INDEX

No.	Title	Page No.	Date	Staff Member's Signature
	Sem II			
1.				
1.	Basic of r software	35-37	25/11/19	AY
2.	Binomial distribution	37	2/12/19 – 16/12/19	
3.	Probability distribution	41	9/12/19.	
4.	Practice	~~23/12~~ 44	23/12/19	AY
5.	Normal distribution	45	6/1/20	
6.	Z & t distribution	47	20/1/20	AY
7.	Large sample test	49	27/1/20	AY
8.	Small sample test	51	3/2/20	AY
9.	large & small samples test	53	10/2/20	
10.	Anova and chi-squared test	55	17/2/20	
11.	Non parametric test	58	24/2/20	AY 4-3

Practical 1.

Basic of r software :

1. R is software for data analysis and stastical computing.
2. It is software by which effective data handling and outcome's storage is possible.
3. It is capable of graphical display.
4. It's a free software.

* $2^2 + |-5| + 4\times5 + 6/5$

2^2 + abs(-5) + 4*5 +6

[1] 30.2

* $x = 20$ → x =20

$y = 2x$ y = 2*x

$z = x+y$ z = x+y

$\sqrt{z}$ z^0.5 → 7.745967

* round (2.567)

→ 3

* $x = 10$

$y = 15$

$z = 5$

$x+y+z$ $\sqrt{xyz}$ → a=(x*y*z)^0.2 → 27.38613

→ 30

xyz → x*y*z round (a) → 27

→ 750

A vector in r software is denoted by the syntax " c ".

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

→ [1] 4 9 25 49

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

→ [1] 4 27 25 343

* c(2,4,6,8)*3

→ [1] 6 12 18 24

* seq(1,6,1)^c(2,3,4)

→ [1] 1 8 81 16 125 1296

* c(2,4,6,8) * c(-2,-3,-5,-7)

→ [1] -4 -12 -30 -56

* c(2,4,6,8) + 10

→ [1] 12 14 16 18

* c(2,4,6,8) + c(-2,-3,-1,0)

→ [1] 0 1 5 8.

Q. Find the sum, product, squareroot of sum and product for the following values.

4, 9, 2, 5, 7, 8, 3, 6, 15, 12, 10, 9, 8, 12, 14.

```
x = c(4,9,2,5,7,8,3,6,15,12,10,9,8,12,14)
y = sum(x)
y
z = prod(x)
z
y ^ 0.5
z ^ 0.5
```

⇒ 125

8.559323e+12

11.18034

2925632

Q. Find the sum, product, maximum, minimum values of (2, 8, 9, 11, 10, 7, 6) ^ 2

```
x = c(2,8,9,11,10,7,6) ^ 2
x
sum(x)
prod(x)
max(x)
min(x)
```

⇒ 4 64 81 121 100 49 36

455

4425974784 00

121

4

Q. matrix (nrow = 3, ncol = 4, seq (1, 8, 1))

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

Q. $\begin{bmatrix} 2 & 8 & 5 & 1 \\ 6 & 9 & 0 & 4 \\ 7 & 4 & 2 & 5 \end{bmatrix}$

→ matrix (nrow = 3 , nrow = 4, c (2, 6, 7, 8, 9, 4, 5, 0, 2, 1, 4, 5)

Q. $\begin{bmatrix} 4 & 7 & 4 \\ 5 & 8 & 0 \\ 6 & 9 & 2 \end{bmatrix}$ and $\begin{bmatrix} 6 & 11 & 9 \\ 4 & 12 & 7 \\ 5 & 8 & 4 \end{bmatrix}$

→ x = matrix (nrow = 3, ncol = 3, c (4, 5, 6, 7, 8, 9, 4, 0, 2)

y = matrix (nrow = 3, ncol = 3, c (6, 4, 5, 11, 12, 8, 9, 7, 4)

x + y

10	18	13
9	20	7
11	17	6

x

Scanned with CamScanner

$x * 2$

$$\begin{matrix} 8 & 14 & 8 \\ 10 & 16 & 0 \\ 12 & 18 & 4 \end{matrix}$$

$x * y$

$$\begin{matrix} 24 & 77 & 36 \\ 20 & 96 & 0 \\ 30 & 72 & 8 \end{matrix}$$

$y * 2$

$$\begin{matrix} 12 & 22 & 18 \\ 8 & 24 & 14 \\ 10 & 16 & 8 \end{matrix}$$

2.12.19

Practical-2

n = Total no. of trails
$p = p(\text{success})$
$q = p(\text{Failure})$
X = no. of success
outcome of n

e.g.
$P(x) = {}^nC_x\, p^x q^{n-x}$
$E(x) = np$
$V(x) = npq$

$n=10, p=0.6, q=0.4$
$p(x \le x) = p(x \le 3)$

$P(x)$	> dbinom(x,n,p)
$E(x)$	> n * p
$V(x)$	> n * p * q
$P(x \le x)$	> pbinom(x,n,p)

Q.1 Toss a coin 10 times with probability (Head) = 0.6
Let x be the no. of heads.
find probability of
1) 7 heads
2) 4 heads
3) atmost 4 heads
4) atleast 6 heads
5) No heads
6) All heads

and also expectation & variances.

→ $n=10$
$p=0.6$
$q=0.4$
$x=7$

dbinom(x, n, p)
[1] 0.2149908
dbinom(4, 10, 0.6)
[1] 0.1114767

pbinom (4,10,0.6)

[1] 0.6177194

1-pbinom (6,10,0.6)

[1] 0.3822806

dbinom (0,10,0.6)

[1] 0.0001048576

~~Q.2 suppose~~ dbinom (10,10,0.6)

[1] 0.006046618

n*p

[1] 1

n*p*q

[1] 24

Q.2 suppose there are 12 MCQ in an english question paper
→ each question has 5 answers $ only 1 of them is correct
Find the probability of having

1) 4 correct answers.
2) Atmost 4 correct answers
3) Atleast 3 correct answers.

Q.3 Find the complete binomial distribution when n is p = 0.1.

Q.4 Find the probability of exactly 10 succes out of hundred trails with p = 0.1

Q.5 x follows binomial distribution with n = 12 $ p = 0.25 find

1) $P(x \leq 5)$ 2) $P(x > 7)$ 3) $P(5 < x < 7)$

Q.6) There are 10 members in a commitee probability of any number attending a meeting is 0.9. what is the probability, that 7 or more members will be present in a meeting.

Q.7 A salesman has a 20% probability of making a sell to a customer. On a typical day he will meet 30 customers what minimum no. of sells he will make with 88% probability.

Q.8 For n=10 & p=0.6 Find the binomial probabilities & plot the graphs of pmf & CDF.

Note:

1) P(X=x) = dbinom (x,n,p)

2) P(X≤x) = pbinom (x,n,p)

= probability of atmost x values

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

4) If x is known & the probability is given as P, to

Find x = qbinom (P,n,P)

Answers:

2) n=12

P=1/5

P1 = dbinom (4,n,P)

P1

[1] 0.1328756

P2 = pbinom (4,n,P)

P2

[1] 0.9274445

P3 = 1- pbinom (2,n,P)

P3

[1] 0.4416543

3) n=5

p=0.1

x=0

dbinom (x,n,p)

[1] 0.59049

x=1

dbinom (x,n,p)

[1] ~~0.0729~~ 0.32805

x=2

dbinom (x,n,p)

[1] 0.0729

x=3

dbinom (x,n,p)

[1] 0.0081

x=4

dbinom (x,n,p)

[1] 0.00045

x=5

dbinom (x,n,p)

[1] 1e-05

4)

4) n=100

p=0.1

x=5

pbinom(x,n,p)

[1] 0.9455978.

5) n=12,

p=0.25

pbinom(5,12,0.25)

[1] 0.9455978

1- pbinom(7,12,p)

[1] 0.01425278

dbinom(6,n,p)

[1] 0.04014945

6) n=10

p=0.9

x=6

1- pbinom(x,n,p)

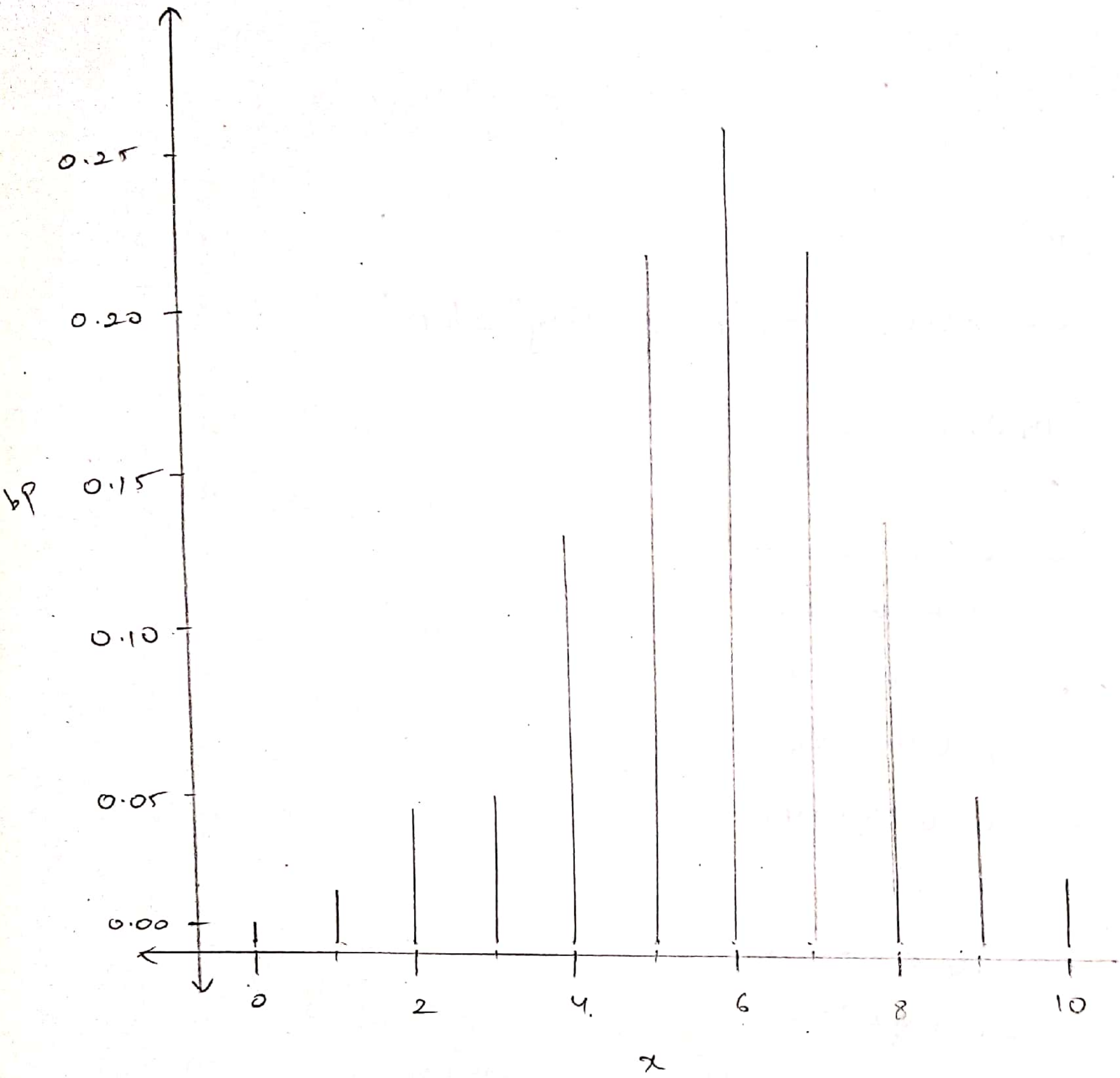
[1] 0.9872048.

40

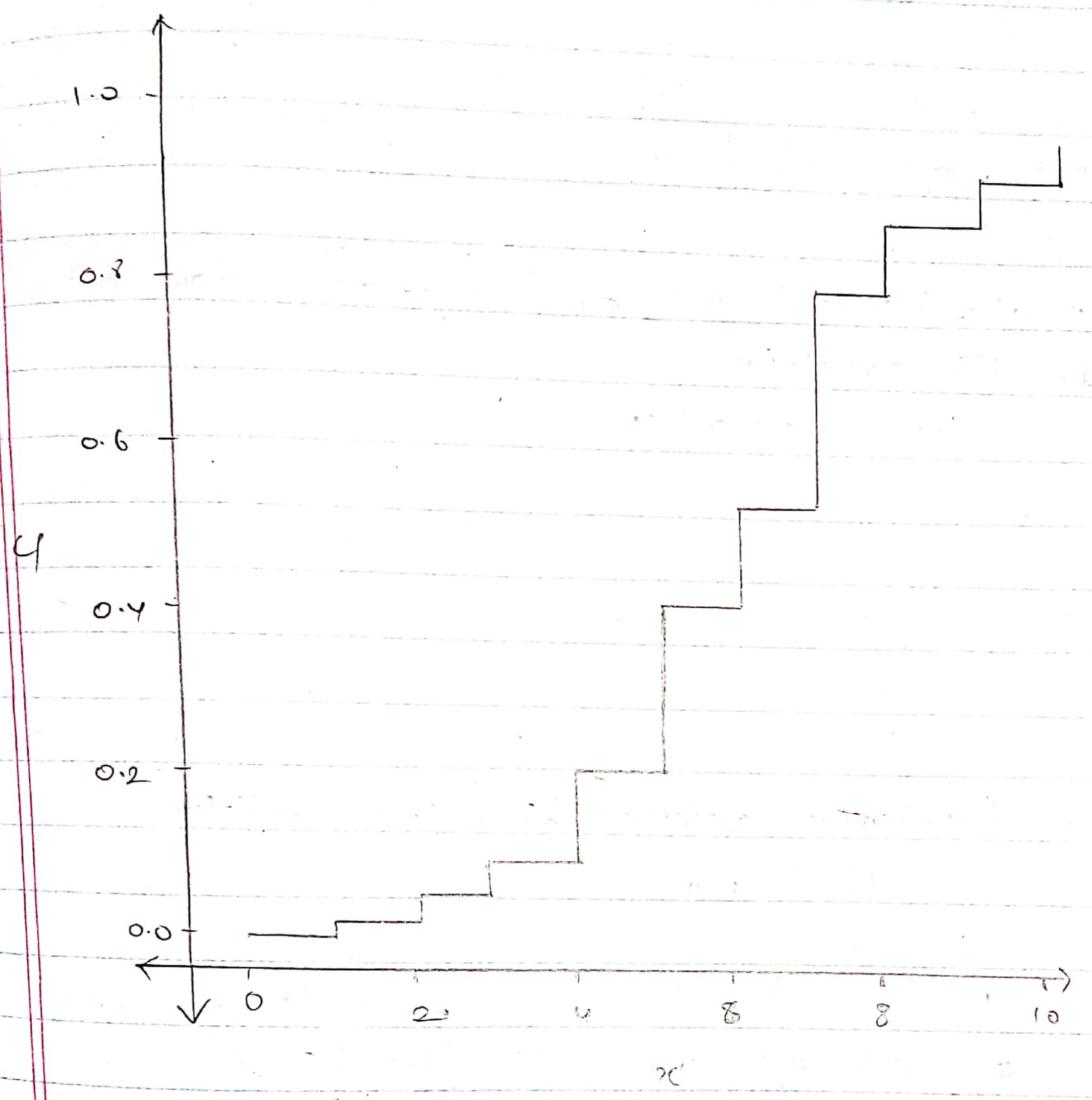
7) n = 30
p = 0.2
p1 = 0.88
qbinom (p1, n, p)
[1] 9.

8) n =10
p=0.6 , x = 0:n
bp = dbinom (x , n, p)
d= data.frame (" x-values"= x, "probability" = b.p)
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	0.2508226560
8	7	0.2149908480
9	8	0.1209323520
10	9	0.0403107840
11	10	0.0060466176



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



AM

Practical - 3

Probability distribution.

Q.1

1)

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

→ The given distribution is not p.m.f because it does not satisfy the 1st condition

i.e. $0 \le p(x) \le 1$

2)

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

→ The given distribution satisfies both the condition.

∴ It is pmf

3)

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

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

Hence, it doesn't satisfy ~~[1]~~ 2nd Condition

i.e Sum of the probability is more than 1

∴ It is not pmf.

Q.2 Following is a pmf of x

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

42

Find mean and variance of x.

→

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
		$\Sigma x.p(x) = 3.45$	$\Sigma x^2 p(x) = 13.55$

Mean $= E(x) = \Sigma x p(x) = 3.45$

Var $= V(x) = \Sigma x^2 p(x) - [E(x)]^2$

$= 13.55 - (3.45)^2$

$= 1.6475$

```
x = c(1,2,3,4,5)
prob = c(0.1,0.15,0.2,0.3,0.25)
a = x * prob
sum(a)
```

[mean] → 3.45

```
b = (x^2)* prob
Var = sum(b) - mean^2
Var
[1] 1.6475
```

SP :

Q.3 Find mean & variance of x

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

```
h =c (5,10 ,15, 20, 25)
hF= c (0.1,0.3, 0.2,0.25, 0.15)
v=h * hF
mean =sum(v)
[1] 15.25
n= (h^2)*hF
[1] 2.5   30.00   45.00   100.00   93.75
var = sum(n)- mean^2
[1] 38.6875
```

Q.4 Find c.d.f of the following p.m.f and draw the graph.

①

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

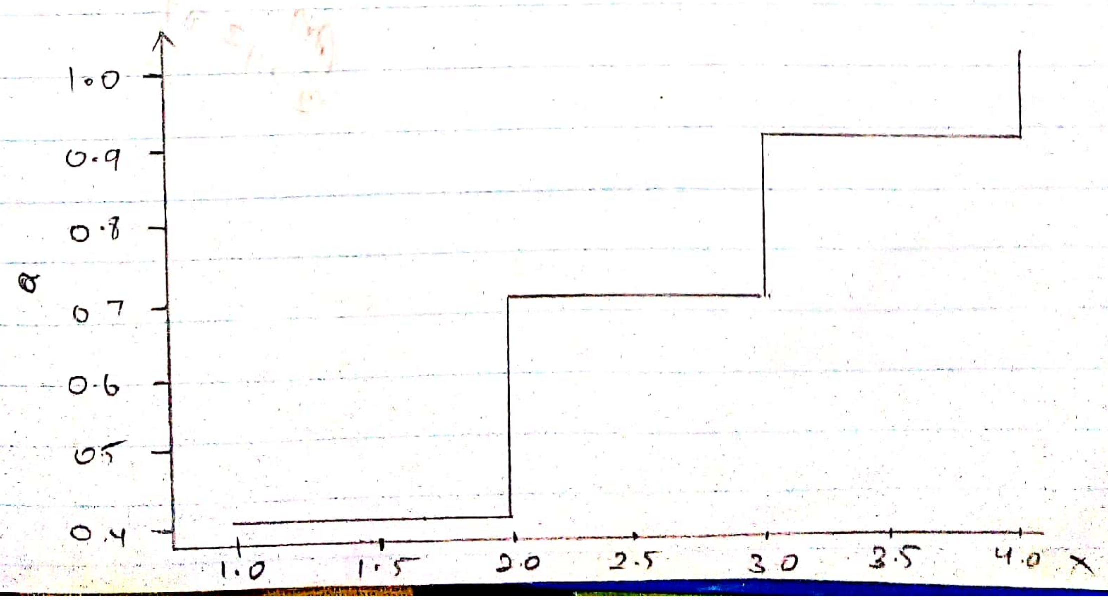
→

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

→ 0.4 0.7 0.9 1.0

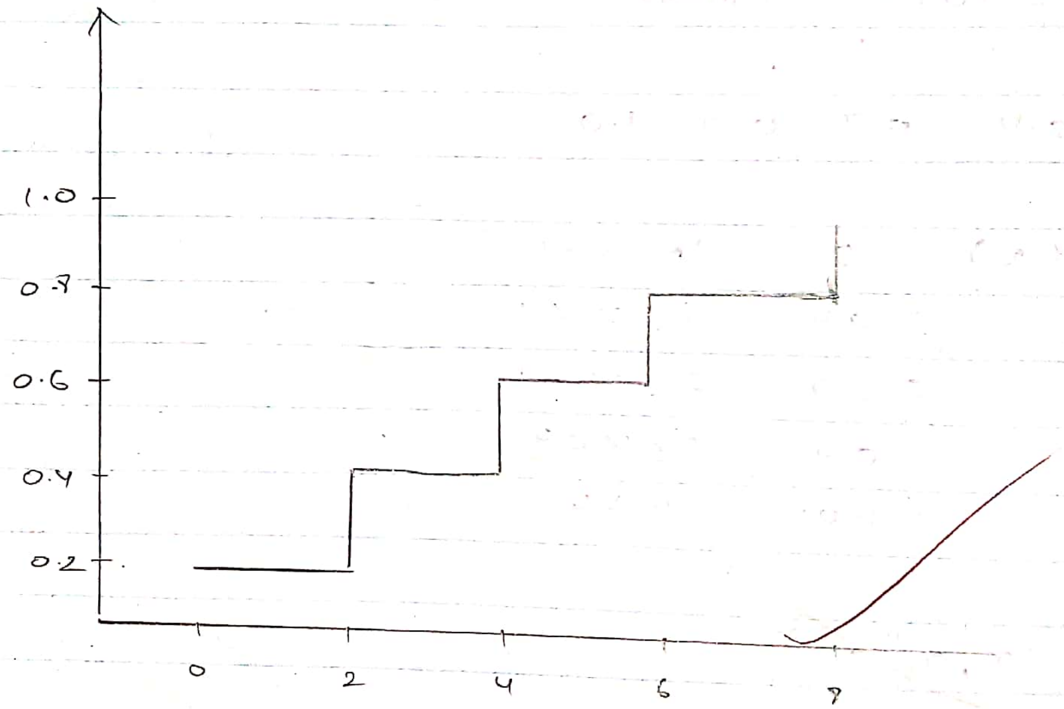
$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.4 & 1 \leq x < 2 \\ 0.7 & 2 \leq x < 3 \\ 0.9 & 3 \leq x < 4 \\ 1.0 & x \geq 4 \end{cases}$$

```
Plot(x,a,"s")
```



②

```
x =c (0,2,4,6,8)
prob =c ( 0.2, 0.3, 0.2, 0.2, 0.1)
a = cumsum (prob)
a
[1] 0.2   0.5   0.7   0.9   1.0
plot (x, a, "s")
```



27/12/19

Practical - 4

(practice)

$$p(X=x) = \binom{n}{x} p^x q^{n-x}$$

① $p(X=7) = \binom{8}{7} (0.6)^7 (0.4)^{8-7}$ ... ($n=8, p=0.6, q=0.4$ given)

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

$= 8 \times 0.2799 \times 0.4$

$= 0.08957$

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

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

$+ {}^8C_2 (0.6)^2 (0.4)^6 + {}^8C_3 (0.6)^3 (0.4)^5$

$= 1 \times 1 \times 0.00065536 + 8 \times 0.6 \times 0.0016384$

$+ 28 \times 0.36 \times 0.004096$

$+ 56 \times 0.216 \times 0.01024$

$= 0.1736704$

③ $p(x = 2 \text{ or } 3) = p(2) + p(3)$

$= {}^8C_2 (0.6)^2 (0.4)^6 + {}^8C_3 (0.6)^3 (0.4)^5$

$= 28 \times 0.36 \times 0.004096 + 56 \times 0.216 \times 0.01024$

$= 0.04128768 + 0.12386304$

$= 0.16515072$

Scanned with CamScanner

Practical - 5

Normal distribution:

Is an example of continuous probability distribution.

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

① $p(X=x) = \text{dnorm}(x, \mu, \sigma)$

② $p(X \le x) = \text{pnorm}(x, \mu, \sigma)$

③ $p(X > x) = 1 - \text{pnorm}(x, \mu, \sigma)$

④ To find the value of K

$p(X \le K) = p$

$\text{qnorm}(p, \mu, \sigma)$

⑤ To generate a random sample of size n,

$\text{rnorm}(n, \mu, \sigma)$

1) A random variable X follows normal distribution with $\mu = 10$, $\sigma = 2$. Find the following

1) $p(X \le 7)$, 2) $p(X > 12)$, 3) $p(5 \le X \le 12)$, 4) $p(X < K) = 0.4$

2) $X \sim N(100, 36)$

$p(X \le 110)$, $p(X > 105)$, $p(X \le 92)$, $p(95 \le x \le 110)$

$p(X < K) = 0.9$

3) Generate a 10 random sample Find the sample mean, median & variance & S.D

$\mu = 10$, $\sigma = 3$

Scanned with CamScanner

4) Plot the standard Normal curve

5) $x \sim N(50, 100)$
$p(X \le 60)$, $p(X > 65)$, $p(45 \le x \le 60)$,

Solution:

1) ① P1 = pnorm (7,10, 2)
cat ("p(x ≤ 7) is =", p1)
→ p (x≤7) is = 0.0668072

② p2 = 1- pnorm (12,10,2)
cat (" p(x≤12) is ~~=~~", p2)
→ p (x<=12) is = 0.1586553

③ P3 = pnorm (12,10,2) - pnorm(5,10,2)
cat ("p(5≤x ≤12) is =", p)
→ p (5≤x≤12) is = 0.8351351

④ K = qnorm (0.4,10,2)
cat ("p (x <k) is =", k)
→ p (x<k) is = 9.493306

2) P1 = pnorm (110,100,6)
p1
→ [1] 0.9522096

P2 = 1 - pnorm(105,100,6)
P2
→ [1] 0.2023284

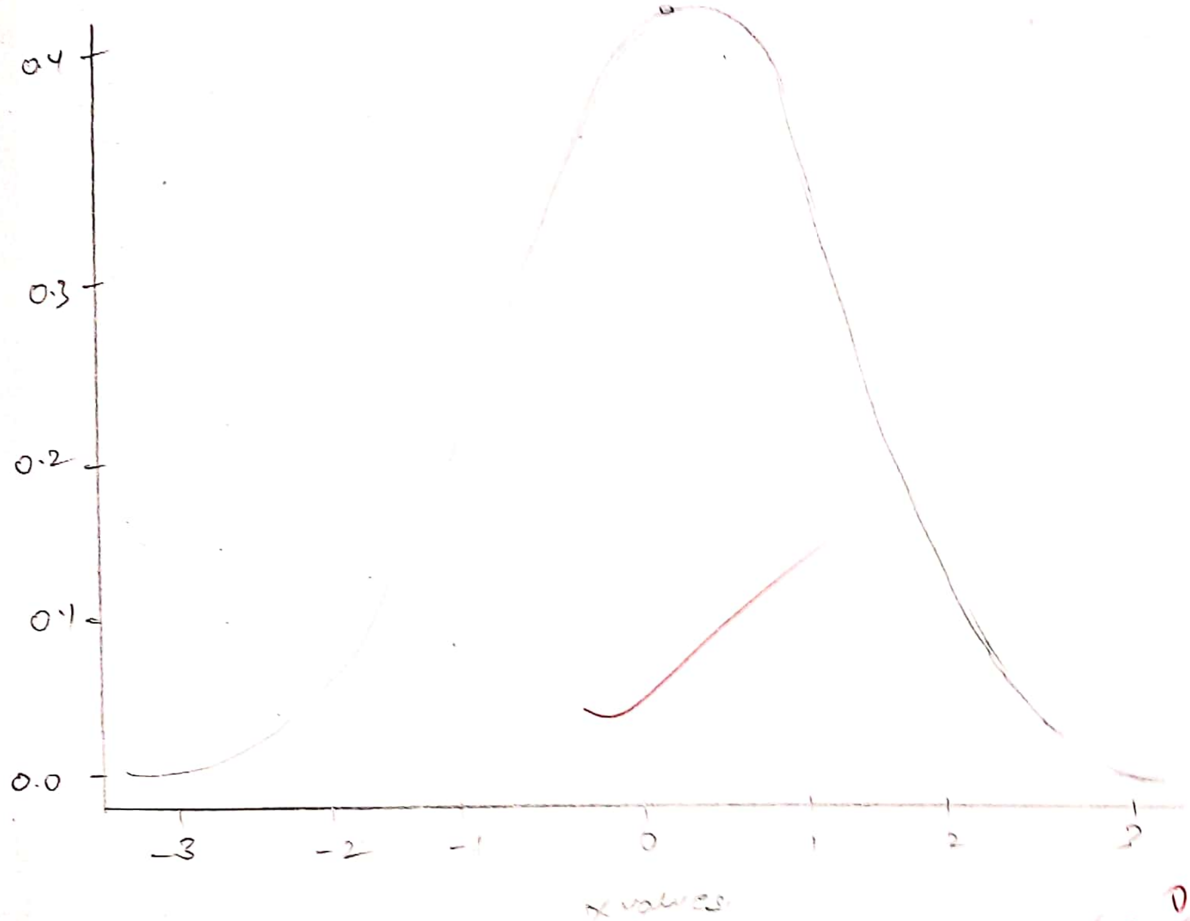
P3 = pnorm (92,100,6)
P3
→ [1] 0.09121122

P4 = pnorm (110 ,100, 6) - pnorm (95,100,6)
P4
→ [1] 0.7498813

K = qnorm (0.9, 100,6)
K
→ [1] 107.6893

3) X = rnorm(10,10,3)
mean (X)
→ [1] ~~10.43946~~ 10.11893
median (x)
→ [1] 10.43946
> n =10
> var6 =(n-1) * var(x) /n
> sqrt(var6)
> [1] 2.189529

4) x = seq (-3, 3, by = 0.1)
y = dnorm (x)
plot (x, y, xlab="x values", ylab = "probability", main= "standard normal curve")



AY 6.1.20

5) X ~ N (50,100)
P(X ≤ 60), P(X ≥ 65)

p1 = pnorm(60, 50, 10)
[1] = 0.843447

p2 = pnorm (60, 50, 10) - pnorm (45, 50, 10)
[1] = 0.5328072

p3 = 1 - pnorm (65, 50, 10)
[1] = 0.0668072.

Aim : z $ t distribution.

Q1) Test the hypothesis Ho, u=20 against H1: u ≠ 20. A sample of size of 400 is selected & the sample mean is 20.2 & the S.D 2.25. Test at 5% level of significance.

```
> mo = 20
> mx = 20.2
> sd = 2.25
> n = 400
> zcal = (mx-mo) / (sd / (sqrt(n)))
> zcal
[1] 1.77778
> cat('zcalculated is =', zcal)
zcalculated is = 1.77778
> pvalue = 2 * (1-pnorm(abs(zcal)))
[1] 0.07545
```

Since, pvalue > 0.05, we accept Ho: u1 = 20

Q2 We want to test the hypothesis Ho: u = 250 against H1: u1 = 250 A sample of size 100 has a mean 275 & SD = 30. Test the hypothesis at 5% level of significance.

```
> mo = 250
> mx = 275
> sd = 30
> n = 100
> zcal = (mx-mo)/ (sd / sqrt(n))
[1] 8.333
```

SD

> Cal ('z calculated is =', zcal)

> Calculated is =8.333

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

[1] 0

Since, pvalue <0.05 we reject Ho

Q.3 We want to hypothesis $H_0 : P = 0.2$ against $H_1 : P_0 \neq 0.2$ (P = population proportion). A sample of 400 is selected $ the sample proportion is calculated 0.125. Test the hypothesis at 1% level of significance

> P =0.2

> Q =1-P

> p =0.125

> n =400

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

> zcal

[1] 8-3.75

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

[1] 0.0001

Since, pvalue < 0.01, we reject Ho.

Q.4 In a big city 3.25 men out of 600 men where found to be self-employed, this info support the conclusion at exactly 1/2 of men are self-employed?

Scanned with CamScanner

> p=0.5
> p = 325/600
[1] 0.541
> n =600
> q =1-p
[1] 0.5
> zcal =(p-P)/(sqrt (P*Q)/n))
[1] 2.041241
> pvalue =2*(1-pnorm (abs(zcal)))
[1] 0.04122
since p value < 0.05 we reject Ho

Q.5 Test the hypothesis Ho μ=50 against H1 : μ≠50 . A sample of 30 collected
{50 ,49 ,52,44,45,48,46,45,49,45, 40,47,55,54, 46, 58,47, 44,59,60, 61,41,52,44,55,56,48,45, 48 ,49}

> s = (50, 49,52,...)
> n = length (s)
> mx = mean (s)
[1] 49.33
> mo = 50
> sd = sd (s)
[1] 5.65
> var= (n-1) * var (s) /n
[1] 639.39
> zcal =(mo -mx) / (sd /sqrt(n))
[1] 0.0215
> pvalue =2 * (1-pnorm (abs(zcal)))
[1] 0.982
Since, pvalue > 0.05

AM

Aim : Large sample test.

Q.1) Two random sample of size 1000 and 2000 are drawn from two population with a standard deviation 2 and 3 respectively. Test the hypothethis that two population means are equal or not at 5% level of signification. Sample means are 67 and 68 respectively

$H_0 : \mu_1 = \mu_2$

against $H_1 : \mu_1 \neq \mu_2$

```
n1=1000
n2=2000
mx1 = 67
mx2 =68
sd1 = 2
sd2 =3
zcal = (mx1-mx2)/sqrt((sd1^2/n1) + (sd2^2/n2))
cat ("zcalculated is =",zcal)
pvalue = 2 * (1-pnorm(zcal))
```

→ zcalculated is = -10.84652
[1] 0

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

Scanned with CamScanner

Q2) A study of noise level in a hospital is done; following data is calculated, first sample size = 84, 1st sample mean = 61.2, 1st sample standard deviation = 7.9
2nd sample size = 34, 2nd sample mean = 59.4
2nd sample s.d = 7.8. Test Ho: $\mu_1 = \mu_1$ at 1% level by significance.

```
n1 = 84
n2 = 34
mx1 = 61.2
mx2 = 59.4
sd1 = 7.9
sd2 = 7.8
zcal = (mx1-mx2)/sqrt((sd1^2/n1)+(sd2^2/n2))
zcal
[1] 1.131117
pvalue = 2*(1-pnorm(abs(zcal)))
pvalue
[1] 0.258006
```

∵ pvalue is greater than 0.01, we accept
$H_1 : \mu_0 = \mu_e$

Q.3) from each of two population of oranges test following samples are collected, test whether the proportion of bad orange are equal or not, 1st sample size = 250, 2nd sample size = 200, no. of bad orange in the 1st sample = 44 and 2nd sample = 30

Ho : p1 = p2 against
H1 : p1 ≠ p2

```
n1 = 250
n2 = 200
p2 = 30/200
p1 = 44/250
P = (n1 * p1 + n2 * p2)/(n1+n2)
q = 1-p
zcal = (p1-p2)/sqrt (p* q * (1/n1 + 1/n2))
zcal
```

⇒ [1] 0.8355556
[1] 0.7293581

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

[1] 0.4596896

∴ pvalue is grater than 0.05 we acept the
Ho : P1 = p2

Q.4 random sample of 400 males and 600 females where asked, whether they want ATM nearby. 200 males and 390 females were in favour of proposals. Test the hypothesis that the proportion of males & females favouring the proposal are equal or not at 5% LOS.

$H_0: p_1 = p_2$ against $H_1: p_1 \neq p_2$

n1 =400

n2 =600

p1 =200/400

p2= 390/600

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

p

[1] 0.59

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

∴ p value is greater than 0.05 we reject

$H_0: p_1 = p_2$

Q The following are the two independence samples from the two population, test equality of two population of mean at 5% level of significance.

Sample 1 = 74, 77, 74, 73, 79, 76, 82, 72, 75, 78, 77, 78, 76, 76

Sample 2 = 72, 76, 74, 70, 70, 78, 70, 72, 75, 79, 79, 74, 75, 78, 72, 74, 80

$H_0: \mu_1 = \mu_2$

x1 = c(74, 77, 74, 73, ... 76).

n1 = length (x1)

mx1 = mean (x1)

variance = (n1 - 1) * var (x1) / n1

sd1 = sqrt (variance) → (0.6714177)

x2 = c (72, 76, 74, ... 80)

n2 = length (x2)

mx2 = mean (x2)

variance = (n2 - 1) * var (x2) / n2

sd2 = sqrt (variance) → (0.7850631)

~~tcal =~~

t.test (x, x2)

→ ~~p~~ pvalue = 0.1387

AM

27.01.20

Practical - 8

Small sample test

Q.1 Random sample of 150 observation is given by
80, 100, 110, 105, 122, 70, 120, 110, 101, 84, 83, 95, 89, 107, 125
Do this data support the assumption that population mean is 100?

x = c(80, 100, 110 ...)

Ho u : u = 100

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 0

95 percent confident interval:

91.37775 109.2892

Sample estimates

mean of x

100.3333

∴ p-value < 100, we reject Ho.

Q.2 Two groups of 10 students score following marks.

Grp1: 18, 22, 21, 17, 20, 17, 23, 20, 22, 21

Grp2: 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

Test the hypothesis that there is ~~Ho~~ no significance difference between score at 1% level of significance.

grpa = c (18, 22, 21, 17, ...)
grpb = c (16, 20, 14, ...)
t.test (grpa, grpb)

welch Two sample t.test

data : grpa and grpb

t = 2.2573 , df = 16.376 , p-value = 0.03798

alternative hypothesis : true difference in mean is not equal to 0

95 percent confidence interval :

0.1628205 5.0371795

sample estimates :

mean of x mean of y
20.1 17.5

∵ p-value > 0.01 , we accept Ho.

Q.3 Two types of medicine are used on 5 & 7 patients for reducing their weights . The decrease in the weight after medicines are given below :

MedA = 10, 12, 13, 11, 14

Med B = 8, 9, 12, 14, 15, 10, 9

is there a significance difference in efficiency of medicines ?

data : meda and medb

t = 0.8034 , df = 9.7594 , p-value = 0.4406

alternative hypothesis : true difference in means is not equal to 0.

95 percent confidence interval :

-1.7811 3.7811

Sample estimates :

mean of x mean of y
12 11

∵ p-value > 0.0506 , we accept Ho.

Q.4 The weight reducing diet program is conducted 2
are noted for 10 participants. Test wether the program is
effective or not

Before: 120, 125, 115, 130, 123, 119, 122, 127, 128, 118

After: 111, 114, 107, 120, 115, 112, 112, 120, 119, 112.

Ho: There is no significant difference in weight against H_1
Diet program reduced weight

x = c (120, 125, 115, ...)

y = c (111, 114, 107, ...)

paired t-test

data: x and y.

t = 17, df = 9, p-value = 1

alternate hypothesis: true difference in means is less than 0

95 percent confidence interval:

-Inf 9.416556

Sample estimates

mean of the differences 8.5

∴ p-value > 0.05, we accept Ho.

Q.5 Sample A: 66, 67, 75, 76, 82, 84, 88, 90, 92

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

Test the population means are equal or not

Ho: $\mu_1 = \mu_2$

> sama = c(66, 67, 75, ...)

samb = c(64, 66, 74, ...)

t.test (sama, samb)

welch Two Sample t-test

data: sama and samb

t = -0.639, df = 17.974, p-value = 0.534

alternative hypothesis : true difference in means is not equal to 0.

95 percent confidence interval :

-12.853992 6.85

Sample estimates :

mean of x mean of y

80 83

∵ p-value > 0.05, we accept Ho.

Q) The following are the marks before & after a training program. Test the program is effective or not

Before : 71, 72, 74, 69, 70, 74, 76, 70, 73, 75

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

Ho : There is no significant difference in marks against H1

```
x = c(71, 72, ...)
y = c(74, 77, ...)
t.test (x, y, paired = T, alternative = "greater" )
```

paired t test

data : x and y

t = -4.4691, df = 9, p-value = 0.9992

alternative hypothesis : true difference in means equal to 0

95 percent confidence interval :

-5.0766 Inf

Sample estimates : mean of the differences

-3.6

∵ p-value > 0.05, we accept Ho.

33

practical - 9

Large & small sample test.

Q.1 The arithmetic mean of sample of 100 items from a large population is 52. If standard deviation to test the hypothesis that the population means is 55 against the alternative is more than 55 at 5% level of significance

→ Ho: $\mu_1 = \mu_2$ against Ho: $\mu_1 \neq \mu_2$

n = 100

mx = 52

mo = 55

sd = 7

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

zcal

[1] -4.285

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

[1] 1.82153e-2

∴ pvalue < 0.05, we reject Ho: $\mu_1 = \mu_2$

Q.2 In big city 350 out of 700 males are found to be smokers thus this information supports that exactly half of the males in the city are smokers. Test at 1% level of significance.

→ P = 0.5

p = 350 / 700

q = 1 - p

n = 700

zcal = (p - P) / (sqrt (P * q / n))

zcal

[1] 0

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

pvalue

[1] 1

∴ pvalue > 0.05, we accept Ho

Q.3 1000 articles from a article A are found to have 2% defectives. 1800 article from a second factory are found to have 1% defective. Test that 5% Loss that two factory similar or not

→ Ho : μ1 = μ2 against H1 : μ1 ≠ μ2

n1 = 1000

n2 = 1500

p1 = 0.02

p2 = 0.01

q = 1 - p

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

[1] 0.014

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

[1] 2.0648

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

[1] 0.037

∵ pvalue < 0.05, we reject the Ho.

Q.4) A sample of size 400 was drawn by a sample of 99 test at 5% LOS that the sample comes from the population with mean 100 variance 64.

→ Ho : $\mu_1 = \mu_2$ against H1 : $\mu_1 \neq \mu_2$

n = 400

mx = 99

mo = 100

Var = 64

sd = sqrt(var)

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

zcal

[1] -2.5

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

pval

[1] 0.0124

∵ pvalue < 0.05, we reject the Ho.

Q.5 The flower stems are selected & the heights are found to be (cm) 63, 63, 68, 69, 71, 71, 72 test the hypothesis that the mean height is 66 or not at 1% LOS.

→ Ho : $\mu_1 = \mu_2$

x = c (data)

t.test(x)

one sample t-test

data : x

t = 47.94 df = 6, pvalue = 5.522e-09

alternative hypothesis : true mean is not equal to 0.

95 percent confidence interval :

64.66479 71.62092

sample estimate:

mean of x

68.44286

∴ Pvalue is less than 0.01 we reject Ho: μo = 68

Q1) Two random sample were drawn from 2 normal population & their 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 the same variance at 5% LOS

Ho = population is same

x = c (data)

y = c (data)

f = var.test(x,y)

f

F test to compare two variance

data: x and y

F = 0.70 686 , num df = 8 , denom df = 10 , pvalue = 0.6307

alternative hypothesis: true ratio of variance is not equal to 1

95 percent confidence interval:

0.1833662   3.0360393

Sample estimate:

ratio of variance

0.7068557

∴ Pvalue is ~~greater than~~ 0.05, we accept Ho: σo= The population have same variance

AM

22

practical - 10

Q1 Use the following data to test wether the cleanliness of home and child are independent or not

Cleanliness of Home

Cleanliness of child	Clean	dirty
Clean	70	50
Fairly clean	80	20
Dirty	35	45

→ Ho : Cleanliness of child & cleanliness of home 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
```

∵ pvalue < 0.05 we reject the Ho: CC & CH. ~~are~~ are independent

Q.2 Use the following data to find a vaccination and a particular diseases are independent or not

Diseases

Vaccination	Affected	Not affected
Given	20	30
Not given	25	35

→ Ho: vaccination & diseases are independent

x = c(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

person's chi-squared test with Yate's continuity correction

data: y

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

∵ pvalue > 0.05 we accept Ho: vaccination & diseases are independent

Q.2 perform Anova for the following data.

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

Ho: The means of the variety 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,x2,x3,x4))
names(d) → [1] "values" "ind"
Oneway.test(values ~ ind, data=d, var.equal=T)

one-way analysis of means

data: values and ind
F= ~~5.7278~~ 11.735, num df=3, denom df=9, p-value =0.00183

anova = aov(values ~ ind, data=d)
anova

call:
aov(formula = values ~ ind, data=d)

Terms:

	ind	Residuals
sum of squares	71.06692	18.16667
Deg of freedom	3	9

Residential standard error: 1.420746
estimated effects may be unbalanced.

∴ pvalue < 0.05 we reject Ho: The means of variety are equal

Q.4 The following data give of life of tyres of four brand

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

Test the hypothesis that average life of four brand is same

H0: average life of tires of four brand is 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(15,14,16,18,14,16)
d=stack(list(b1=x1,b2=x2, b3=x3, b4=x4))
names(d)
[1] "value" "ind"
oneway.test(values~ind, data=d, var.equal=T)
        oneway analysis of means

data: values and ind
F= 6.8445, num df = 3, denom df = 20, p-value = 0.002349
anova = aov(values~ind, data=d)
anova
  call:
   aov(formula = values~ind, data=d)

  Terms:
                     ind        Residuals
   sum of squares    91.4381    89.0619
   deg of freedom    3          20
Residual standard error: 2.110231
estimated effects may be unbalanced
```

∴ pvalue < 0.05, we reject H0 : The average life of four type of brand are same.

57

Q. 1000 students of a college are graded according to their IQ and economic condition of their homes. Check that is there any association between IQ and economic condition of their home.

economic condition		I.Q. High	I.Q. Low
	High	460	140
	med	330	200
	Low	240	160

→ 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 = 39.726, df = 2, pvalue = 2.364e-09
```

∵ pvalue < 0.05 we reject Ho: economical condition & IQ are independent

AM
17.2.20

Scanned with CamScanner

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 the 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, ...)

median = 21.5

sp = length (x [x > me])

sn = length (x [x < me])

n = sp + sn

n

[1] 20

pv = pbinom (sp, n, 0.5)

pv

[1] 0.419015

∵ p-value > 0.05 ∴ we accept Ho: population median is 21.5

Q.2 Following are the 10 observation

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

apply sine test to test hypothesis that population median is 625 against the alternative it is greater than 625 at 1% LOS

If the alternative is greater use sn in pbinom.

```
x = c (612, 619, ...)
med = 625
sp = length (x [x>med])
sn = length (x [x<med])
n = sp + sn
n = 10
pv = pbinom (sn, n, 0.5)
pv
[1]  0.0546875
```

∵ p-value is more from 0.01 ∴ we accept H0 : population median is 625

Q.3 10 observation are
36, 22, 21, 30, 34, 25, 20, 22, 20, 18
using sign test that Hypothesis is 25 against the alternative it is less than 25 at 5% los.

```
x = c (36, 32, ...)
med = 25
sp = length (x [x>med])
sn = length ( x [x < med])
n = sp + sn
n
pv = pbinom (sp, n, 0.5)
pv
[1]  0.0546875
```

∵ P-value > 0.05 we accept H0: population median is 25

Q.4 The following are measurements
63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 65
using wilcoxon signed rank test - Test the Hypothesis that the population median is 60 against the alternative it is greater than 60. at 5% LOS.

→ x=c(63, 85, 60, ...)
wilcox.test(x, alt="greater", mu=60)

wilcoxon signed rank test with continuity correction

data: x
V = 68, p-value = 0.06186
alternative hypothesis: true location is greater than 60.
∴ p-value > 0.05 we accept

Q.5 15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 20;
we wilcoxon signed rank test to test the hypothesis the pop median is 20 against the alternative it is less than 20

→ x=c(15, 17, 24, ...)
wilcox.test(x, alt="greater", mu=20)
wilcoxon signed rank test with continuity correction

~~data~~

data = x
V = 43.5, p-value = 0.9232
alternative hypothesis: true location is less than 20.
∴ p-value > 0.05 we accept

Q.6 20, 25, 27, 30, 18

Test the hypothesis that population median is 25 against the alternative it is not 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

AM 9.7.20

Scanned with CamScanner