SEMESTER - II

Practical - 1

Basic of R software.

(i) R is a software for data analysis and statistical computing.
(ii) It is a software by which effective data handling and outcome storage is possible.
(iii) It is capable of graphical display.
(iv) It's a free software.

Problems.

1) $2^2 + |-5| + 4 \times 5 + 6/5$
$\Rightarrow$ 2^2 + abs(-5) + 4x5 + 6/5
= 30.2

2) $x = 20$
$y = 2x$
$z = x + y$, $\sqrt{z}$
$\Rightarrow$ x = 20
y = ~~2+x~~
z = x + y
sqrt(z)
= 7.74

3) $x = 10$
$y = 15$
$z = 5$
a) $x + y + z$
b) xyz

c) $\sqrt{xyz}$

d) round of $\sqrt{xyz}$

$\Rightarrow$ $x = 10$

$y = 15$

$z = 5$

$a = 10 * 15 * 5$

$b = \sqrt{a}$

round (a).

A vector in R software is denoted by the syntax C.

1) (2,3,5,7)^2

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

4 9 25 49

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

4 27 25 343

3) c(2,3,5,7,9,11)^ c(2,3)

4 27 25 343 81 1331

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

1 8 81 16 125 1296

5) c(2,3,5,6)* 3

6 9 15 18

6)
⇒ c(2,4,6,8) * c(-2,-3,-5,-7)
[1] -4 -12 -30 -56

7)
⇒ x = c(2,4,6,8)
x + 10
[1] 12 14 16 18

8)
⇒ c(2, 4, 6, 8) + c(-2,-3,-1,0)
[1] 0 1 5 8.

→ sum product

Q] Find the sum, product, sqrt of the following sum and product for the following vali
(4, 9, 2, 5, 7, 8, 3, 6, 15, 12, 10, 9, 8, 13, 14)
x = c(4, 9 ....)
y = sum(x)
z = Prod(z)
y ^ 0.5
y
[1] 125
z
[1] 5
> y ^ 0.5
[1] 11.18034

Adding of Matrix

```
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

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

x+y

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

y

	[,1]	[,2]	[,3]
[1,]	6	11	9
[2,]	4	12	7
[3,]	5	8	4

x* y

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

x* 2

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

A-b
2.12.9

y* 2

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

BEO 2^0.5
[1] 2925632

Q13) Find the sum, product, max, min, values of
c(2,8,9,11,10,7,6)^2
x = c(2,8....)^2
[1] 4 64 81 121 100 49 36
sum(x)
[1] 455
Prod(x)
[1] 44 25 9 2748400

max(x)
[1] 121
min(x)
[1] 4.

Q Matrix
x ← matrix (nrow=3, ncol=4, c(2,6,7,8,9,4,5,10,
x

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

n=10, P= 0.6, q = 0.4

dbinom(7,10,0.6)

[1] 0.2149908

dbinom(4,10,0.6)

[1] 0.1114767.

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

2) Suppose there are 12 mcq's in an english question papers. Each question has 5 answers and only 1 is correct. Find the probability

(i) 4 correct answers

(ii) Atmost 4 correct answers

(iii) Atleast 3 correct answers

→ n= 12, P= 1/5

(i) P(X=4)

> dbinom(4,12,1/5)

[1] 0.13287

(ii) Pbinom(4,12,1/5)

[1] 0.927

(iii) 1-Pbinom(2,12,1/5)

[1] 0.4416

3) Find the complete binomial distribution where a=5, b= 0.1

> dbinom(0,5,0.1)

[1] 0.59049

PRACTICAL - 2
Binomial Distribution

n = Total no of trials
p = P(success)
q = P(failure)
x =

$P(x) = {}^nC_x \, p^x q^{n-x}$; $x = 0, 1, \ldots$
$E(X) = np$
$V(X) = npq$

dbinom (x,n,p)	P(x)
n * p	E (x)
pbinom (x,n,p)	

Problems.

1) Toss a Coin 10 times with probability (heads = 0.6]
let X be the number of heads.
i) Find the probability of ① 7 heads, ② 4 hea
③ Atmost 4 heads, ④ atleast 6 heads
⑤ no head ⑥ All heads
Also find exceptation and variance.

> dbinom (1,5,0.1)
[1] 0.328
> dbinom (2,5,0.1)
[1] 0.0729
> dbinom (3,5,0.1)
[1] 0.0081
> dbinom (4,5,0.1)
[1] 0.00045
> dbinom (5,5,0.1)
[1] 1e-05

4) Find the probability of exactly 10 success out of 100 trials with P=0.1

→ Pbinom (100,10,0.1)
[1] 1

5) X follows binomial distribution with n=12 & p=0.25 Find (i) $P(X \le 5)$ (ii) $P(X>7)$ (iii) $P(5<X<7)$

> Pbinom (5,12,0.25)
[1] 0.94755
> 1- Pbinom (7,12,0.25)
[1] 0.0027
> Pbinom (6,12,0.25)
[1] 0.985

6) There are 10 members in a Committee. Probability of any member attending a meeting is 0.9. What is the probability 7 members are attending the meeting?

> Pbinom (7,10,0.9)
[1] 0.070190

7) A
of
on
Cus
wi

⇒ >
[1]

8)

Scanned with CamScanner

7) A salesman has 20% probability of making a sale to a customer. On a typical day he will meet 30 customer. What minimum no of sale will he make with 88% probability?

⇒ > qbinom
[1] 9

8) For n = 10, p = 0.6. Find the binomial probability of that & plot the graphs of pmf & cdf

```
> n=10
> p = 0.6
> x = 0:n
> bp = dbinom(x,n,p)
> d = data.frame("x.values" = x, "Probability" = bp)
> d
```

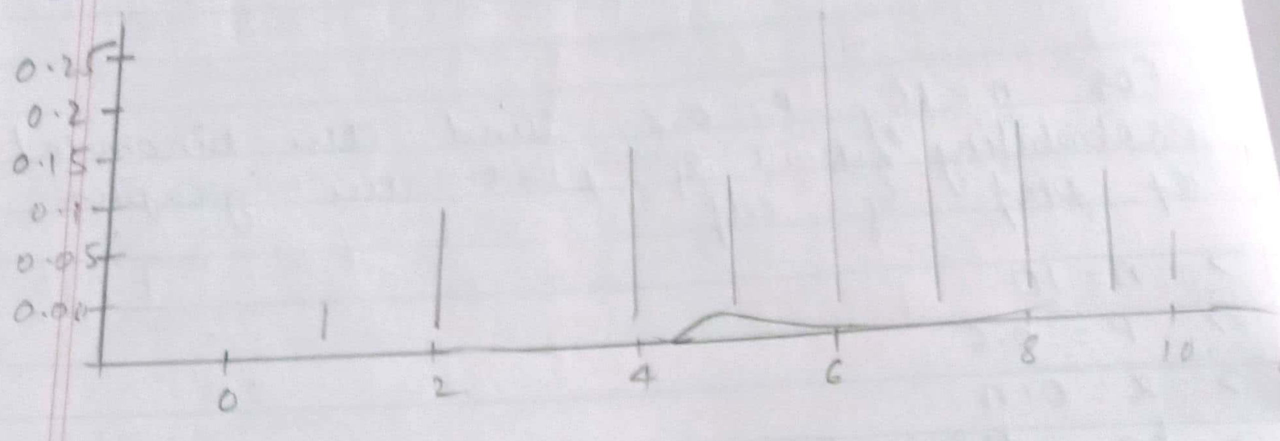
	x.values	Probability
1	0	0.00104
2	1	0.0015
3	2	0.0010
4	3	0.42467
5	4	0.11147

of

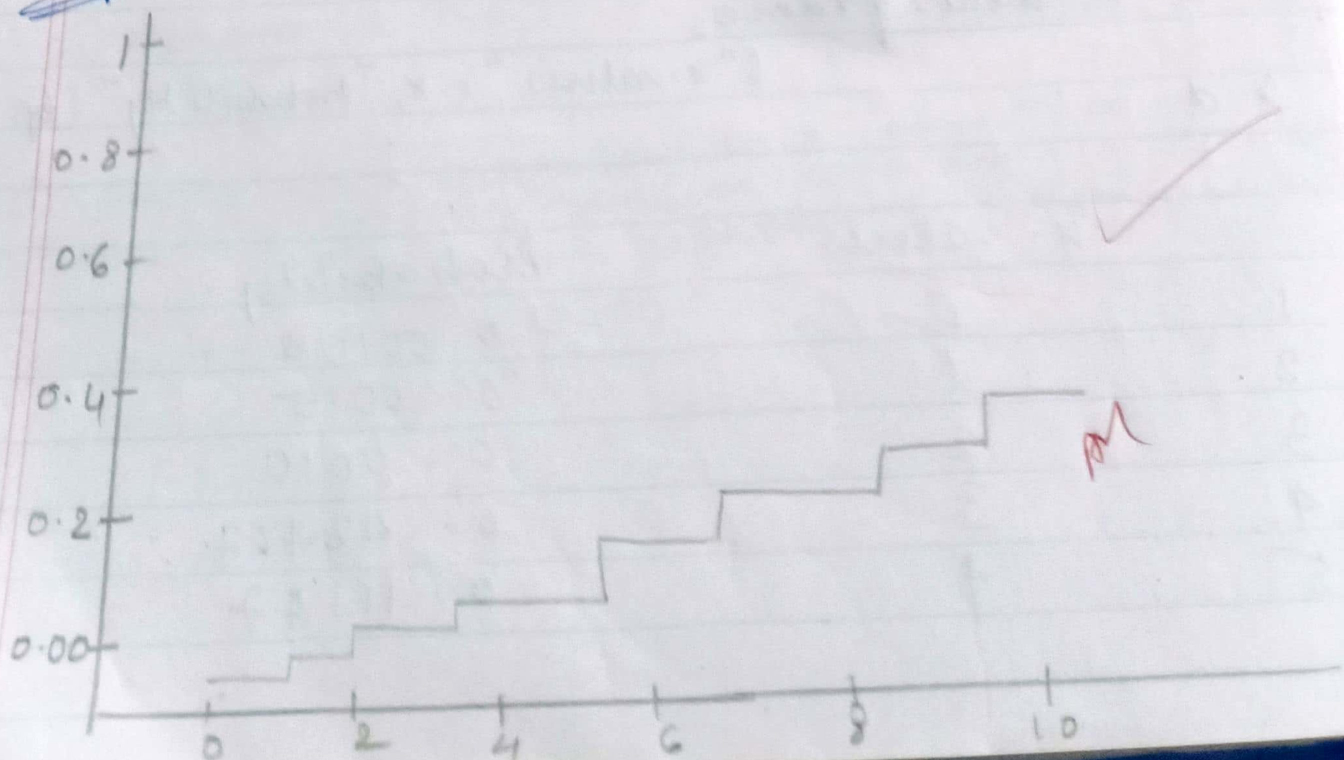
P = 0.25

		0.2006
	5	0.250822
6	6	0.21499
7	67	0.12093
8	8	0.04031
9	87	0.00604
10	10	
11		

pmf



cdf



Q2. Following is a pmf of (x)

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

Find mean & variance.

x_i	$P(x)$	$x P(x)$	$x^2 P(x)$
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

$\Sigma x P(x) = 3.45 \quad \Sigma x^2 P(x) = 13.55$

Mean $= E(X) = \Sigma x P(x) = 3.45$

Var $= V(X) = \Sigma x^2 P(x) - [\Sigma(x)]^2$

$= 13.55 - (3.45)^2$

$= 1.6475$

Using R

```
> x = seq(1,5,1)
Prob = c(0.1, 0.15, 0.2, 0.3, 0.25)
a = x * Prob
sum(a)
sum(b)
mean = sum(a)
mean
[1] 3.45
b = (x^2) * Prob
var = sum(b) - mean^2
var
[1] 1.6475
```

PRACTICAL- 3

Q1. check the following are pmf or not.

1)

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

condition for it is 1) $0 \leq P(x) \leq 1$

2) $\sum_{x} P(x_i) = 1$

Since, it dosen't satisfy the given condition

∴ It is not a pmf.

2)

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

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

Hence it is satisfy both the condition.

∴ It is a 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 is not satisfing the Condition

It is not a pmf.

Find mean and variance of x.

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

Using R

```
x = seq(5,25,5)
Prob = c(0.1, 0.3,0.2, 0.25,0.15)
a = x* Prob
mean = sum(a)
mean
[1] 15.25
b = (x^2)* Prob
Var = sum(b)-mean^2
Var
[1] 38.6875
```

Q4) i) Find cdf of the following pmf and draw the graph of cdf

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
[1] 0.4, 0.7 0.9 1.0
```

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

```
> Plot (x, 9, "s")
```

Q2.

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

```
> x = seq(0,8,2)
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
```

$$f(x) = \begin{cases} 0 & \text{iff } x < 0 \\ 0.2 & \text{iff } 0 \leq x < 2 \\ 0.5 & \text{iff } 2 \leq x < 4 \\ 0.7 & \text{iff } 4 \leq x < 6 \\ 0.9 & \text{iff } 6 < x < 8 \\ 1.0 & \text{iff } x \geq 8 \end{cases}$$

```
> plot (x, a, "s")
```

PRACTICAL - 04

BD

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

$n=8,\ p=0.6,\ q=0.4$ → given

1. $P(X=7) = \binom{8}{7}(0.6)^7(0.4)$

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

$= 0.08957$

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

$= {}^8C_0 (0.6)^0 (0.4)^8 + {}^8C_1 (0.6)^7 (0.4)^{8-1} + {}^8C_2 (0.6)^2$

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

$= 0.1736704$

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

$= 0.04128768 + 0.12386304$

$= 0.161515012.$


1.0
1.5
2.0
2.5
3.0
3.5
4.0
x

PRACTICAL - 05
Normal distribution

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

$P(X = x)$ = dnorm (x, μ, σ)

1. $P(X \leq x)$ = Pnorm (x, μ, σ)
2. $P(X > x)$: 1 - Pnorm (x, μ, σ)
3. To find the value of k, so that $P(X \leq k) = P$
4. qnorm (P, μ, σ)
5. To generate a random sample of size n,
rnorm(n, μ, σ)

Q1. a random variable X follows normal distribution with $\mu = 10$, $\sigma = 2$, find :
i) $P(X \leq 7)$ (ii) $P(X > 12)$ (iii) $P(5 \leq X \leq 12)$
iv) $P(X < K) = 0.4$

```
> P1 = Pnorm (7,10,2)
> P1
[1] 0.0668072
> P2 = 1- Pnorm (12,10,2)
> P2
[1] 0.1586553
> P3 = Pnorm (12,10,2) - Pnorm (5,10,2)
> P3
[1] 0.8351351
> K = qnorm (0.4, 10,2)
> K
[1] 9.493306
```

Q2. $X \sim N(100, 36)$

(i) $P(X \leq 110)$ (ii) $P(X > 105)$ (iii) $P(X \leq 92)$

(iv) $P(95 \leq X \leq 110)$ (v) $P(X < K) = 0.9$

```
> a = Pnorm (110, 100, 6)
> a
[1] 0.9522096
> b = 1- pnorm (105, 100, 6)
> b
[1] 0.2023284
> c = Pnorm (92, 100, 6)
> c
[1] ~~a Rao~~ 0.09121122
> d = Pnorm (100, 110, 6) - Pnorm (95, 100, 6)
> d
[1] 0.7498813
> K = qnorm (0.9, 100, 6)
> K
[1] 107.6893
```

Q3. Generate 10 random sample & find sample mean, median, variance & S.D

042

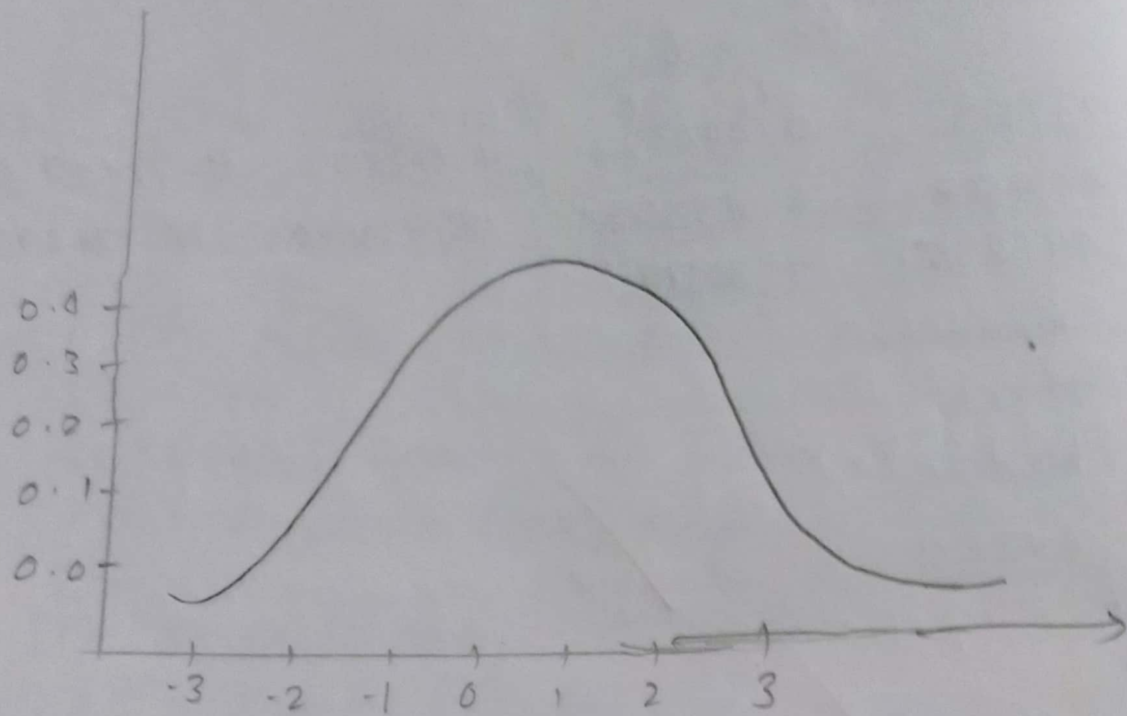
```
> x = rnorm(10, 10, 3)
> x
[1] 13.937393   12.040522   11.512806   9.774476
    5.884007    8.433805    15.813018   14.174675
    6.495235    9.305191

> am = mean(x)
[1] 10.73711
> mc = median(x)
[1] 10.64364
> n = 10
[1] 10
> var = (n-1) * var(x)/n
> var
[1] 9.995247
> sd = sqrt(var)
> sd
[1] 3.161526
```

Q4) Plot the standard curve

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

SDO



Q5. X ~ N (50, 100)

Find : (i) P(x ≤ 60) (ii) P(x > 65) (iii) P(45 ≤ x ≤ 60)

> Pnorm (60, 50, 10)

[1] 0.8413447

> Pnorm (60, 50, 10) - Pnorm (45, 50, 10)

[1] 0.5328072

> 1- Pnorm (65, 50, 10)

[1] 0.0668072.

PRACTICAL NO : 6
Z & T distribution

1) Test the hypothesis; $H_0 : \mu = 20$ against $H_1 : \mu \neq 20$
A sample of size 400 is selective & the sample mean is 20.2 & the standard deviation is 2.25. Test at 5% level of significance

```
> Mo = 20 ; mx = 20.2; sd = 2.25 ; n = 400
> mo ; mx; sd; n
[1]    20
[1]    20.2
[1]    2.25
[1]    400

> zcal =(mx-mo) / (sd / sqrt(n))
> zcal
[1] 1.777778
> cat (" Zcalculated is =", zcal)
 zcalculated is = 1. 777778 >
> Pvalue = 2* (1- pnorm (abs (zcal)))
> Pvalue
[1]  0.0754
```

Since the value is greater then 0.05 we accept the value of μ.

2) We want to test the hypothesis $H_0 : \mu = 250$ against $H_1 : \mu \neq 250$. A sample of sye 100 has a mean of 275 & S.D is 30 Test the hypothesis at 5% level of significa

```
> mo= 250; mx = 275; sd= 30; n = 100
> mo; mx; sd; n
[1] 250
[1] 275
[1] 30
[1] 100
> zcal = (mx-mo)/(sd/sqrt(n))
> zcal
[1] 8.3333
> cat("zcalculated is =", zcal)
zcalculated is = 8.3333 >
> Pvalue = 2*(1-pnorm(abs(zcal)))
> Pvalue
[1] 0
```

We reject H_0 : 4

3] We want to test the hypothesis $H_0: P =$ against $H_1: P \neq 0.2$ (P = Population proportion). A sample of 400 is selected and the sample proportion is calculated 0.125. Test the hypothesis at 1 %.

```
> P = 0.2
> Q = 1-P    [1] 0.8
> p = 0.125
> n = 400
```

zcal = (p - P)/(sqrt(P * q/n))
> zcal
> -3.75
[1]

3) In a big city 325 men out of 600 men were found to be self-employed thus this information supports the conclusion that half of the men in the city are self employed ? (P= 0.5)

> p = 325/600
> p
[1] 0.541667
> n = 600
> Q = 1 - P
> Q
[1] 0.5
> zcal = (p - P)/(sqrt(P*Q/n))
> zcal
[1] 2.041241
> Pvalue = 2* (1- Pnorm (abs (zcal)))
> Pvalue
[1] 0.04122683

0.2

4) Test the hypothesis H0: μ = 50 against H1: μ≠50
a sample of 30 is collected φ
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

```
> s= c(50, 49, 52, 44, 45, 48, 46, 45, 49, 45, ...
      47, 44, 59, 60, 61, 41, 52, ...
      55, 54, 46, 58, ...
      56, 46, 45, 48, 49 )

> s
[1]    " ---
> length(s)
[1] 30
> mx= mean(s)
> mx
[1] 49.3333
> sd= sd(s)
> sd
[1] 5.658886
> m0 = 50
> var =(n-1)* var(s) /length(s)
> var
[1] 639-3423
> n= length(s)
> n
[1] 30
> zcal = (m0 - mx) /sd /(sqrt(n))
> zcal
[1] 0.02150885
> Pvalue = 2*(1-Pnorm(abs(zcal)))
> Pvalue
[1] 0.9828397
```

3

AM
20-01-20

PRACTICAL NO : 07.
Large sample test.

1) Two random sample of size 1000, 2000 are drawn from 2 population respectively. Test the hypothesis that the 2 population means are equal or not at 5% level of significance with the SD 2 & 3 the 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 (" z calculated is :", zcal)
z calculated is : -10.84652.
> Pval = 2* (1- pnorm (abs (zcal)))
> Pval
[1] 0
```

∴ Pval < 0.05, we reject H_0 $\mu_1 = \mu_2$

2) A study of noise level in two hospitals is demonstrated. following data is calculated first sample is

⇒ > n1 = 84

> n2 = 34

> mx1 = 61.7

> mx2 = 59.2

> sd1 = 7.9

> sd2 = 7.8

> zcal = (mx1 - mx2)/sqrt((sd1^2/n1) + (sd2^2/n2))

[1] 1.131111

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

> Pvalue

[1] 0.2580

∴ p-value is > 0.01, we accept the $H_0: \mu_1 = \mu_2$

3) from each of the 2 population of oranges the following samples are collected whether the proportion of bad oranges are equal or not

sample size = 250
no. of bad oranges = 44

2) sample size = 200
no of bad oranges = 30

Ans) $H_0: P_1 = P_2$ against $H_1: P_1 \neq P_2$

```
> n1 = 250
> n2 = 200
> P1 = 44/250
> P2 = 30/200
> P = (n1 * P1 + n2 * P2)/(n1 + n2)
> P
[1] 0.1644
> q = 1-P
> Zcal = (P1 - P2)/ sqrt(P * q * (1/n1 + 1/n2))
[1] 0.7343
> Pvalue = 2 * (1 - pnorm (abs(Zcal)))
> Pvalue
[1] 0.4590
```

$\because$ Pvalue > 0.05, we accept $H_0: P_1 = P_2$

Q4). Random Sample of 400 males & 600 females where as ned whether they want a ATM nearly. 200 males & 340 females were in favour of proposal, test the hypothesis that the proportion of males & females favouring proposal are equal or not at 5%

320

→ $H_0 : P_1 = P_2$ against $H_1 : P_1 \neq P_2$

```
> n1 = 400
> n2 = 600
> P1 = 200/400
> P2 = 340/600
> 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] -1.7247.
> Pvalue = 2*(1-pnorm(abs(zcal)))
[1] 2.303922e-06.
```

∵ Pvalue < 0.05, we reject $H_0 : P_1 = P_2$.

Q5. Following are the two independent sample from the 2 population test. equally bc of 2 population mean at 5% level of significance

x_1 = 74, 77, 74, 75, 79, 76, 82, 72, 75, 78, 77, 76, 76.

x_2 = 72, 76, 74, 70, 70, 78, 70, 72, 75, 79, 75, 76, 72, 74, 80.

$H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$

```
> x1 = c(74,77,74 ...)
> n1 = length(x1)
> mx1 = mean(x1)
> variance = (n1 - 1) * var(x1)/n1
> variance
[1] 0.4508
> sd 1 = sqrt(variance)
> sd 1
[1] 0.6714
> x2 = c(72,76,74 ...)
> n2 = length(x2)
> mx2 = mean(x2)
> var = (n2+1) * var(x2)/n2
[1] 0.6163
> sd2 = sqrt(var)
[1] 0.7850
> tcal = (mx1 - mx2)/sqrt((sd1^2/n1)+(sd2^2/n2))
> tcal
[1] 1.5757
> cat("t calculated is :", tcal)
[1] t calculated is 1.5757
```

FAO

PRACTICAL NO: 08
Small sample test.

1) The random sample of 15 observation are given by
80, 100, 110, 105, 122, 70, 120, 110, 101, 88, 8[illegible]
95, 89, 107, 125.
Do this data. support the assumption that population mean is 100
Ho $\mu_1 = \mu_2$

> x = c(80, 100, 110 .... 125)
> x
> 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 confidence interval:
91.3775 109.28892
sample estimates:
mean of x
100.3333
Since p-value is less than 0.05 the value of $H_0: \mu$ is rejected.

2) Two groups of 10 students score the following marks

group 1 - 18, 22, 21, 17, 20, 17, 23, 20, 22, 21

group 2 - 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

Test the hypothesis that there is no significant difference between the scores at 1% level of significance.

$H_0 = \mu_1 = 100$

```
> grp A = c(18,22 .....)
  grp B = c(16,20 .....)
> 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 means 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 is greater than 0.05 it is accepted.

3) Two types of medicine are used ~~to~~ on 5 and 7 patients for reducing their weight.

The decrease in the weight after using medi[cine] given below.

~~$H_0: \mu_1 = \mu_2$~~

A = 10, 12, 13, 11, 14

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

is There a significant difference in effi[cacy] of the medicine.

SPO

Ho : $\mu_1 = \mu_2$

⇒ > x = c(10, 12 ....)
y = c(8, 9 .....)

> x
[1] 10, 12 ...

> y
[1] 80, 92 ...

> t.test(x, y)

Welch two sample t test.

data : x and y
t = 0.80384, df = 9.7594, p-value = 0.4406
alternative hypothesis : true difference in means is not equal to 0.
-1.781171 3.781171

sample estimates:
mean of x and y
12 11

since p-value is greater than 0.05 we accept the value.

4) The weight reducing diet program is conducted and the observations are noted by 10 participants. Test whether the program is effective or not

data before - 120, 125, 115, 130, 123, 119, 122, 127, 118

after - 111, 114, 107, 120, 115, 112, 112, 120, 111

H_0 $\mu_1 = \mu_2$.

8) > a = c (66, ...)
> b = c (64, ...)
> t. test (a, b)

Welch two sample t-test.

data : a & b
t = - 0.6303, df = 13.836, p-value = 0.5387
alternative hypothesis : true difference in mean is not equal to 0
95 percent confidence interval :
- 15.422359 8.422359
mean of x mean of y
80.0 83.5

Since p-value is greater than 0.05

Ho : There is no significant difference in weight.
against H_1 : The diet program reduce weight
Ho $\mu_1 = \mu_2$.

```
> d.before = c (120, 125 .... )
> d.after = c (111, 114 ... )
> t.test (d.before, d.after, paired = T, alternative = "Less")
```

Paired t-test.

data : d.before and d.after
t = 9.0711, df = 9, p-value = 1
alternative hypothesis : true difference in mean is less than 0
95 percent confidence interval :
- Inf 9.616654
sample estimates :
mean of the differences.
8.

Since p-value is greater than 0.05 we accept the value

5) Sample A = 66, 67, 75, 76, 82, 84, 88, 90, 92

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

Test the population means are equal or not

Ho $\mu_1 = \mu_2$

050

6) The following are marks before & after ~~are~~ a training prgm 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: $\mu_1 = \mu_2$.

> before = c(71, ...)

> after = c(74, ...)

> t.test(before, after, paired = T, alternative "greater")

Paired t-test.

data : before & after

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

alternative hypothesis : true ~~rate~~ difference in means is greater than 0.

95 percent confidence interval :

-5.076639 Inf

sample estimates.

mean of the differences

-3.6

∴ p-value is greater than 0.05 we accept the value.

AM
07.07.20

B- 64, 66, 74, 82, 85, 87, 92, 95, 93, 97.
Test whether the population has the same variance at 5% level of significance.

1.
⇒ n = 100
mx = 52
mo = 55
sd = 7
zcal = (mx - mo) / {sd / (sqrt(n))}
> zcal
[1] -4.285714
> Pvalue = 2* (1-pnorm (abs(zcal)))
> Pvalue
[1] 1.82153e-05
∵ pvalue is less than 0.05 we reject the value $H_0 : \mu_1 = \mu_2$.

2.
⇒ > P = 0.5
> p = 350/700
> p
[1] 0.5
> n = 700
> q = 1-p
> q
[1] 0.5

020

PRACTICAL - NO-9.

large & small ~~number~~ sample test

Q1. The airthmatic mean of sample of 100 items from a large population is 52 - if a standard deviation is 7. test the hypothesis that the population mean is 55 against the alternative is more than 55 at 5%. LOS (level of significance)

Q2. 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%. LOS

Q3. 1000 articles form a factory A are found to have 2% defectives. 1500 article from a second factory are found to have 1%. defective. Test that 5%. level of significance that the two factory similar or not

Q4. A sample of size 400 was drawn & a sample mean 99 test at 5% los that the sample comes from the population wich mean 100 variance 64?

Q5. The flower stems are selected & the heights are found to be (cm) 63, 63, 68, 69, 71, 71, 72. Test the hypothesis that mean weight is 66 or not at the 1%. of significance.

Q6. Two random sample were drawn from Two normal population and three values are A - 66, 67, 75, 76, 82, 84, 88, 90, 92 -

B - 64, 66, ...
Test ...
variance ...

1.
⇒ n = 100
mx = 5...
mo = ...
sd = ...
Zcal = ...
> Zcal ...
[1] ...
> P...
> R...
[1]

val...

2.
⇒ > P...
> P...
> ...

```
> p = 0.5
> zcal = (p-P) / (sqrt(P* q/n))
> zcal
[1]  0
> Pval = 2* (1- pnorm (abs (zcal)))
> Pval
[1]      1
```

∵ Pvalue > 0.05 we accept the value.

3.

⇒

```
> n1 = 1000
> n2 = 1500
> P1 =  0.02
> P2 =  0.01
```

~~> q = 1-p~~

```
>  p = (n1* p1 + n2 * p2) / (n1+n2)
>  p
[1]  0.014
>  q = 1-p
>  q
>   0.986
> zcal = (p1- p2) /sqrt (p* q* (1/n1 + 1/n2))
> zcal
[1]  2.084842
> Pvalue = 2* (1-pnorm (abs (zcal)))
> Pvalue
[1]  0.03708364.
```

∵ Pvalue is less than 0.05 we reject the value $H_0 : \mu_1 = \mu_2$

4.
```
> n = 400
> mx = 99
> mo = 100
> var = 64
> sd = sqrt(var)
> sd
[1] 8
> zcal = (mx-mo) / {sd / sqrt(n)}
> zcal
[1] -2.5
> Pvalue = 2* (1-pnorm (abs(zcal)))
> Pvalue
[1] 0.01241933
```
∴ Pvalue < 0.05 we reject the null $H_0: \mu_1 = \mu_2$.

5.
```
> x = c(63, 63, 68, 69, 71, 71, 72)
> x
[1] ...
> t.test(x)
```
One sample t test

data : x

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

alternative hypothesis : true mean is equal to 0 95 percent confidence interval :

64.6647 71.622092

Sample estimates :

mean of x

68.14286

∵ Pvalue is less than 0.01. we reject $H_0 : H_1 = H$

Q6.

H_0 = population is same
H_1 = variance in population

> x = c(Data)
> y = c (Data)
> F = var.test(x, y)
> F

F test to Compare 2 variance.

data = x & y

F = 0.70686, num def = 8, denom df = 10, p-value = 0.6359 alternative hypothesis: true solution of variance.s not equal to 1.

95 percent Confidence interval:

0.1833 3.036039

Sample estimates:
ratio of variances
0.70687.

∵ P-value is greater than 0.05: we accept H_0 : H_0 = The population have same variance.

PRACTICAL - 10

Anova and chi-square test

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

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

→ H_0 : CC & CH

```
> x = c(20,80,35,50,20,45)
> m = 3 (rows)
> n = 2 (columns)
> y= matrix (x, nrow= m, ncol = n)
> y
        [,1]   [,2]
[1,]     70     50
[2,]     80     20
[3,]     35     45
> PV= chiq.test(y)
> PV
        Pearson's Chi - squared test

data : y
X-squared = 25.646, df= 2, p-value = 2.698e-06
```

∴ Pvalue is ~~greater~~ less than 0.05 we ~~reject~~ accept

Ho: vacination by disease are independent.

Q3. Perform. ANOVA for the data.

Varieties

A

B

C

D

observations

50, 52

53, 55, 53

60, 58, 57, 56

52, 51, 54, 55

Ho: The means of the varieties are equal

> x_1 = c(50,52)

> x_2 = c(53...)

> x_3 = c(60...)

> x_4 = c(52...)

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

> names(d)

> one way.test(values ~ ind, data = d, var.equal = T)

data : values and ind

F = 11.735, num df = 3, denom df = 9, pvalue = 0.00183

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

> anova

Call :

aov(formula - values ~ ind, data = d)

terms :

	ind	Residuals
Sum of squares	71.06416	18.16667
Deg. of freedom	3	9

∴ Pvalue is less than 0.05 we rej...
H_0 : CC & CH

Q9. Use the following data to find vaccination & a particular disease are independent

	AFF	Not AFF
Vaccination Given	20	30
Not given	25	35

⇒ H_0 : Vaccination and disease 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
    Pearson's Chi-squared test with Yates' continuity
    correction
data: y
X-squared = 0, df = 1, p-value = 1
```

Residual standard error: 1.420746
Estimated effects maybe unbalanced
∴ Since pvalue is less than 0.05, we reject
H0: The means of the varieties are equal

Q4. The following data gives the life of tyre of 4 brand

Type	Observation
A	20, 23, 18, 17, 18, 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 the average life for the four brand are same

→ H0: Avg life of 4 brands 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, 13, 14, 16)
  d = stack(list(b1 = x1, b2 = x2, b3 = x3, b4 = x4))
> names(d)
> oneway.test(values ~ ind, data = d, var.equal = T)
```

one way analysis of means

data: values & 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)

Terms:

	ind	Residuals
Sum of squares	91.4381	89.0619
Deg. of freedom	3	20

Residual standards error: 2.110236
Estimated effects may be unbalanced

∵ the p-value is less than 0.05, we reject.

→ One thousand students of a college are graded according to their IQ and the economic condition of their home. Check that is there any association between IQ and their economic condition of their home.

		IQ High	IQ Low
EC	High	460	140
	Med	330	200
	Low	240	160

PRACTICAL NO: 11

Non- parametric test

21. Following are amounts of sulphuric oxide emited by industries in 20 days apply sign test to test hypothesis that population medium is 21.5 ~~to test~~

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

H0: population median is 21.5

```
> x = c(17, 15, ....)
> x
> median = 21.5
> sp = length(x[x> med])
> sp
[1] 9
> sn = length(x[x < med])
[1] 11
> n = sp+sn
> n
[1] 20
> pv = pbinom(sp, n, 0.5)
> pv
[1] 0.4119015
```

∴ pvalue > 0.05 we accept H0.

→ 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, p-value = 2.364e
```

∵ p-value is less than 0.05, we reject Ho : EC & IQ are independent

PM

Q2. Following are the 10 observations
612, 619, 631, 628, 643, 640, 655, 649, 670, 663
Apply sign test to test hypothesis that population median is 625 against alternative it is greater than 625 at 10%. LOS.

Ho : population median is 625

```
> x = c(612, 619, 631, 628, 643, 640, 655, 649, 670, 663)
> x
[1] 625
> sp = length(x[x>m])
> sp
[1] 8
> sn = length(x[x<m])
[1] 2
> n = sp+sn
[1] 10
> pv = pbinom(sp,n,0.5)
> pv
[1] 0.9892578
```

∵ pvalue > 0.01, we accept Ho.

The 10 observations are -
36, 22, 21, 30, 24, 25, 20, 27, 20, 18
Using sign test, test hypothesis is 25 against alternative is less than 25 at 5% LOS

Ho: population median is 25

```
> x = c(36, 32...)
> m = 25
> m
[1] 25
> sp = length (x[x>m])
[1] 3
sp = length (x[x<m])
[1] 96
> n = sp + sn
[1] 9
> pv = pbinom (sn, n, 0.5)
> pv
[1] 1
```

∵ pvalue > 0.05, we accept Ho

Q4. Following are measurements -
63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63,
72, 65
Using wilcoxon signed rank test, test that population median is 60 against alternative it is greater than 60 at 5% LOS.

Ho : population median is 60

```
> x = c(63, 65, 60 --)
> wilcox.test(x, alt = "greater", mu = 60)
```

wilcoxon signed rank test with Continuity Correction

Q.5. Observations 15, 17, 25, 24, 20, 21, 32, 28, 12, 25, 27, 25

use wilcoxon signed rank test the hypo that pop. median is less than 20.

> H0 : population median is 20
> x = c(15, 17, 25 ...)
> x
> wilcox.test (x, alt = "less", mu = 20)

wilcoxon signed rank-test with Continuity Correction

data : x

V = 48.5, p-value = 0.92392

alternative hypothesis: True location is less than 20.

Q6. 20, 25, 27, 30, 18 wilcoxon signed rank test the hypothesis at population median is 25 against the alternative population median is not equal to 25 at 5% LOS.

⇒ H0 : population median is 25
> x = c(20, 25 ...)
> wilcox.test (x, alt = "two sided", mu = 25).

wilcoxon signed rank test with Continuity Correction.

data : x
V = 3.5, pvalue = 0.71

alt hypo : true location is not equal to 25

∴ Pvalue > 0.05

∴ we accept the H0.