Semester II

Practical no: 1

Topic: Basic of R software

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

2) This software is used for effective data handling and output storage is posible

3) It is capable of graphical display

4) It a free software

1) $2^2 + \sqrt{25} + 35$

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

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

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

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

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

4) $42 + |-10| + 7^2 + 3 \times 9$

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

5)
```
> 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
> x^2 + y^2
[1] 1300
> sqrt(y^3 - x^3)
[1] 137.8405
> abs(x-y)
[1] 10
```

6)
```
c(2,3,4,5)^2
[1] 4 9 16 25
```

7)
```
c(4,5,6,8)*3
[1] 12 15 18 24
```

8)
```
c(2,3,5,7) * c(-2,-3,-5,-4)
[1] -4 -9 -25 -28
```

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

10.> c(2,3,5,7) * c(1,2,3)
Warning message:
longer object length is not a multiple of shorter object length

11.> c(1,2,3,4,5,6) ^c(2,3)
[1] 1 8 9 64 25 216

12.> Find the sum, product, maximum, minimum of the values
5, 8, 6, 7, 9, 10, 15, 5

Soln > 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
> range (x)
[1] 5 15

13.> x c(1,2,3,4,5,6,7,8) matrix
> 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

14.> $x = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ $y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ -10 & 6 & 12 \end{bmatrix}$

To find $x+y$, $x*y$, $2x+3y$

> 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

	[,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

> 2 * x + 3 * y

	[,1]	[,2]	[,3]
[1,]	8	20	44
[2,]	-2	34	-17
[3,]	36	30	54

Q15) a = 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(a)
[1] 23
> b = table(a)
> transform(b)

a	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)
d = cut(a, breaks, right = FALSE)
e = table(d)
transform(e)

d	Freq
[0, 5)	8
[5, 10)	5
[10, 15)	4
[15, 20)	6

Arun
4/2/19

Practical no: 2

Problem on p.d.f and c.d.f

1) Can the following be p.d.f

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

$\int f(x)\,dx = 1$

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

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

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

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

$\neq 1$

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

$\int f(x)\,dx = 1$

$= \int_0^1 3x^2$

$= 3 \int_0^1 x^2$

$= 3 \cdot \left[\frac{x^3}{3}\right]_0^1$

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

It is a p.d.f

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

$= \int_0^2 \frac{3x}{2}\left(1-\frac{x}{2}\right)$

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

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

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

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

$= \frac{3}{4}[4-0] - \frac{1}{4}[8-0]$

$= 3 - 2$

$= 1$

It is a p.d.f

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 one propobility is negative it not a p.m.f

2.)

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

Since $P(x) \geq 0 \; \forall x$ and $\Sigma P(x) = 1$

It is a p.m.f

x = c(0, 1, 2, 3, 4, 5)
prob = c(0.1, 0.3, 0.2, 0.2, 0.1, 0.1)
sum(prob)

iii.)

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

y = c(10, 20, 30, 40, 50)
probl = c(0.2, 0.3, 0.3, 0.2, 0.2)
sum(probl)
[1] 1.2

$\because$ Sin $P(x) \geq 0 \; \forall x$ and $\Sigma P(x) = 1$

But the value is 1.2

$\therefore$ It is not p.m.f

3.) Find $P(x \leq 2)$, $P(2 \leq x \leq 4)$, 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

i.) $P(x \leq 2) = P(0) + P(1) + P(2)$

$= 0.1 + 0.1 + 0.2$

$= 0.4$

ii.) $P(2 \leq x \leq 4) = P(2) + P(3) + P(4)$

$= 0.2 + 0.2 + 0.1$

$= 0.4$

iii.) P(atleast 4) $= P(4) + P(5) + P(6)$

$= 0.1 + 0.2 + 0.1$

$= 0.4.$

iii) $P(3 < x < 6) = P(4) + P(5)$

$= 0.1 + 0.2$

$= 0.3$

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

```
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
```

$F(x) = 0$ if $x < 0$

$= 0.2$ if $0 \le x < 1$

$= 0.4$ if $1 \le x \le 2$

$= 0.6$ if $2 \le x < 3$

$= 0.7$ if $3 \le x < 4$

$= 0.9$ if $4 \le x < 5$

$= 1.$ if $5 \le x < 6$

ii)

x	1.0	12	14	16	18
p(x)	0.2	0.35	0.15	0.2	0.1

```
probl = c(0.2, 0.35, 0.15, 0.2, 0.1)
cumsum(probl)
[1] 0.20 0.55 0.70 0.90 1.00
```

$F(x) = 0$ if $x < 10$

$= 0.20$ if $10 \le x < 12$

$= 0.55$ if $12 \le x < 12$

$= 0.70$ if $12 \le x < 14$

$= 0.90$ if $14 \le x < 16$

$= 1.00$ if $16 \le x < 18$

AW 4.12.19

11/12/18

Practical no. 3

Topic:- Probability distribution and Binomial distribution

1) Find the c.d.f of the following p.d.f and draw the graph

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

```
> x = c(10, 20, 30, 40, 50)
> prob = c(0.15, 0.25, 0.30, 0.2, 0.1)
> cumsum(prob)
[1] 0.15 0.40 0.70 0.90 1.00
```

$F(x) = 0$ if $x < 10$

$= 0.15$ $10 \le x < 20$

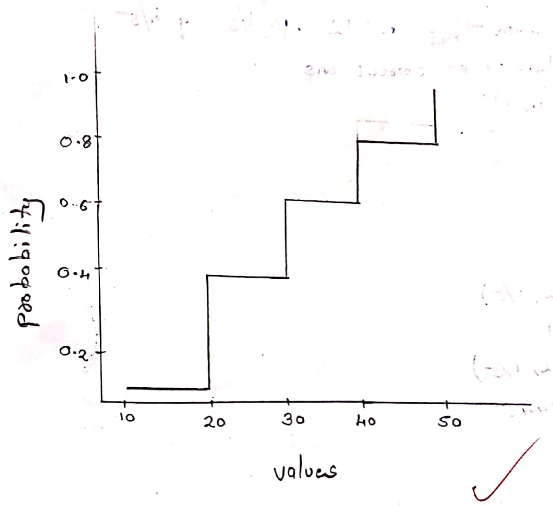
$= 0.40$ $20 \le x < 30$

$= 0.70$ $30 \le x < 40$

$= 0.90$ $40 \le x < 50$

$= 1.00$ $x \geq 50$

graph of c.d.f



2.) Binomial distribution :-

1.) Suppose there are 12 mcq in a test each questions has five option and only one of them is correct. Find a probability of having ~~no~~ one
1.) Five correct ans
2.) Atmost 4 correct answer.

Soln It is given that $n = 12$, $p = 1/5$ $q = 4/5$
x = Total no. of correct ans
$x \sim B(n, p)$

```
> n=12
> p=1/5
> q=4/5
> x=5
> dbinom(5, 12, 1/5)
[1] 0.0531
> pbinom(4, 12, 1/5)
[1] 0.927445
```

2.) There are 10 members in a committe the probability of any member attending a meeting is 0.9. find the probability
1.) 7 member attended
2.) atleast 5 member attended
3.) atmost 6 member attended

$n = 10$
$p = 0.9$
$q = 0.1$
x = Total no of members attended

```
1.) > dbinom(7, n, p)
[1] 0.05739

3.) > pbinom(6, n, p)
[1] 0.0127952

2.) > 1-pbinom(5, n, p)
[1] 0.998365
```

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(0,1,2,3,4,5,6)
> 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
```

$F(x) = 0$ if $x < 0$

$= 0.1$ if $x \le < 1$

$= 0.1$ if $1 \le x < 2$

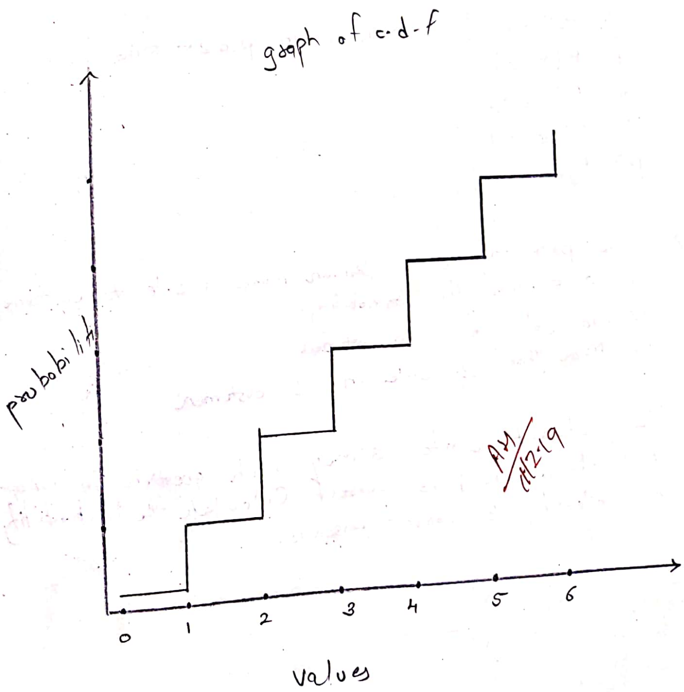
$= 0.2$ if $2 \le x < 3$

$= 0.2$ if $3 \le x < 4$

$= 0.1$ if $4 \le x < 5$

$= 0.2$ if $5 \le x < 6$

$= 1.0$ if $x \geq 6$



Practical:- 4

Binomial Distribution

1) find the complete b.d when $n=5$ $p=0.1$

2) find Probability of exactly 10 success in 100 trial with $p=0.1$

Q.3) X follows B.D with $n=12$, $p=0.25$ find

i) $P(x=5)$

ii) $P(x \le 5)$

iii) $P(x>7)$

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

Q.4) The probability of salesman make a sale to customer is 0.15 find the Probability.

i) No sale for 10 customer

ii) More than 3 sale in 20 customer

Q.5) A student write 5 mcq. Each question has 4 op. out of which 1 is correct. Calculate the Probability for atleast 3 correct answers.

Note:

i) To find the value of x for which the Probability is p, the command is qbino (p,n,p)

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

iii) P(X ≤ x) = pbinom (x,n,p)

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

1)
```
> n=5 , p=0.1
> dbinom (0:5, 5, 0.1)
[1] 0.59049   0.32805   0.7290   0.00810
```

2)
```
> n=100
> p=0.1
> x=10
> dbinom (10, 100, 0.1)
[1] 0.1318653
```

3)
```
> n=12
> p=0.25
```

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

ii.) n = 12
> p = 0.25
> x = 5
> pbinom (5, 12, 0.0.25)
[1] 0.9455978

iii.) > n = 12
> p = 0.25
> x > 7
> 1 - pbinom (7, 12, 0.25)
[1] 0.0027815-1

iv] > n = 12
> p = 0.25
> 5 < x < 7
> dbinom (6, 12, 0.25)
[1] 0.04014945

4.) i.) n = 10
> p = 0.15
> x = 0
> dbinom (0, 10, 0.15)
[1] 0.1968744

ii.) > n = 20
> p = 0.15
p (x > 3) = 1 - P(x ≤ 3)
> 1 - pbinom (3, 20, 15)
[1] ...

5.) > n = 5
> x = 3
> P = 1/4 = 0.25
> 1 - P(x ≤ 2)
> 1 - pbinom (2, 5, 0.25)
[1] 0.1035156

Q6.) X follow binomial distribution with n = 10, P = 0.4 Plot the graph of p.m.f & c.d.f

Sol.
> n = 10
> p = 0.4
> x = 0:n
> prob = dbinom (x, n, p)
> comprob = pbinom (x, n, p)
> d = data.frame (x. values = x, "Probability" = prob)
> Print (d)

X. values	Probability
0	0.0060466176
1	0.0403107840
2	0.1209323520
3	0.2508226560
5	0.2006581248
6	0.1114767360
7	0.0424673280
8	0.0106168320
9	0.0015728640
10	0.0001048576

> Plot (x, prob, "h")
> Plot (x, comprob, "s")

Practical : 05

5.) Normal Distribution

① $P[X = x]$: dnorm (x, μ, σ)
② $P[X \leq x]$: 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) - pnorm (x_1, μ, σ)
⑤ To find the value of k so that
$P[X \leq k] = p$; qnorm (p_1, μ, σ)
⑥ To generate 'n' random numbers
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.) $P(X \leq K) = 0.7$; $K = ?$

Sol^n

①
```
> a = pnorm (40, 50, 10)
> cat ("p(x<=40)=", a)
p(x<=40)= 0.1586553
```

②
```
> b = 1 - pnorm (55, 50, 10)
> cat ("p(x>55)=", b)
p(x>55) = 0.3085375
```

(5) > c = pnorm(60,50,10) - pnorm(42,50,10)
> cat("p(42 ≤ x ≤ 60): ", c)
p(42 ≤ x ≤ 60): 0.6294893

(4) > d = qnorm(0.7,50,10)
> cat("p(x<=k)=0.7 , k=", d)
p(x<=k)=0.7 , k= 55.24401

(2) $X \sim N(\mu = 100, \sigma^2 = 36)$

Find: i.) $P(X \le 100)$

ii.) $P(X \le 95)$

iii.) $P(X > 115)$

iv.) $P(95 \le X \le 105)$

v.) $P(X \le k) = 0.4$, k = ?

Soln: i.) > a = pnorm(110, 100, 6)
> cat("p(x ≤ 110): ", a)
p(x ≤ 110): 0.9522096

ii.) > b = pnorm(95, 100, 6)
> cat("p(x≤95): ", b)
p(x ≤ 95): 0.2023284

iii.) > c = 1 - pnorm(115, 100, 6)
> cat("p(x>115): ", c)
P(x>115) = 0.006209665

iv.) > d = pnorm(105, 100, 6) - pnorm(95, 100, 6)
> cat("p(95 ≤ x ≤ 105): ", d)
p(95 ≤ x ≤ 105): 0.5953432

v.) > e = qnorm(0.4, 100, 6)
> cat("p(x ≤ k) = 0.4, k= ", e)
p(x≤k): 0.4 , k= 98.47992

3.) Generate 10 random number from normal distribution with mean = 60 & S.D = 5 also calculate sample mean, variance median and standard deviation

Soln: > x = rnorm(10, 60, 5)
> x

61.99112	55.80205	55.07745	50.96845	58.17959
58.15321	61.55288	71.91887	66.57875	66.94407

> am = mean(x)
> am
60.01123
> me = median(x)
> me
59.86623
> n = 10
> variance = (n-1) * var(x)/n
> variance
44.45974
> sd = sqrt(variance)
> sd
6.667814

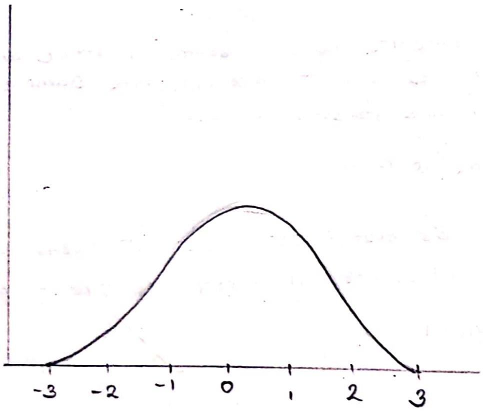
Scanned with CamScanner

④ Draw the graph of standard distribution

Sol^n > x = seq(-3,3,by=0.1)

> y = dnorm(x)

> plot (x,y,xlab="x values", ylab="probability", main="Standard normal Graph")



Practical no: 6

Topic : Z Distribution

1) Test the hypothesis (Ho) μ=10 against H₁: μ ≠ 10. A sample of size four hundred be selected which gives a mean 10.2 & standard deviation 2.25. Test the hypothesis at 5% level of significant

Sol^n mo = mean of population = 10   mx = mean of sample 10.2   sd = mean of standard = 2.25   n = sample set = 400

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

> cat ("zcal is =", zcal)

zcal is 1.77778

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

> pvalue

[1] 0.07544036

Since 0.075 is more than 0.05 ∴ we will accept the Ho.

2) Test the hypothesis μ=75 H₁: μ ≠ 75. A sample of size 100 is selected and sample mean is 80 with sd=3. Test the hypothesis at 5% level of significant

Sol^n mo = 75, mx = 80, sd = 3, n = 100

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

> cat ("zcal is =", zcal)

zcal is : 16.66667

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

> pvalue

[1] 0

3) Test the hypothesis (H_0): $\mu = 25$ against $H_1: \mu \neq 25$ at 5% level of significant. The following sample of 30 is selected.

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

Sol)
```
> x = c(20, 24, 27 . . . 18)
> n = length(x)
> mx = mean(x)
[1] 26.06667
> variance = (n-1) * var(x)/n
> variance
[1] 52.99556
> sd = sqrt(variance)
> sd
[1] 7.279805
> z cal = (m-25) / (sd / sqrt(n)))
> cat ("zcal is =", zcal)
zcal is = 0.8025454
> pvalue = 2 * (1-pnorm(abs(zcal)))
> pvalue
[1] 0.4222375
```

4) Experience has show that 20% students of a college are smoke a sample of four hundred student reveal that out of four hundred only 50 smoke test the hypothesis that the experience the give the correct porportion or not.

```
> P = 0.2
> Q = 1-P
> p = 50/400
> p
[1] 0.125
> n = 400
> zcal = (p-P) /(sqrt (P*Q/n))
> cat("zcal is", zcal)
zcal is -3.75
> pvalue = 2* (1- pnorm (abs (zcal)))
> pvalue
[1] 0.0001768346
```

Scanned with CamScanner

5) Test the hypothesis Ho: P = 0.5 against H1: P ≠ 0.5
A sample of 200 is selected and the sample proportion is cal p = 0.56 test the hypothesis 1% level of significance

Sol:
```
> n: 200
> p = 0.56
> P = 0.5
> Q : 1 - P
> z cal : (p-P)/(sqrt (P*Q/n))
> cat("z cal is ", zcal)
z cal is 1.697056
> pvalue : 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.08968602
```

AM
22-01-20

Practical : 7 49

Topic:- large Sample Tests:

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

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

Sol: Ho: The noise level are same
```
n1 = 84
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 =", z)
zcalculated is = 1.273682
> pvalue = 2*(1-pnorm(abs(z)))
> pvalue
[1] 0.04550026  ~~2.???~~ 0.2027
```
Since pvalue < 0.05, we reject Ho at 5% level of significance.

2.) Two random sample of size 1000 and 2000 are drawn from two population with a mean: 67.5 and 68 respectively & with the same standard deviation of 2.5 test the hypothesis The mean of two population are equal.

Ho: Two population are equal

n1 = 1000

n2 = 2000

mx1 = 67.5

my1 = 68

sdx = 2.5

sdy = 2.5

y = (mx1 - my1) / (sqrt((sdx^2/n1) + (sdy^2/n2))

y

[1] -5.163978

cat("ycalculated is=", y)

ycalculated is = -5.163978

p value = 2*(1-pnorm(abs(y)))

pvalue

[1] 2.417564e-07

Since pvalue < 0.05 we reject the Ho of 5% level significant.

3.) In a first year class 20% of a random sample of 400 students had defect eye sight in sy class 15.5% of 500 students sample had the same defect, is the difference of proportion is same?

Soln Ho: The proportion of population are equal

n1 = 400

n2 = 500

p1 = 0.2

p2 = 0.155

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

q = 1-p

p

[1] 0.175 ~~0.18 75378~~

> q

[1] = 0.822225

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

> z

[1] 1.766547

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

~~Since pvalue > 0.05~~

[1] 0.0774

Since pvalue < 0.05 we ~~reject~~ accept the Ho at 5% level of significiat

4) From each of the boxs of the apple a sample size 200 is collective. It is found that there are 44 bad apple in the first sample and 30 in the second sample test the hypothesis that the two boxes are equivalent in term of number of bad apple.

Soln

Ho.. the two boxes are equivalent.

p1 = 44/200

p2 = 30/200

n1 = 200

n2 = 200

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

p

[1] 0.185

q = 1 - p

[1] 0.815

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

z

[1] 1.802741

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

[1] 0.07142

∴ pvalue < 0.05 we ~~reject~~ accept the Ho of 5% level of significant

5) In MA class out of a sample of 60, mean height is 63.5 inch with a sd 2.5 in a Mcom class out of 500 students mean height 69.5 inch with a sd 2.5. test the hypothesis that the mean of MA class & Mcom are same

Soln Ho: hypothes of the mean of MA & Mcom class are same

> n1 = 60

> n2 = 500

> mx = 63.5

> my = 69.5

> sd1 = 2.5

> sd2 = 2.5

> y = (mx - my)/sqrt((sd1^2/n1) + (sd2^2/n2))

> y

[1] -12.53359

> cat("y calculated is =", y)

> y calculated is = -12.53359

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

pvalue

[1] : 0

Since pvalue < 0.05 we reject the Ho of 5% level significant

AM 5.2.20

13

Practical: 8

Topic :- Small sample test

1) The flower 10 are selected and the height are found to be 63, 63, 68, 69, 71, 71, 72, cms. Test the hypothesis that the mean height is 66 cm or not at 1% inrerent.

Sol: Ho:
Mean = 66 cms

```
> 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
```

∴ pvalue < 0.01 we reject the Ho of 1% level of significant.

52

2) Two random sample where drawn from two different population
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 differentence b/w the 2 population mean at 5% los.

Sol: Ho: There is no difference in the population means

```
x = c(8,10,12,11,16,15,18,7)
y = c(20,15,18,9 ,8,10,11,12)
        Welch Two Sample t-test
data: x and y
t = -0.36247, df = 13.837, p-value = 0.7225
alternative hypothesis: true difference in means is not
 equal to 0
95 percent confidence interval:
 -5.192719 3.692719
sample estimates:
 mean of x mean of y
   12.125    12.875
```

pvalue > 0.05 we accept the Ho of 5% level of significant

52

3) Following are the weight of 10 people before and after a diet program. Test the Ho that the diet program is effective or not

Before(kg)	100	125	95	96	98	112	115	104	109	
After(kg)	95	80	95	98	90	100	110	85	100	

Sol Ho: The diet program is not effective

> x: c(100, 125, 95, 96, 98, 112, 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 & y
t: 2.6089, df: 9, pvalue: 0.9858
alternative hypothesis: true difference in means is less than
95 percent confidence interval:
-Inf 18.72908
sample estimates
mean of the differences
11

∴ Pvalue is greater the 0.05 the we accept the 5% level of significant.

4) The marks before and after a training program are given below

Before	20	25	32	28	27	36	35	25
After	30	35	32	37	37	40	40	23

Test the hypothesis the training program is or not

Sol Ho: The training program is not effect

> 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 diff in means is greater than 0
95 percent confidence interval:
-8.967399 Inf
sample estimates:
mean of the differences
-5.75

∴ Pvalue > 0.05 We accept the 5% level of significant.

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

Soln Ho : The varience of a two 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 & y
F= 0.70686, num df= 8, denom df= 10
pvalue = 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

pvalue < 0.05 we accept the 5% level of significant.

6.) The arthematic mean of a sample of 100 observation is 52 if the sd is 7 test the hypothesis that the population is 55 are not at 5% los

Soln Ho : The population mean equal to 55

n= 100
mx= 52
mo = 55
sd= 7

zcal = (mx-mo) /(sd/sqrt(n)))
> cat("zcal is =", zcal)
zcal is : -4.28574
> pvalue = 2 * (1-pnorm(abs(zcal)))
> pvalue
[1] 1.82153e-05

pvalue is <0.05 we reject the 5% level of significal

A.Vr
12.2.24

Practical: 9

Chi square distribution and ANOVA Ban

1) Use the following data to wheather the cleanliness of home depends upon the child

Codn of Home

		Clean	Dirty
Codn of Child	Clean	70	50
	fairly	80	20
	Dirty	35	45

H°: codn of the home and 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 chi-squared test

data: y

∴ P value is less than 0.05 we reject Ho at 5%.

2) Table below show a relation between the performance of mathematics and computer of CS student.

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

Ho: Performance b/w Maths and Computer are independent

```
> x = c(56,47,14,71,163,42,12,38,85)
> m=3
> n=2
> y= matrix(x,nrow=m,ncol=n)
        [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-square = 145.78 df=4 , p-value < 2.2e-16 ,

pvalue < 0.05 we reject Ho at 5% less

3) Perform ANOVA for the following data

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

Ho: The mean of varity A, B, C, D are

x1 = c(50, 52)
x2 = c(53, 55)
x3 = c(60, 58, 57, 56)
x4 = c(52, 54, 54, 55)
d: stack c list (b1 = x1, ~~b2~~ = x2, b3 = x3, b4 = 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 = 11.735, num df = 3, denom df = 9, p-value: 0.00
> anova: aov [values ~ind, data = d]
> Sumary (anova)

	Df	Sum sq	Mean sq	F value	Pr (>F)
ind	3	71.06	23.688	11.73	0.00183
Residuals	9	13.17	2.019		

∴ PV < 0.05 we reject Ho at 5% los

) Perform ANOVA for the following data

types	observation
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

Ho: The mean of variety A, B, C, D 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))
> name (d)
[1] "values" "ind"
> oneway test (value ~ind, data = d, var equal = 3
one way analysis of means
data: values and ind
F = 0.667, nom df = 3, denom df = 8, p-value: 0.1139
> annova = aov (values, ~ind, data = d
> summary (annova)

	Df	sum sq	mean sq	F value	Pr (>F)
ind	3	18	6.00	2.667	0.119
Residuals	8	18	2.25		

p. value > 0.05 we accept Ho at 0.05 at 5% los
> x = read. csv ("C: / users / Administrator / Desktop
> x

Practical: 10

Topic: Non Parametri · Test

1) Following are the amount of sulphur oxide emitted by a factory

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 the population median is 21.5 against the alternative it is less than 21.5

Sol^n

Ho: Population median equal to 21.5
H1: It is less than 21.5
x = c(
m = 21.5
sp = length(x[x>m])
sn = length(x[x<m])
n = sp+sn
pv = pbinom(sp,n,0.5)
pv = 0.411

pvalue > 0.05 we accept Ho at 5% level of significant

Note: If the alternative is greater than median pv=pbinom(sn,

2) For the observation 12, 19, 31, 28, 43, 40, 55, 49, 70, 63 apply sign test to test population median is 25 against the alternative it is more than 25

Sol^n Ho: population median equal to 25
H1: It is more than 25
x = c(12, 19, 31, 28, 43, 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
[1] 0.0546875

pvalue is less than 0.05 we reject

58

3) For the following data
60, 65, 63, 89, 61, 71, 58, 51, 48, 66 test the Hypothesis using wilcoxon sign rank test for testing the hypothesis The median is 60 against the alternative it is greater.

Sol:

Ho: The median is 60
H1: It is greater than 60
x=c(60, 65, 63, 89, 61, 71, 58, 51, 48, 66)
mu=60
wilcox.test(x, "greater", mu=60)

Wilcoxon signed rank test with continuity correction
data: x
V=29, p-value=0.2386
alternative hypothesis: true location is greater
pvalue < 0.5 we reject the Ho

Note: If the alternative is less
wilcox.test(x, alter="less", mu=60)
If the alternative is not equal to
wilcox.test(x, alter=2.sided)

4) Using wilcoxn test test the hypothesis median is 12 against the alternative is less than 12
12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20

Sol:

Ho: The median is 12
H1: It is less than 12
> x=c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20)
> wilcox.test(x, alter="less", mu=12)

Wilcoxon signed rank test with continuity correction
data: x
V=25, p-value=0.2521
alternative hypothesis: true location is less than 12.
p value value is

AM 4/3/20