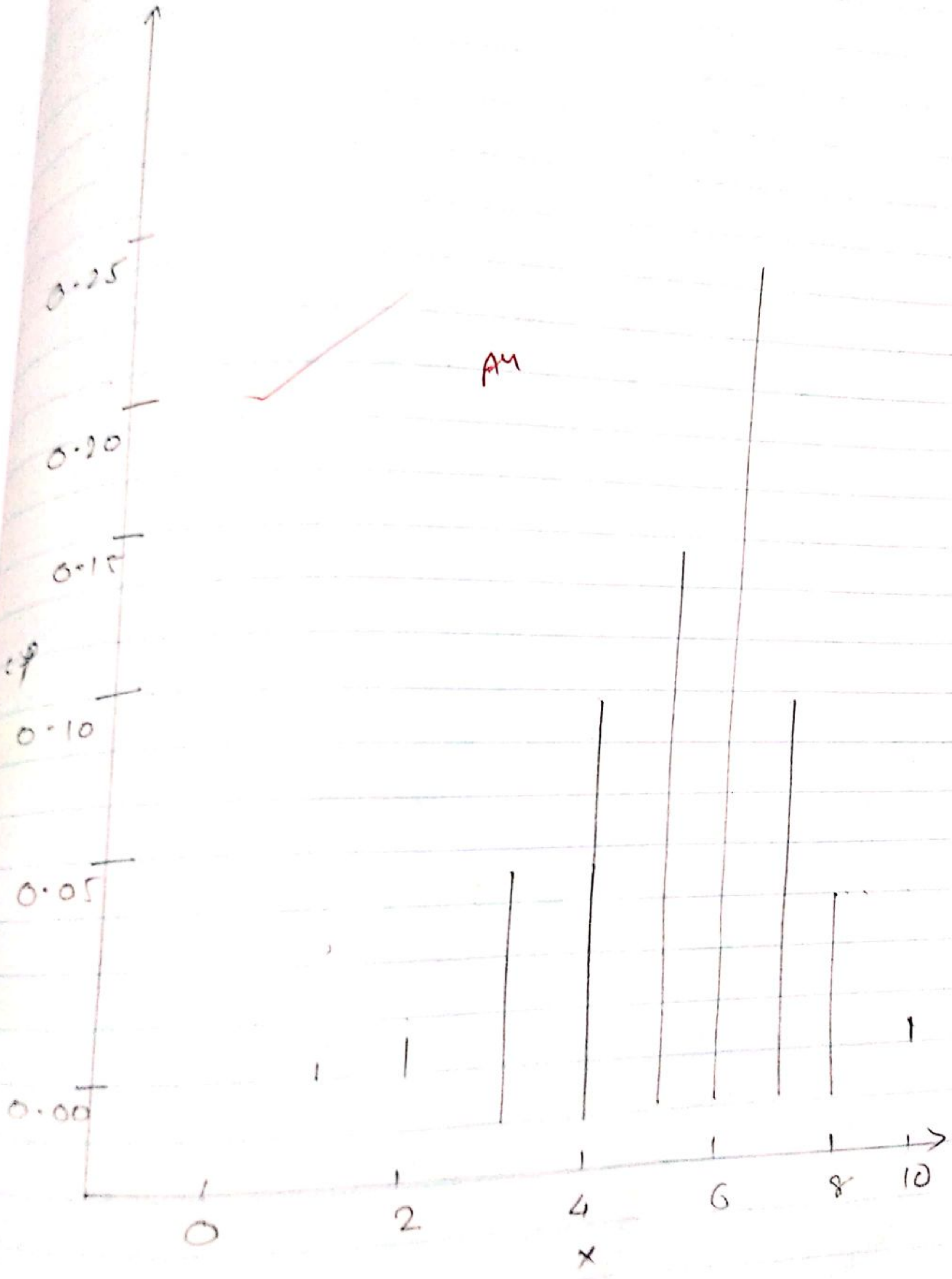
> bp = dbinom (x, n, p)

> d = data.frame("x-values"=x, "probability"= bp)

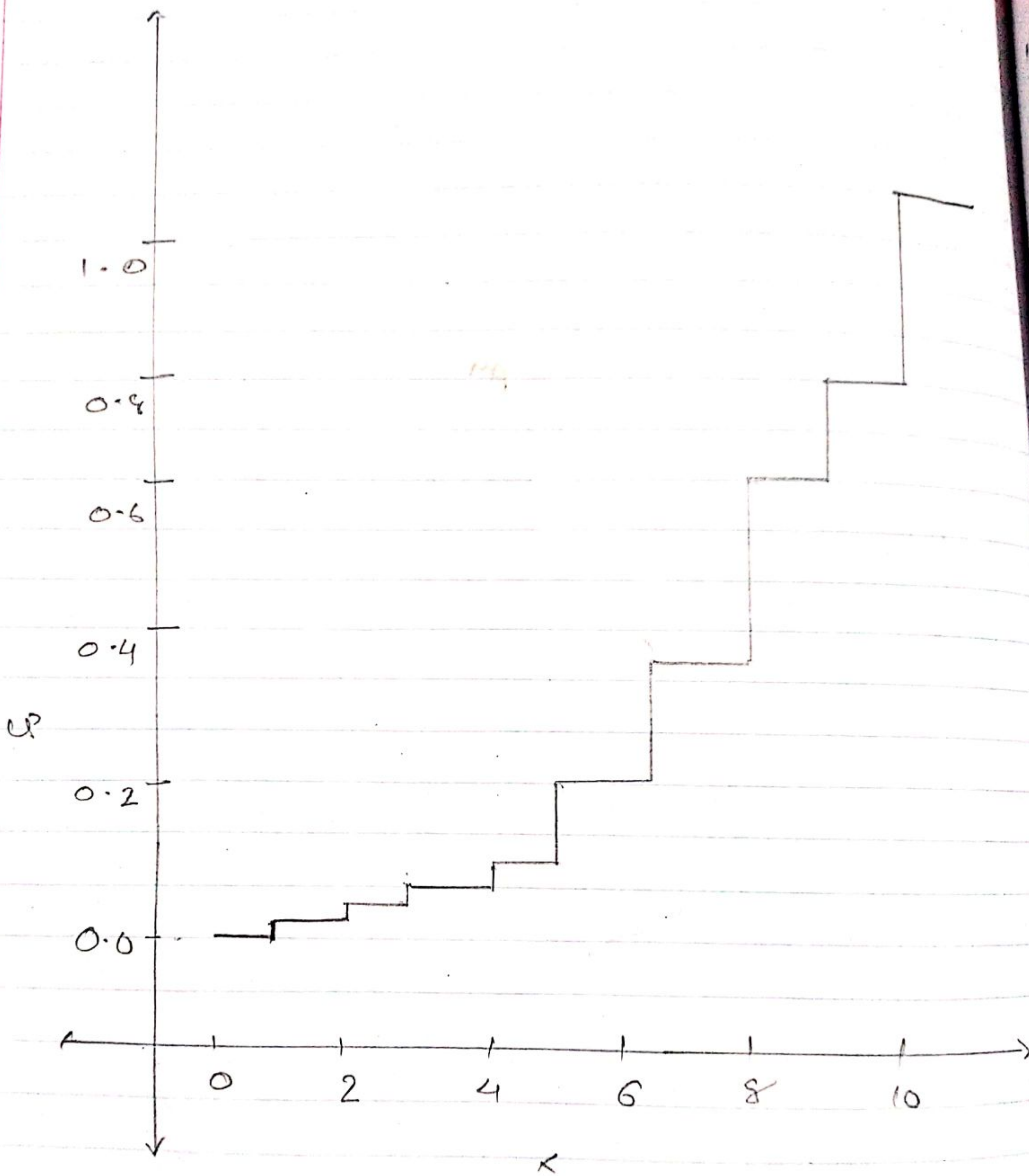
> print (d)

	x.values	probability
1	0	0.0001048576
2	1	0.0015728640
3	2	0.0106168320
4	3	0.0424673280
5	4	0.1114767360
6	5	0.2006581248
7	6	0.2508226568
8	7	0.2149908480
9	8	0.1704223520
10	9	0.0403107340
11	10	0.0060 4641767

```
> plot (x, bp, "s")
> cp = pbinom (x, n, p)
> plot (x, cp, "s")
```



Scanned with CamScanner


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

Practical no: 3

Check the following are pmf (Probability mass function) or not

x	1	2	3	4	5
$p(x)$	0.2	0.5	-0.5	0.4	0.4

Since the Probability is negative.
∴ It is not a pmf.

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

The condition for pmf is

① $0.5\, P(x) \leq 1$ — ①

② $\Sigma P x = 1$ — ②

```
> prob = c(0.3, 0.2, 0.3, 0.1, 0.1)
> sum(prob)
[1] 1
```

Since both the condition are satisfied it is a p.m.f

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

The condition to check pmf are:

① $0 \leq P(x) \leq 1$

② $\Sigma P(x) = 1$

```
> prob = c(0.4, 0.2, 0.3, 0.2, 0.1)
> sum(Prob)
[1] 1.2
```

Here the second condition does not satisfy
∴ It is not a p.m.f.

Q.2] Following is a p.m.f of x. Find mean & variance.

x	1	2	3	4	5
P(x)	0.1	0.15	0.2	0.3	0.25

x	P(x)	x·P(x)	x^2 P(x)
1	0.1	0.1	0.1
2	0.15	0.3	0.6
3	0.2	0.6	1.8
4	0.3	1.2	4.8
5	0.25	1.25	6.25
		3.45	13.55

```
→ >x = c(1,2,3,4,5)
>prob = c(0.1, 0.15, 0.2, 0.3, 0.25)
>a = x * prob
>mean = sum(a)
>mean
[1] 3.45
>b = (x^2) * prob
>b
[1] 0.10 0.60 1.80 4.80 6.25
>var = sum(b) - mean^2
>var
[1] 1.6475
```

3] following is a p.m.f of x. find mean and Variance

x	5	10	15	20	25
p(x)	0.1	0.3	0.2	0.25	0.15


```
> x = c (5, 10, 15, 20, 25)
> prob = c( 0.1, 0.3, 0.2, 0.25, 0.15)
> a = x * prob
> a
[1] 0.50   3.00   3.00   5.00   3.75
> mean = sum (a)
> mean
[1] 15.25
> b = (x^2) * prob
> b
[1] 2.50   30.00   45.00   100.00
> var = sum (b) - mean^2
> var
[1] 38.6875
```

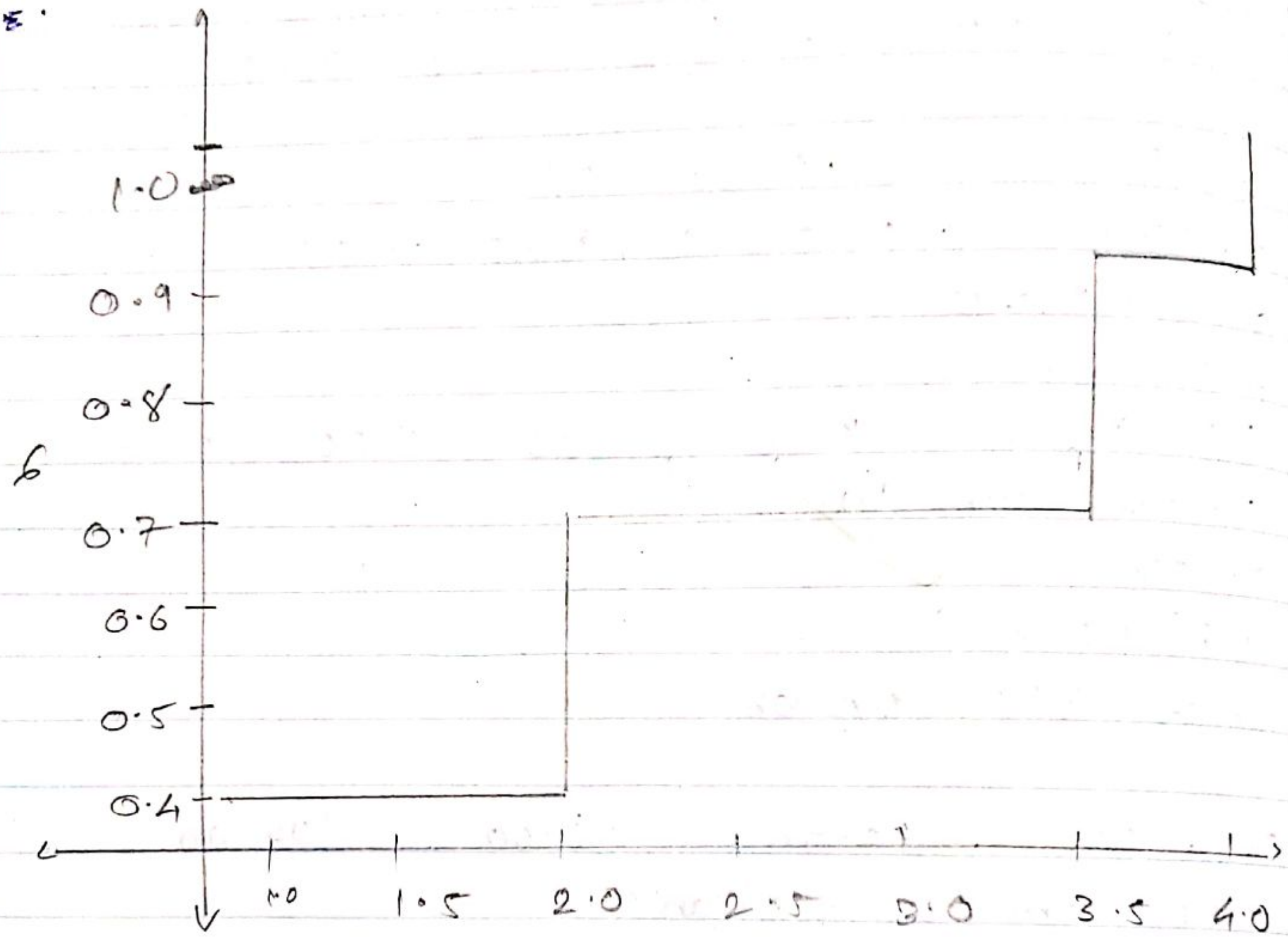
4] Find cdf of the following pmf and draw the graph of cdf.

1]

x	1	2	3	4
P(x)	0.4	0.3	0.2	0.1

```
> x = c (1, 2, 3, 4)
> prob = c ( 0.4, 0.3, 0.2, 0.1)
> a = cumsum (prob)
```

> a
[1] 0.4 0.7 0.9 1.0
> plot (x, a, "s")



$F(x) = 0 \qquad x < 1$

$= 0.4 \qquad 1 \le x < 2$

$= 0.7 \qquad 2 \le x < 3$

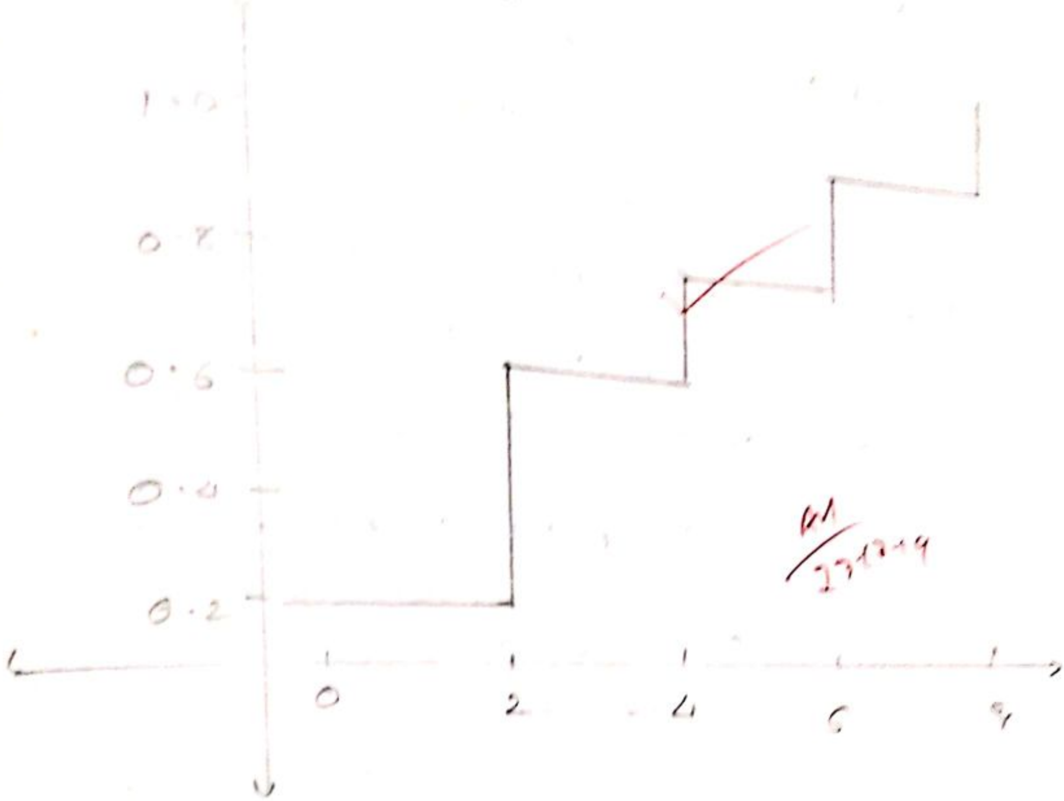
$= 0.9 \qquad 3 \le x < 4$

$= 1.0 \qquad x \ge 4$

Q.]

x	0	2	4	6	8
$P(x)$	0.2	0.3	0.2	0.2	0.1

```
> x = c(0,2,4,6,8)
> prob = c(0.2, 0.3, 0.2, 0.2, 0.1)
> fd = cumsum(prob)
> fd
[1] 0.2  0.5  0.7
> plot (x, fd, "s")
```



AA 23/12/19

$F(x) = 0$	$x < 0$
$= 0.2$	$0 \leq x < 2$
$= 0.5$	$2 \leq x < 4$
$= 0.7$	$4 \leq x < 6$
$= 0.9$	$6 \leq x < 8$
$= 1.0$	$x \geq 8$

Practical No. 4

Practise:→

Q.] X follows binomial distribution with $P = 0.6$ $q = 0.4$.
find

a] $P(X=7)$ b] $P(X \leq 3)$

→ $P(X=7) = \binom{8}{7} (0.6)^7 (0.4)^{8-7}$

$= {}^8C_7 \times 0.2799 \times 0.4$

$= 0.8957.$

$P(X \leq 3) = P(0) + P(1) + P(2) + P(3)$

$= {}^8C_0 (0.6)^0 (0.4)^8 + {}^8C_1 (0.6)^1 (0.4)^7$

$+ {}^8C_2 (0.6)^2 (0.4)^6 + {}^8C_3 (0.6)^3 (0.4)$

$= 1 \times 1 \times 0.00065536 \times 8 \times 0.6 \times 0.001638$

$+ 428 \times 0.36 \times 0.0004096 +$

$56 \times 0.216 \times 0.101024$

$= 0.1736704$

$P(X = 2 \text{ or } 3) = P(2) + P(3)$

$= {}^8C_2 (0.6)^2 (0.4)^6 + {}^8C_3 (0.6)^3 (0.4)$

$= 2^8 \times 0.36 \times 0.004096 + 56 \times$

0.216×0.01024

$= 0.04128768 + 0.12386304$

$= 0.16515012$

Scanned with CamScanner

Practical No. 5

Normal Distribution →

Normal Distribution is an example of continous probability distribution

$x \sim N(\mu, \sigma^2)$

a] $P(X = x)$ dnorm (x, μ, σ)

b] $P(X < x)$ = pnorm (x, μ, σ)

c] $P(X > x)$ = 1 - pnorm (x, μ, σ)

d] To find the value of k so that the command is given (p, μ, σ)

$k \quad P(X \leq k) = p$

qnorm (p, μ, σ)

e] To generate a random sample of size n,

rnorm (n, μ, σ)

] A random variable $x \sim N(10, 2)$. find

1.] $P(X \leq 7)$ 2.] $P(X > 12)$ 3.] $P(5 \leq x \leq 12)$

4.] $P(X < k) = 0.4$ with $\mu = 10$, $\sigma = 2$

] $x \sim N(100, 36)$ $\sigma = \text{sqrt}(36)$

1.] $P(X \leq 110)$ 2.] $P(X > 105)$ 3.] $P(X \leq 92)$

4.] $P(95 \leq X \leq 110)$ 5.] $P(X \geq k) = 0.9$

] Generate 10 random sample and find the sample mean, median and variance, standard deviation

```
>m= 10
>s=2
>p1= pnorm (7,m, s)
> cat ("p (x<=7) is =", p1)
 p (x<=7) is = 0.0668072 >
>p2 =1- pnorm (12,m,s)
>cat ("p(x>12) is =", p2)
 p(x>12) is = 0.1586553 >
>p3= pnorm(12,m,s) - pnorm (5,m,s)
>cat ("p(5<=x<=12) is =", p3)
 p (5<=x<=12) is = 0.835135
>k = qnorm (0.4, m,s)
> cat ("p (x<k)= 0.4 , k is =", k)
 p (x<k) = 0.4 , k = 9.493306 >
```

```
] >m= 100
>s = sqrt (36)
>p1 = pnorm (110,m,s)
> cat ("p (x<=110) is = ", p1)
 p (x<=110) is= 0.9522096 >
>p2 = 1- pnorm (105 ,m, s)
>cat ("p (x>105) is = ", p2)
 p (x >105) is = 0.2023284 >
>p3= pnorm (92 ,m, s)
> cat ("p (x <=92) is = ", p3)
 p (x<=92) is = " 0.09121122 >
>p4 = pnorm (110 ,m, s) - pnorm (95, m, s)
> cat (" p (95 <= x <= 110 ) is = ", p4)
 p (95 <= x <= 110 ) is = 0.7498813 >
```

Scanned with CamScanner

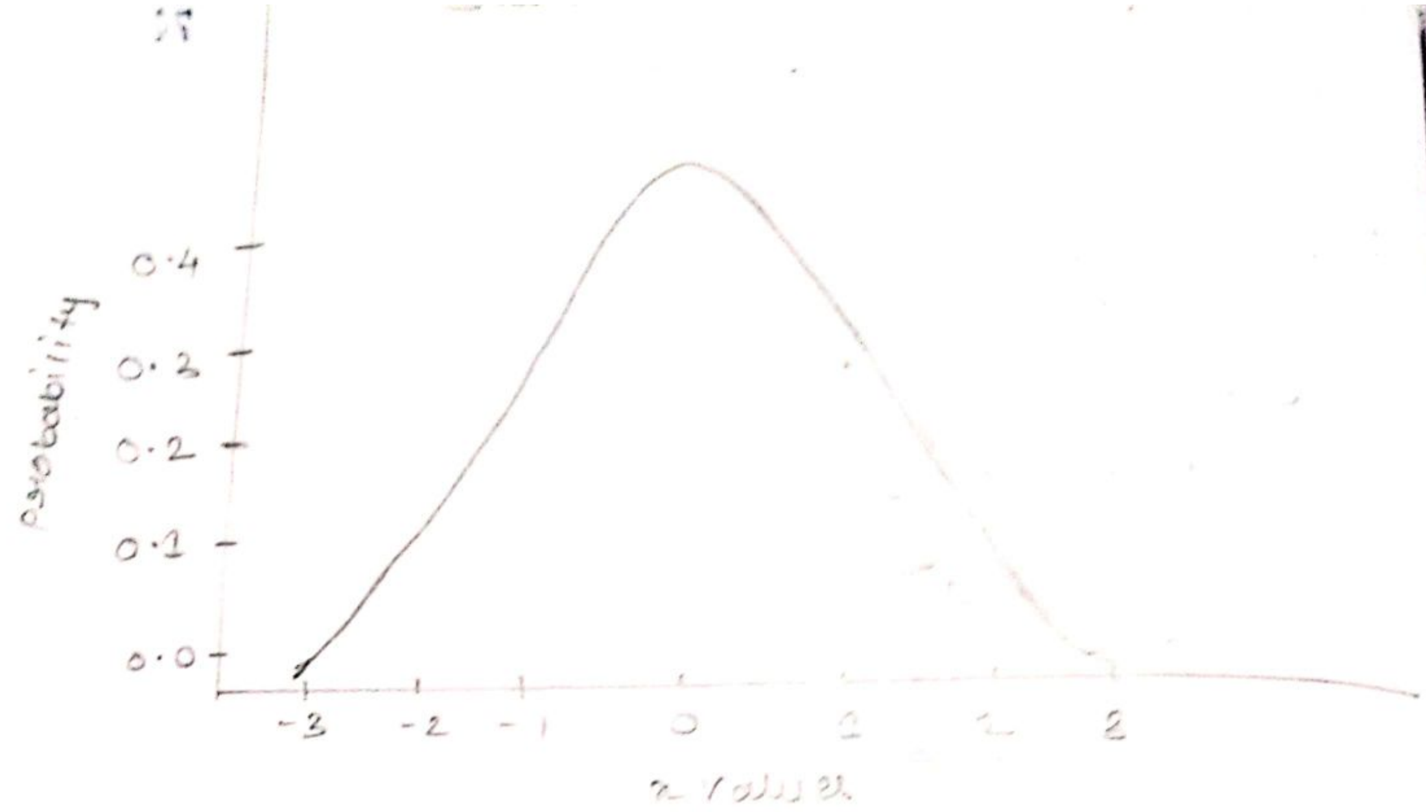
= qnorm
(iii) p (x < k) = 0.9, k
p (x < k) = 0.9, k = 15 = 10 + ...


```
> μ = 10
> s = 8
> n = 10
> x = rnorm (10,10,8)
> x
[1] 11.622405   9.636588   12.259680
     9.051027  15.317658   12.706422
    10.825069  10.117840    8.762215
     6.842165

> m = mean (x)
> m
[1] 11.18102
> m = median (x)
> m
[1] 10.77636
> v = (n-1) * ( var (x) /n)
> v
[1] 8.977855
> sd = sqrt (v)
> s d = sqrt (v)
> sd
[1] 2.996307
```

] Plot the Standart Normal Curve

```
> x = seq (-3, 3, by = 0.1)
> y = dnorm (x)
> plot ( x, y, xlab = "x value", ylab = " Probability",
   main = " Standard normal curve")
```



Q.5.] X ~ N (50, 100)

Find 1.] P(X ≤ 60)   2.] P(X > 65)

3.] P(45 ≤ X ≤ 60)

```
→ > u = 50
> s = sqrt(100)
= s
[1] 10
> p1 = pnorm(60, u, s)
> p1
[1] 0.8413447
> p2 = 1 - pnorm(65, u, 5)
> p2
[1] 0.668072
> p3 = pnorm(60, u, 5) - pnorm(45, u, 5)
> p3
[1] 0.5328072
```

Scanned with CamScanner

Practical No. 6

Topic :- z and t distribution Sums.

Q1.] Test the hypothesis Ho : $\mu = 20$ against $H_1 : \mu \neq 20$ sample of size 400 with selective and sample mean 20.2 and a Standard deviation 2.25 Test that 5% level of significance.

```
> mo = 20
> mx = 20.2
> sd = 2.25
> n = 400
> zcal = (mx - mo) / (sd / sqrt (n))
> zcal
[1] 1.777778
> cat ("z calculates is =", zcal)
z calculated is = 1.777778>
> pvalue
[1] 0.0754036
```

Since pvalue is more than 0.05 we accept Ho: μ= 20.

Q.2.] We want to test the hypothesis Ho: $\mu = 250$ against Ho: $\mu \neq 250$ a sample of size 100 has a mean of 275 and sd as 30. Test the hypothesis at 5% level of significance.

```
-> mo = 250
> mx = 275
> sd = 30
> n = 100
```

```
>zcal = (mx - mo) / (sd- sqrt (n))
>zcal
[1] 8.333333
> cat ("z calculated is =", zcal)
z calculated is = 8.333333 >
> pvalue = 2* (1 -pnorm (abs (zcal)))
>pvalue
[1] 0
```

Since pvalue is less than 0.05 we reject Ho

Q8:] we want to test the hypothesis Ho: P = 0.2 against H1: P ≠ 0.2 (P = population proportion) A sample of 400 is selected and the sample proportion is calculated 0.125 test the hypothesis at 1% level of significance.

Ho'.

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

Since pvalue is less than 0.05 we reject Ho

Scanned with CamScanner

] In a big city 325 men out of 600 men where found to be self employed. Thus this information support the conclusion that exactly that half of the men are self-employed 76

→ > p = 0.5
> n = 600
> p = 325/600
> Q = 1-p
> zcal = (p-P)/sqrt(P * Q/n))
> zcal
[1] 2.041241
> pvalue = 2 * (1-pnorm(abs(zcal)))
> pvalue
[1] 0.04122683

Since pvalue is less than 0.05 we reject H0.

Ho.

ition)
e
ypothesis

05.] Test the hypothesis of H0: U = 50 against H0: U ≠ 50
A sample of 30 is collected -
50, 49, 52, 44, 45, 48, 46, 45, 49, 45, 40, 47, 55
54, 46, 58, 47, 44, 59, 60, 61, 41, 52, 44, 55, 56, 46,
45, 48, 49

Ho!

→ m0 = 50
> x = c(50, 49, 52, ..., 48, 49)
> n = length(x) > n [1] 30
> mx = mean(x) > mx [1] 49.3333
> variance = (n-1) * var(x)/n [1] 30.45556
> sd = sqrt(Variance) > sd [1] 5.563772
> zcal = (mx - m0) / sd / (sqrt(n))
> zcal
[1] -0.6562964

> pvalue = 2 * (1 - pnorm ...

> pvalue

[1] 0.5116234

Since pvalue is more than 0.05 we accept

Ho : U = 50

AM 27-11-20

Practical - 7.

Topic :-> Large Sample test

Q.1] Two random sample are drawn from two population with a standard deviation 2 & 3 Respectively test the hypothesis that the two population means are equal or not at 5% level of significance. sample means are 67 and 68 Respectively. of size 1000 & 2000

```
-> >n1 = 1000
>n2 = 2000
>mx1 = 67
>mx2 = 68
>sd1 = 2
>sd2 = 3
>zcal = (mx1 - mx2) / sqrt((sd1^2/n1) + (sd2^2/n2))
> cat ("z calculated is =", zcal)
    z calculated is =  - 10.84652
>pvalue = 2 * (1-pnorm (abs (zcal)))
>pvalue
[1] 0
```

Since pvalue is less than 0.05 we reject H0:

$\mu_1 \neq \mu_2$.

]A study of noise level in 2 hospitals is done following data is calculated. First sample size 84, First sample mean 1.2, First standard deviation 7.9 second sample size = 34, second sample mean = 54.4, second standard deviation = 7.8 test Ho: u1 = u2 at 1 % of los.

>n1 = 84
>n2 = 34
>mx1 = 61.2
>mx2 = 54.4
>sd1 = 7.9
>sd2 = 7.8
>zcal = (mx1 - mx2)/sqrt (sd1 ^2/n1) + (sd2^2/n2))
>cat ("z calculated is =", zcal)
~~z calculated is = 1.3~~
z calculated is = 1.31 +177)
>pvalue = 2 * (1 - pnorm (abs (zcal)))
>pvalue
[1] 0.258006.

Since pvalue is greater than 0.01, we accept Ho u1 = u2.

3.] From each of two population of oranges the following samples are collected test whether the proportion of bad oranges are equal or not. First sample size = 250 second sample size = 200; n/o of bad oranges

In the first sample = 44 and second sample = 30

Ho: P1 = P2 against H1 : P1 ≠ P2

>n1 = 250
>n2 = 200
>p1 = 44/250
>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] 0.7393581
>pvalue = 2 * (1 - pnorm (abs(zcal)))
>pvalue
[1] 0.4596596

Since p value is greater than 0.05, we

Q.4] Random Sample of 400 males and 600 females when asked whether they want the ATM nearly 200 males and 390 females were in favour of the proposal. Test the hypothesis that the proportion of males and females proposal are equal or not at 5% level of significance

>n1 = 400
>n2 = 600
>p1 = 200/400
>p2 = 390/600
>p = (n1 * p1 + n2 * p2) / (n1 + n2)
>Q = 1 - P
>zcal = (p1 - p2) / sqrt (P * Q * (1/n1 + 1/n2)

>zcal
[1] -4.724751
>pvalue = 2 * (1 - pnorm (abs (zcal)))
>pvalue
[1] 2.303972e-06
Since pvalue is less than 0.05 we reject Ho: $\mu_1 = \mu_2$

5] Following are the two independent sample from the two population test equality of two means at 5% LOS

x1 = 74, 73, 79, 77, 76, 82, 72, 75, 78, 77, 78, 76, 76, 74
x2 = 72, 76, 74, 70, 70, 78, 70, 72, 75, 79, 79, 74, 75, 78, 72, 74, 80.

-> x1 = c(74, 77, 79, 77, ... 74)
>n1 = length (x1) > [1] 14
>mx1 = mean(x1)
>variance1 = (n1 - 1) * var (x1) / n1 [1] 0.4508
>sd1 = sqrt (variance 1) [1] 0.6711
>x2 (c 72, 76 .... 74, 80)
>n2 = length (x2)
>m2 = mean (x2)
>variance 2 = (n2 - 1) * var (x2) / n2 [1]
0.6163
>sd2 = sqrt (variance2) [1] 0.7850
>t.test () (x1, x2)
>pvalue = 0.1387
>0.05 is accepted.

Practical-8

Topic :-> Small Sample test

Q1.] Random Sample of 15 observation are given by 80, 100, 110, 105, 122, 70, 120, 110, 101, 88, 83, 95, 89, 107, 125. Do this data Support that Population mean is 100

Soln -> x = c (80, 100, 110, 105, 122, 70, 120, 110, 101, 88, 83, 95, 89, 107, 125)

> x

[1] 80 100 110 105 122 70 120 110 101 88 83 95 89 107 125

> t.test (x)

one sample t-test

data : x

t = 24.029, df = 14, p-value = 8.819e-13

alternative hypothesis : true mean is not equal to

91.37775 109.28892

Sample estimates

mean of x

100.3333

Since the Pvalue is greater than (0.005)

the Pvalue is ~~greater~~ Rejected

Ho: $\mu \neq 100$

Q8 [illegible] medicines are [illegible] on [illegible]
patients test starting the weight [illegible]
in the weight after medicine are given
below

medicine A (kg)	10 12 13 11 14
medicine B	8 9 12 14 15 10 9

Is there any significance efficiency of the medicine ?

Soln -> $H_0 = \mu_1 = \mu_2$

> x2 = c(8, 9, 12, 14, 15, 10, 9)

> x2

[1] 8 9 12 14 15 10 9

> x1 = c(10, 12, 13, 11, 14)

> x1

[1] 10 12 13 11 14

> t.test(x1, y2)

Welch two Sample t-test

data : x1 and y2

t = 0.80384, df = 9.7594, p-value = 0.4406

alternative hypothesis : true difference in means is not equal to 0.95%.

interval :

-1.781171 3.781171

Sample estimates

mean of x mean y

12 11

Q.4.] The weight Reducing diet Program conducted and observation are noted for 10 participants test whether for program is effective or not.

Before = (120, 125, 115, 130, 123, 119, 122, 127, 123, 118)

After - (112, 114, 107, 120, 115, 112, 112, 125)

Ho - there is no significant difference in weight

against H1 = the diet program reduce weight

Soln -> > x = c(120, 125, 155, 130, 123, 119, 120, 127, 128, 118)

> x

[1] 120 125 155 130 123 119 120 127 128 118

> y = c(112, 114, 107, 120, 115 112, 112, 120, 119, 11)

> t.test(x, y paired = T, alternative = "less")

paired t-test

data x and y

t = 3.0792, df = 9, p-value = 0.9934

alternative hypothesis: term difference in mains is less than 0 95 % confidence interval

-Inf 19.62252

Sample estimates:

mean of diff

12-3

Q5] Sample A = 66, 67, 75, 76, 82, 84, 88, 90, 92
Sample B = 64, 66, 74, 82, 85, 87, 92, 93, 95, 97
Test the Population mean are equal or not

Q6] following are

Soln-> Ho : u1 = u2

> x = c (66, 67, 75, 76, 82, 84, 88, 90, 92)

> x

[1] 66 67 75 76 82 84 88 90 92

> y = c (64, 66, 74, 82, 85, 87, 92, 43, 95, 97)

> y

[1] 64 66 74 82 85 87 92 43 95 97

> t.test (x, y)

> t

t -0.63968, df 17.974 pvalue = 0.5304

95 percent confidence interval

-12.8359 6.85399

mean of x mean y

80 83

∴ the Pvalue is greater that 0.05 we

accept

Ho : u1 = u2

5% level of significance

Q.6] Following are the marks and before and After of a marking program test the Program is effective or not

Before : 71, 72, 74, 69, 70, 74, 74, 70, 73, 75

After : 74, 77, 74, 73, 79, 76, 82, 72, 75, 78

→ H_0 = No Significance differences in marks

H_1 the test Program increases the mark

> x = c (71, 72, 74, 69, 70, 74, 76, 70, 73, 75)

> y = c (73, 77, 74, 73, 79, 76, 82, 72, 75, 78)

> t. test (x, y, paired = T, alternative = "greater")

t = -44691 df = 9, pvalue = 0.9992.

95 ~~pr~~ percent of condifidence interval

-5.076639

mean of the differences

= 3.6.

the Pvalue is greator that 0.05, we accept $H_0 = H_1$

increase in marks at 5% level of significance.

AM

Topic :→ Large and small sample test.

1.] The arithmetic mean of a sample of 100 items from a large population is 52 if the SD is 7 Test the hypothesis that the population mean is 55 Against the alternative it is more then 55 at 5% los level of significance.

2.] In a big city 350 out of 700 males are found to be smoker thus this info - support that excatly half of the males in the city are smokers ? test thus 1% LOS.

3.] Thousand artical from a factory A are found to have 2% defectives. 1500 articals from a second factory B are found to have 1% defective. Test that 5% of LOS. That the 2 factory are. similar or not

4.] A Sample of size 400 was drawn and the sample mean is 99 test of 5%. LOS that the Sample comes from a population with mean 100 and Variance 69?

5.] The flower stairs are selected and the heights are found to be (cm) 63, 63, 68, 69, 71, 71, 72, Test the hypothesis that the mean height is 66 at 1%. of 1% LOS.

6.] Two Random samples were drawn from two normal Population and the Values are

A :→ 66, 67, 75, 76, 83, 84, 88, 90, 92

B :→ 64, 66, 74, 78, 82, 85, 87, 92, 93, 95, 97

Test whether the population have the same Variance at 5% LOS

Solution :->

Ho $\mu_1 = \mu_2$ against $H_1 \mu_1 \neq \mu_2$

1.] >n= 100

>mx = 52

> mo= 55

> sd = 7

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

>zcal

[1] -4.285719

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

>pvalue

[1] 1.82153e -05

Pvalue is <0.05 we reject

(2) Ho $\mu_1 = \mu_2$ against H $\mu_1 \neq \mu_2$

(2) >p =0.5

>p = 350/700

>p

[1] 0.5

>n = 700

> Q = 1-p

>Q

[1] 0.5

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

>zcal

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

pvalue > 0.05 we accept

(3) >n1 = 1000
>n2 = 1500
>p1 = 20/1000
>p2 = 20/1500
>p = (n1 * p1 + n2 * p2)/(1/n1 + 1/n2)
>p
>zcal = (p1 - p2) /sqrt (p * q * (1/n1 + 1/n2))
>zcal
[1] 2.082731
>pvalue = 2 * (1 - pnorm (abs (zcal)))
>pvalue
[1] 0.03727577

Since pvalue < 0.05

∴ we reject pvalue.

4] > n = 400
> mx = 99
> mo = 100
> Variance = 64
> sd = sqrt (Variance
> sd
[1] 8
> zcal = (mx - mo) / (sd / sqrt(n))
> zcal
[1] -2.5
> pvalue = 2 * (1 - pnorm (abs (zcal)))
> pvalue
[1] 0.0124193

Since pvalue < 0.05
∴ we reject H0 $\mu_1 = \mu_2$

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

One Sample t-test

data: x

t = 47.54 , df = 6, p-value = 5.522e-09

alternative hypothesis: true mean is not equal to 0

95 % Confidence interval:

64.56479 73.62092

Sample estimate
mean of x

65.14286

pvalue is < 0.05 we reject

H_0 $\mu_2 = \mu_2$ against H_0 $P_2 = P_0$

```
> x = c(66, 67, ....... , 90, 92)
> y = c (64, 66, ....... , 95, 97)
> F = var.test(x,y)
> F
```

F test to compose two Variances

data: x and y

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

alternative hypothesis: true ratio of variances is not equal

$\lambda = 1$

95 percent confidence interval:

0.1833662   3.0360393

Sample estimates

ratio of variance

0.7068567

AM 17.2.2

H_0 $\mu_1 = \mu_2$ against H_0 $P_1 = P_2$

$\because$ pvalue > 0.05 we accept H_0 $\mu_1 = \mu_1$

Practical 10.

Anoba and chi squares.

Use the following data to test whether the cleanliness of home and cleanliness of child are independent or not.

Cleanliness of Child

	Clean	Dirty
~~Clean~~ ~~Fairly~~ Clean	70	50
Fairly ~~Dirty~~ Clean	80	20
Dirty	35	45

Ho: CC & CH

```
> x = c ( 20, 80, 35, 50, 20, 45)
> x
[1] 20  80 35 50  20 45
> m = 3
> n = 2
> y = matrix (x, nrow = m , ncol = n)
> y
        [,1]  [,2]
[1,]     20    50
[2,]     80    20
[3,]     35    45

> pv = chisq.test (y)
> pv
```

pearson's chi-squared test

data: y

x-squared = 49.82, df = 2, p-value

Q.2-] Use the following data to find a Vaccination ~~Variation~~ and Particular disease are independent or not

Soln->

		Aff	Not Aff
VGC	Given	20	30
	Not Given	25	35

Soln-> Ho: Vaccination and diseases are independent

```
>x=c(20,25,30,35)
>x
[1] 20  25  30  35
>m=2
>n=2
>y= matrix(x, nrow=m, ncol=n)
>y
      [,1]  [,2]
[,1]   20    30
[,2]   25    35

>pv = chisq.test(y)
>pv
```

Pearson Chi-squared test with
03 yates's continuity correction

data : y

X-squared = 0 , df = 1, p-value = 1

Q.3.] Perform Anova for following data

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

Soln -> Ho: Mean of the Varieties 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))
>d
```

	values	ind
1	50	b1
2	52	b1
3	53	b2
4	55	b2
5	53	b2
6	60	b3
7	58	b3
8	57	b3
9	56	b3

10	52	b4
11	54	b4
12	54	b4
13	55	b4

```
>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 , dnom df=9 , p-value=
0.001
> anova= aov (values ~ ind , data=d )
> anova
 call:
   aov (Formula = Values ~ind , data=d)
Terms:                          ind          Residual
  Sum of squares            71.06410         18.1667
  Deg of Freedom               3                9

Residual   Standard error :  1.420746
estimated . effects   may  be  unbalanced
```

∵ p-value is less than 0.005
we reject Ho

Q.4] The following data gives life of tree

Type	Observations
A	20, 33, 18, 17, 18, 22, 24
B	19. 15, 17, 20, 16, 17
C	21, 19, 22, 17, 20
D	15, 14, 16, 18, 16

Test the hypothesis average life of four branches are same

Soln → Ho: the average life of a branches are same

```
>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, 16)
>d = stack( list( b1 = x1, b2 = x2, b3 = x3, b4 = x4))
>d
```

	Values	ind
1	20	b1
2	23	b1
3	18	b1
4	17	b1
5	18	b1
6	22	b1
7	24	b1
8	21	b2
9	19	b2
10	22	b2

	17	b2
11	20	b2
12	15	b2
13	14	b3
14	22	b3
15	17	b3
16	20	b3
17	15	b4
18	14	b4
19	16	b4
20		
21	18	b4
22	14	b4
23	16	b4
24		

```
> names (d)
[1] "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 = 
                                              0.00234
> anova = aov (values ~ ind, data = d)
> anova

Call:
   aov ( Formula = values ~ ind, data = d)

                          ind       Residuals
 Terms
                          91.4381   84.0619
Sum of Squares
                          3         20
 Deg of Freedom

Residual Standard error:  2.110236
```

estimated effects may be unbalanced

∴ the P-value is less than 0.005 we
reject the values.

Q.5] 2000 students of a college are graded accordingly to their IQ and economic condition of home check that is the any association between their economic of house and they're IQ?

	IQ	
EC	High	low
High	460	140
Mid ~~Medium~~	330	200
Low	240	160

Soln → Ho : EC & IQ are independent

```
>x=c(460, 330, 240, 140, 200, 160)
>m=3
>n=2
>y=matrix(x, nrow=m, ncol=n)
>y
        [,1]   [,2]
[,1]    460    140
[,2]    330    200
[,3]    240    160
>pv=chisq.test(y)
>pv
```

pearson's chi-squared test

data : y

X-squared = 39.426, df = 2, pvalue = $2.364e^{-09}$
Since p value is less than 0.005 we reject
H0 : EC & IQ are independent.

Practical 11

Non-Parametric Test

Q1] Following are the amounts of sulphur oxide emitted by industries in 20 days apply sign test to test the hypothesis that population median is 21.5 ?

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

Ho: population median is 21.5

```
> x=c(17,15,20,29,19,18,22,25,27,9,24,20,17,6,
   24,14,15,23,24,26)
> sp= length (x [x>median])
> sp
> [1] 9
> ~~A= sp~~ sn= length (x [x<median])
> sn
> [1] 11
> n =sp+sn
> n
> [1] 20
> pv = pbinom (sp, n, 0.5)
[1] 0.4119015
```

∴ P value is greater than 0.059

we accept the value at 5% level of significance.

Q.2] Apply sign test hypothesis that population median 6:5 against alternative is greater than 6:5 at 1% los?

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

soln -> 9D the Alternative is greater than

```
> x = c( 612, 619, 631, 628, 643, 640, 655, 649, 670, 670,
     683)

> median = 24
> sp = length (x[x > median])
> sp
[1] 10
> sn = length(x[x < median])
> sn
[1] 10
> n = sp+sn
> n
[1] 20
> pv = pbinom(sn, n, 0.1)
> p
[1] 0.9999993
```

Q.3] 10 Observations are 36, 32, 21, 30, 24, 25, 20, 22, 20, 18 using sign test hypothesis that population median is 25 against alternative is less than 25 at 5% level of significance?

Soln →
```
>x = c(36, 32, 21, 30, 24, 25, 20, 22, 20, 18)
>median = 25
>sp = length(x[x>median])
>sn = length(x[x<median])
>n = sp+sn
>n
[1] 8
>pv = pbinom[sp, n, 0.5]
>pv
[1] 0.3632813
```

Q.4] The following are some measurements 63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 65 using wilcoxn signed rank test. Test the hypothesis that population median is 60 against alternative it is greater than 60, at 5% los?

Soln → Ho: population median is 60

```
>x=c(63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 65)
> wilcox.test(x, alt = "greater", mu = 60)
```

wilcoxon sign rank test with continuity correction

data: x

V = 48.5, p-value = 0.9232

alternative hypothesis: true location is less than 20.

5] 15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26.
use WSRT to test hypothesis that population median 20 against alternative is lesser than 20.

>x=c (15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26.)
>wilcox ~~.text~~ test (x)
wilcoxon sign rank test with continuity correction
data: x
V=48.5, p-value = 0.9232
alternative hypothesis : true location is less than 20.

6] 20, 25, 27, 30, 18 test the hypothesis that population median is 25 against alternative is not less than 25?
>x=c (20, 25, 27, 30, 18)
>wilcox.test (x, alt = "two sided", mu = 25)
wilcoxon signed. Rank. test with continuity correction.
data: x
V=3.5, p-value = 0.7127
alternative hypothesis: true location is not equal to 25.

PM 9.3.20