★★ INDEX ★★

* Basics of R software :-

R is a software for data analysis of statistical computing.
This software is used for effective data handling & output storage is possible.
It is capable of graphical display.
It is a free software.

① $2^2 + \sqrt{25} + 35$

→ 2^2 + sqrt 25 + 35

= 44

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

→ 2 * 5 * 3 + 62 / 5 + sqrt 49

= 49.4

③ $\sqrt{76 + 4 \times 2 + 9 \div 5}$

→ sqrt (76 + 4 * 2 + 9/5)

= 9.262829

④ $42 + |-10| + 7^2 + 3 \times 9$

→ 42 + abs (-10) + 7^2 + 3 * 9

= 128

Scanned with CamScanner

$x = 20$; $y = 30$

Find, $x+y$; x^2+y^2 ; $\sqrt{y^3-x^3}$; abs(x-y)

$x+y$
$= 50$

iii) x^2+y^2
$= 1300$

$\sqrt{y^3-x^3}$
$= 137.8405$

iv) abs(x-y)
$= 10$

Calculate the following.

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

4 9 16 25

c(4,5,6,8) * 3

12 15 18 24

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

-4 -9 -25 -28

c(2,3,5,7) * c(8,9)

16 27 40 63

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

1 8 9 64 25 216

Scanned with CamScanner

Find the sum, prod, min, max of the given values

5, 8, 6, 7, 9, 10, 15, 5

x = c(5,8,6,7,9,10,15,5)

length(x)
= 8

sum(x)
= 65

max(x)
= 15

prod(x)
= 11340000

min(x)
= 5

Matrix calculation:

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

x <- matrix(nrow=4, ncol=2, data=c(1,2,3,4,5,6,7,8))

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

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))

Scanned with CamScanner

$x + y$

$$\begin{bmatrix} 3 & 8 & 17 \\ 0 & 13 & -3 \\ 13 & 12 & 21 \end{bmatrix}$$

$x * y$

$$\begin{bmatrix} 2 & 186 & 70 \\ -4 & 40 & -88 \\ 30 & 36 & 108 \end{bmatrix}$$

$2 * x + 3 * y$

$$\begin{bmatrix} 8 & 20 & 44 \end{bmatrix}$$

Scanned with CamScanner

7	1
9	1
10	1
12	1
14	2
15	1
16	1
17	1
18	2
19	1

Scanned with CamScanner

PRACTICAL - 02

Problems on PDF & CDF

Q] Can the following be pdf?

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

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

$= \int^{2} (2-x)\,dx$

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

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

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

$\neq 1$

Using, $\int x^n dx = \frac{x^{n+1}}{n+1}$

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

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

$= \int_0^1 3x^2\,dx$

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

Scanned with CamScanner

Find ; $P(x \le 2)$; $P(2 \le x < 4)$; $P(\text{atleast } 4)$, $P(3 < x < 6)$

$\to P(x \le 2) = P(0) + P(1) + P(2)$
$= 0.1 + 0.1 + 0.2$
$= 0.4$

$P(2 \le x < 4) = P(2) + P(3)$
$= 0.2 + 0.2$
$= 0.4$

$P(\text{atleast } 4) = P(4) + P(5) + P(6)$
$= 0.1 + 0.2 + 0.1$
$= 0.4$

$P(3 < x < 6) = P(4) + P(5)$
$= 0.1 + 0.2$
$= 0.3$

Q.4] 1]

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

$F(x) = 0$ if $x < 0$
$= 0.1$ if $0 \le x < 1$
$= 0.2$ if $1 \le x < 2$

Scanned with CamScanner

36

②

x	10	12	14	16	18
$P(x)$	0.2	0.35	0.15	0.25	0.1

$$P(x) = \begin{cases} 0.2, & x < 10 \\ 0.55, & 10 \le x < 12 \\ 0.70, & 12 \le x < 14 \\ 0.90, & 14 \le x < 16 \\ 1.0, & x \ge 18 \end{cases}$$

AM

Scanned with CamScanner

Scanned with CamScanner

PRACTICAL - 03

Probability Distribution & Binomial Distribution

Find the PDF of the following CDF and draw the graph.

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

$$
p(x) = \begin{cases} 0 & \text{if } x < 10 \\ 0.15 & 10 \le x < 20 \\ 0.40 & 20 \le x < 30 \\ 0.70 & 30 \le x < 40 \\ 0.90 & 40 \le x \; 50 \\ 1.0 & x \ge 50 \end{cases}
$$

* Commands *

```
x = c(10,20,30,40,50)
prob = c(0.15, 0.25, 0.3, 0.2, 0.1)
cumsum(prob)
[1] 0.15  0.40  0.70  0.90  1.00
plot(x, cumsum(prob), xlab="values", ylab="probability",
     main="graph of CDF", "s")
```


Scanned with CamScanner

1] Suppose their are 12 MCQ's in a test, each question has 5 options & only one of them is correct. Find a probability of having
① five correct answers.
② Atmost 4 correct ans.

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

① 5 correct answers.
→ $n = 12$, $p = 1/5$, $q = 4/5$; $x = 5$
dbinom (5, 12, 1/5)
[1] 0.0531

② Atmost 4 correct answers.
→ $n = 12$, $p = 1/5$, $q = 4/5$; $x = 4$
pbinom (4, 12, 1/5)
[1] 0.9274

2] Their are 10 members in a committee, the probability of any member attending a meeting is 0.9 find the probability
① 7 members attended
② Atleast 5 members attended
③ Atmost 6 members attended

→ It is given that $n = 10$, $p = 0.9$, $q = 0.1$
x = Total no. of members attended
$x \sim B(n, p)$

① 7 members attended
→ $n = 10$, $p = 0.9$, $q = 0.1$; $x = 7$
dbinom (7, 10, 0.9)
= [1] 0.0573

② Atleast 5 members attended
→ $n = 10$, $p = 0.9$, $q = 0.1$; $x = 5$
1 − pbinom (5, 10, 0.9)
[1] 0.9983

③ Atmost 6 members attended
→ $n = 10$, $p = 0.9$, $q = 0.1$; $x = 6$
pbinom (6, 10, 0.9)
[1] 0.01279

* CDF *

1) Find the CDF & 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

→ $p(x) = 0$ ~~if x < 0~~
$= 0.1$ $0 \le x < 1$
$= 0.2$ $1 \le x < 2$
$= 0.4$ $2 \le x < 3$
$= 0.6$ $3 \le x < 4$
$= 0.7$ $4 \le x < 5$
$= 0.9$ $5 \le x < 6$

Scanned with CamScanner

COMMANDS *

*

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

plot (x, cumsum (prob), xlab = "values", ylab = " probability", main = " graph of CDF", "s")

Ay 11/2/19

Scanned with CamScanner

02 PRACTICAL – 04

4. BINOMIAL DISTRIBUTION

① Find the complete binomial distributions when n=5, p=0.1

Solution :-

Note :-

① P(X=x) = dbinom (x,n,p)

② P(X ≤ x) = pbinom (x,n,p)

③ P(X ≥ x) = 1- pbinom (x,n,p)

n=5, p=0.1

\> dbinom (0:5, 5, 0.1)

= 0.59049 0.32805 0.7290 0.00810 0.00045 0.0[illegible]

② Find the probability of exactly 10 successes in 100 trials with p=0.1

→ n= 100, p = 0.1, X= 10

\> dbinom (10, 100, 0.1)

= 0.1318653

③ x follows binomial distribution with n=12, p=0.2[illegible]

Find ① P (X=5)

② P(X ≤ 5)

③ P(X >7)

④ P(5 < X < 7)

Solutions :-

n=12, p=0.25

① n=12, p=0.25, x=5

\> dbinom (12,5,0.25)

= 0.1032414

② n=12, p=0.25, x≤5

\> pbinom (5,12,0.25)

= 0.9455978

③ n=12, p=0.25, x>7

\> 1- pbinom (7,12,0.25)

= 0.0278151

④ n=12, p=0.25, x>5 & x<7

\> dbinom (6,12,0.25)

= 0.0414945

NOTE :-

① P(X>7) = 1- P (X≤7)

= 1- pbinom (6,n,p)

② P(X ≥7) = 1- P (X≤5)

= 1- pbinom (7,n,p)

④ The probability of salesman makes a sale to customer is 0.15; Find the probability:

1) No sale for 10 customer.

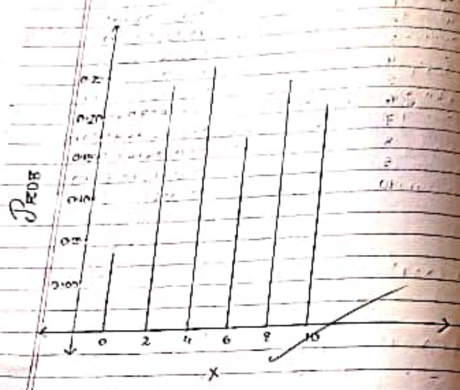
2) More than 3 sale in 20 customer.

x. values	Probability
0	0.0060466176
1	0.0403107840
2	0.1209323520
3	0.2149908480
4	0.2508226560
5	0.2006581248
6	0.1114767360
7	0.0424673280
8	0.0106168320
9	0.0015728640
10	0.0001048576

Scanned with CamScanner

52

plot (x, prob, "h")
plot (x, cumprob, "s")



PRACTICAL - 5

Normal Distribution

43

$P[X=x] =$ dnorm (x, μ, σ)

$P(X \le x) =$ pnorm (x, μ, σ)

$P(X \ge x) = 1 -$ pnorm (x, μ, σ)

$P(x_1 < X < x_2) =$ pnorm (x_2, μ, σ)

To find the value of k so that $P(X \le k) = p$; qnorm (p, μ, σ)

To generate 'n' random numbers:
rnorm (n, μ, σ)

$X \sim N$ [$\mu = 50$, $\sigma^2 = 100$)

find : i) $P(X \le 40)$

ii) $P(42 \le X \le 60)$

iii) $P(X > 55)$

iv) $P(X \le k) = 0.7$; k = ?

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

i) $P(X \le 110)$

ii) $P(X \le 95)$

iii) $P(X > 115)$

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

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

i) P (X ≤ 110)

```
> a = pnorm (110 , 100, 6)
> cat ("P (X<=110) is : ", a)
[1] P (X <= 110) is : 0.9522096
```

ii) P(X ≤ 95)

```
> b = pnorm (95, 100, 6)
> cat ("P (X<=95) is : ", b)
[1] P(X <= 95) is: 0.2023284
```

iii) P(X > 115)

```
> c = 1- pnorm (115, 100, 6)
> cat ("P (X >115) is : ", c)
[1] P(X >115) is : 0.006209665
```

iv) P(95 <= X <= 105)

```
> d = pnorm (95, 100, 6) - pnorm (105, 100, 6)
> cat ("P (95 <= X <=105) is : ", d)
[1] P(95 <= X <= 105) is : 0.5953432
```

v) P(X ≤ k) = 0.4; ~~k is 9~~

```
e = qnorm (0.4, 100, 6)
cat ("P (X <=k) = 0.4 , k is : ", e)
P (X< =k) = 0.4 , k is : 98.47992
```


Scanned with CamScanner

14

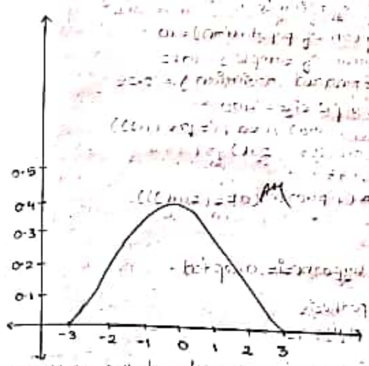
3)

```
> x = rnorm (10, 60, 5)
> x
[1] 57.4310   59.0626   63.72034   61.2825   61.5102
    61.8524   66.9736   62.6194    61.3294   62.4702

> am = mean (x)
> am
[1] 60.44768

> median = median (x)
> median
[1] 61.6196

> n = 10
> variance = (n-1) * var (x)/n
> variance
[1] 20.37836

> sd = sqrt (variance)
> sd
[1] 4.5142
```

Draw the graph of standard normal distribution.

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

45



Scanned with CamScanner

PRACTICAL – 06

Z – DISTRIBUTION

$H_0 : \mu = 10$ against $H_1 : \mu \neq 10$. A sample of size 400 we selected which give a mean 10.2 and standard deviation 2.25. Test the hypothesis at 5% level of significance.

→ μ_0 (mean of population) = 10
μ_x (mean of sample) = 10.2
sd (standard deviation) = 2.25
n = sample size = 400

```
> zcal = (mx - m0) /(sd /(sqrt(n)))
> cat("z cal is:", zcal)
> zcal is: 1.77
> pvalue= 2*(1-pnorm(abs(zcal)))
> pvalue
> 0.079
```

Thus, hypothesis accepted.

Test the hypothesis
$H_0 : \mu = 75$ $H_1 : \mu \neq 75$

A sample of size 100 is selected and the sample mean is 80 with the standard deviation of 3 test the hypo. at 5% level of significance.

→ $\mu_0 = 75$
$\mu_x = 80$
sd = 3

Shot on OnePlus


```
zcal = (mx - m0) / (sd /(sqrt(n)))
cat("zcal is:", zcal)
zcal is: 16.66667
pvalue = 2*(1-pnorm(abs(zcal)))
pvalue
0
```

Thus, hypothesis rejected

Test the hypothesis
$H_0 : \mu = 25$ against $H_1 : \mu \neq 25$
A sample of 30 is selected test the hypothesis at 5% level of significance. A sample of following 30 is selected 20, 24, 27, 35, 30, 46, 26, 27, 10, 20, 30, 57, 35, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 19, 24, 15, 19, 22, 20, 18

```
x = c(20, 24, 27, 35, ......, 19, 20, 20, 18)
mx = mean(x)
mx = 26.066
n = length(x)
variance = (n-1) * var(x) /n
sd = sqrt(variance)
>> sd = 7.2798
m0 = 25
zcal = (mx - m0) / (sd / sqrt(n))
z cal = 0.8025
pvalue = 2 * (1- pnorm (abs (zcal)))
pvalue = 0.4223
```

This, hypothesis accepted.

Scanned with CamScanner

Experience has shown that 20% students of a college smoke. A sample of 400 students reveal that out of 400 only 50 smokes. Test the hypothesis that the experience keeps the correct proportion or not.

> P = 0.2
q = 1-P
p = 50/400
n = 400
z cal = (p-P) / (sqrt (P*q /n))
z cal
>> -3.75
pvalue = 2*(1-pnorm (abs(zcal)))
pvalue

AT

Scanned with CamScanner

2] write, two random sample of size 1000 & 2000 are drawn from 2 population with the means 67.5 & 68 respectively and with the same SD of 2.5. Test the hypothesis that the mean of two population are equal.

→ H0 two populations are equal.

	A	B
population	1000	2000
mean	67.5	68
SD	2.5	2.5

```
n1 = 1000
n2 = 2000
mx1 = 67.5
my2 = 68
sdx = 2.5
sdy = 2.5
y= (m1-m2)/sqrt((sdx^2/n1) +(sdy^2/n2))
y
[1] -5.163978
cat ("y calculated is=", y)
z calculated is= -5.163978
pvalue =2*( 1-pnorm (abs(z)))
pvalue
2.417564e-07
```

Since, pvalue >0.05, we accept H0 at 5% level of significance.

In fybsc 20% of a random sample of 400 students had defective eye site in sybsc 15.5% of 500 students had same defect. Is a difference of proportion in same?

H0 the proportion of the population are equal

```
n1 = 400
n2 = 500
p1 =0.2
p2 = 0.155
p= (n1 * p1 + n2 * p2) / (n1+n2)
p
[1] 0.175
q= 1-p
q
0.825
z = (p1 -p2) / sqrt(p*q*(1/n1 + 1/n2))
z
[1] 1.76547
cat( "z calculated is=", z)
z calculated is= 1.76547
pvalue = 2* (1- pnorm (abs(z)))
pvalue
[1]0.07748487
```

Since, pvalue > 0.05, we accept H0 at 5% level of significance.

Scanned with CamScanner

58

4) From each of the box of the apples in sample of of 200 is collected. It is found that there are 44 bad apples in 1st sample & 30 in 2nd sample. Test the hypothesis that two boxes are equivalent in terms of no. of bad apples.

→ H_0: the two boxes are equivalent.

```
n1 = 200
n2 = 200
p1 = 44/200 = 0.22
p2 = 30/200 = 0.315
p = (n1*p1 + n2*p2)/(n1+n2)
q = [1] 0.37 0.185
q = 1-p
[1] q = 0.8815
z = (p1-p2)/sqrt(p*q*(1/n1+1/n2))
z
[1] 1.802741
cat("z calculated is =", z)
z calculated is = 1.802741
pvalue = 2*(1-pnorm(abs(z)))
pvalue
[1] 0.07142888
```

Since, pvalue > 0.05, we accept H_0 at 5% level of significance.

5) In a HA class out of a sample of 60 mean height is 63.5" with SD of 2.5. In a N-low class out of 50 students mean height 69.5" with SD of 2.5. Test the hypothesis that the mean of HA & N-low are same.

H_0: height of two classes are equal.

```
n1 = 60
n2 = 50
mx = 63.5
my = 69.5
sdx = 2.5
sdy = 2.5
z = (mx-my)/sqrt((sdx^2)/n1)+(sdy^2/n2))
[1] 145.1808
pvalue = 2*(1-pnorm(abs(z)))
[1] 0
```

Since, pvalue < 0.05, we reject the H_0 at 5% level of significance.

Scanned with CamScanner

PRACTICAL- 8

Small Sample Test :-

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

→ H_0 : mean = 66cm

```
> 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

Since, pvalue is <0.01, we reject H_0 at 1% level of significance.

Two random sample was drawn from two different population,

sample1 = 8, 10, 12, 11, 16, 15, 18, 17

sample2 = 20, 15, 18, 9, 8, 10, 11, 12

Test the hypothesis that there is no difference between the population mean at 5% level of significance.

→ H_0 there is no difference in the population mean

```
> x = c(8,10,12,11,16,15,18,17)
> y = c(20,15,18,9,8,10,11,12)
> t.test(x,y)
```

Welch Two Sample test

data: x and y

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

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

-5.19 2719

Sample Estimates:

mean of x   mean of y

12.125   12.875

Since, pvalue is <0.01 we reject H_0 at 1% level of significance.

Scanned with CamScanner

Q.3] Following are the weights of 10 ppl before and after a diet program, test the hypothesis that the diet program is effective or not.

Before (kg) = (100, 95, 96, 98, 112, 115, 125, 109, 110, 104)

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

Solution :- H_0 the diet program is not effective.

```
>x = c(100,125,95,96,98,102,115,104,109,110)
>y = c(95,80,95,97,90,100,110,85,100,101)
>t.test (x,y, paired = T, alternative = "less")
```

Paired test.

data : x & y

t = 2.3215 , df = 9 p-value = 0.9773

alternative hypothesis test difference in mean is less than 0 - inf 17.89635

sample estimates :

mean of the differences.

10

pvalue is < 0.01 we reject H_0 at 5% level of significance.

Q.4] The marks before & after training program are given below.

Before (20, 25, 32, 28, 27, 30, 35, 25)

After (30, 35, 32, 37, 37, 40, 40, 23)

Test the hypothesis the training program is effective or not.

```
>x = c(20,25,32,28,27,36,35,21)
>y = c(30,35,32,37,40,40,23)
>t.test (x,y, paired = T, alternative = "greater")
```

Paired t-test

data : x & y

t = -3.3859, df = 7 , p-value = 0.9942

alternative hypothesis : true difference in mean is greater than 0 as 5% confidence interval

-8.967395 inf

Sample estimates:

mean of the difference.

-5.75

Scanned with CamScanner

Q.5] Two random samples were drawn from two normal population & the values are

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

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

test whether the population has same variance at 5% level of significance.

→ Ho variance of the population are equal

> x = c(66, 67, 85, 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, p-val = 0.6159

alternative hypothesis : true ratio of variances is not equal to 1 as 5% confidence interval

0.1833662 3.0360398

Sample estimates

ratio of variances

0.7068567

pvalue is <0.01, we reject Ho at 5% los.

The arithmetic mean of a sample of 100 obs is 52. If the sd is 7, test the hypothesis that the population mean is 55 or not at 5% of los.

Ho population mean = 55

n = 110

mx = 52

m0 = 55

sd = 7

xcal = (mx - m0) / (sd / sqrt(n))

xcal

[1] -4.494895

n = 100

zcal = (mx - m0) / (sd / sqrt (n))

zcal

[1] -4.285714

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

pvalue

[1] 1.82153e-05

pvalue is <0.05 we reject H₀ at 5% los.

12.2.20

52

PRACTICAL - 9

Chi Square Distribution & ANOVA BARK

Q.1 Use the following data to test whether the cleanliness of the home depends upon the child.

		Cond. of Home Clean	Dirty
Cond. of child	Clean	70	50
	Fairly D.	80	20
	Dirty	35	45

H_0 condition of the home & 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's Chi-squared test

53

data: y
X-squared = 25.646, df = 2, p-value = 2.689e-06

Since, pvalue < 0.05 we ~~accept~~ reject H_0 at 5% LOS

Q.2 Table below shows the relation between the performance of mathematics & computer of CS students.

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

H_0 can relation between performance of mathematics & computer are independent

```
x=c(56,47,14,71,163,42,12,38,85)
m=3
n=3
y=matrix(x, nrow=3, ncol=3)
y
      [,1] [,2] [,3]
[1,]   56   71   12
[2,]   47  163   38
[3,]   14   42   85
```


Scanned with CamScanner

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

Pearson's chi-squared test

data = y

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

Since, pv < 0.05 we reject Ho at 5% LOC.

Q.3) Perform ANOVA for the following data.

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

H_0 the mean of variety A,B,C,D 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))
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)
summary (anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ind	3	71.06	23.688	11.73	0.00184
residuals	9	18.17	2.019		

pv < 0.05 we reject H_0 at 5% LOS.

Perform ANOVA for the following data.

Types	Observations
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

Ho the mean of type A, B, C, D are 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))
names (d)
[1] "values"  "ind"
oneway.test (values ~ ind, data = d, var.equal=T)
```


Scanned with CamScanner

52

One-way analysis of means

data: values and ind
F = 2.6667, num df = 3, denom df = 8,
p value = 0.1189

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

	Df	Sum sq	Mean sq	F value	Pr(>F)
ind	3	18	6	2.667	
Residuals	8	18	2.25		0.11

pv > 0.05 we accept H_0 at 5% LOC.

Q.5

x = read.csv ("C:/Users/admin/Desktop/MARKS...
x

	STATS	CA1
1	40	60
2	45	48
3	42	47
4	15	20
5	37	25
6	36	27
7	49	57
8	59	58
9	20	25
10	27	27

PRACTICAL-10

Non-Parametric test.

54

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

17 15 20 29 18 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 that it is less than 21.5.

H_0 :- Population median = 21.5

H_1 :- It is less than 21.5

x = c(17, 15, 20, 29, 18, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26)

m = 21.5
sp = length (x[x>m])
sn = length (x[x<m])
n = sp + sn
pv = pbinom (sp, n, 0.5)
pv
0.4119015

NOTE :- If the alternative is > median pv = pbinom (sn, n, 0.5)

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 > 25.

H_0 : The median is 25 $\quad$ H_1 : The median > 25

```
> x = c (12, 19, 31, 28, 43, 40, 55, 49, 70, 68)
  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
```

Note:- If the alternative is < median
pv = pbinom (sp, n, 0.5)

For the following data 60, 65, 63, 89, 61, 71, 58, 51, 48, 66 test the hypothesis using wilcoxon signed rank test for testing the hypothesis that the median is 60 against the alternative > 60.

H_0 : The median is 60
H_1 : The median > 60

```
x = c (60, 65, 63, 89, 61, 71, 58, 51, 48, 66)
wilcox.test (x, alter = "greater", mu = 60)
   wilcoxon signed rank test continuity correction
data : x
V = 29 , p value = 0.2386
alternative hypothesis: true location is greater than
   60.
```

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

Scanned with CamScanner

The alternative is > median pv= pbinom(59,...
0.5)

28

4) Using wilcoxon sign rank test test the hypothesis median = 12, against the hypothesis alternative less than 12, 13, 10, 20, 15, 7, 6, 11, 9, 20.

H_0 : The median is 12

H_1 : The median < 12

```
x = c(12, 13, 10, 20, 15, 5, 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.

Note : If the alternative is greater than
wilcox.test(x, alter = "greater", mu = 12)

Scanned with CamScanner