PRACTICAL - 01

* Aim: Basic of R Software.

Q.1> Solve
$4+6+8\div2-5$
> 4+6+8/2-5
[1] 9

2. $2^2+|-3|+\sqrt{45}$
> 2^2+abs(-3)+sqrt(45)
[1] 13.7082

3) $5^3+7\times5\times8+46/5$
> 5^3+7*5*8+46/5
[1] 414.2

4) $\sqrt{4^2+5\times3+7/6}$
> sqrt(4^2+5*3+7/6)
[1] 5.61567

5) round off
$46\div7+9\times8$
> round(46/7+9*8)
[1] 79

Q.2) C(2,3,5,7)*2
[1] 4 6 10 14
> c(2,3,5,7)*c(2,3,6,2)
> c(2,3,5,7)^2
[1] 1.50 0.40 1.75 1.00
> c(6,2,7,5)/c(4,5)
[1] 1.50 0.40 1.75 1.00

Q.3) > x = 20
> y = 30
> z = 2
> x ^ 2 + y ^ 3 + z
[1] 27402
> sqrt (x ^ 2 + y)
[1] 20.73644
> x ^ z + y ^ z
[1] 1300

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

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

Q.6) Marks of Statics of CS batch B
x = c(58, 20, 35, 24, 46, 55, 45, 27, 54, 40, 50, 36, 29, 35, 39)
> x = c(data)
> breaks = seq (20, 60, 5)
> a = cut (x, breaks, right = FALSE)
> b = table (a)
> c = transform (b)
> c

Scanned with CamScanner

	x	freq
1	[20, 25]	3
2	[25, 30]	2
3	[30, 35]	1
4	[35, 40]	4
5	[40, 45]	1
6	[45, 50]	3
7	[50, 55]	2
8	[55, 60]	4

02

PRACTICAL - 2

Topic : Probability Distribution

1) Check whether the foll are pmf or not

x	$P(x)$
0	0.1
1	
2	0.2
3	0.5
4	0.4
5	0.3
	0.5

If the given data is pmf then $\Sigma P(x) =$

$\therefore P(0) + P(1) + P(2) + P(3) + P(4) + P(5) = P(x$

$\therefore 0.1 + 0.2 + 0.5 + 0.4 + 0.3 + 0.5$

$\therefore 1.0$

$P(2) = -0.5$ it can be a probability mass function

$\therefore P(x) \geq 0 \quad \forall, x$

2)

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

The condition for pmf is $\Sigma P(x) = 1$

$\Sigma P(x) = P(1) + P(2) + P(3) + P(4) + P(5)$

$= 0.2 + 0.2 + 0.3 + 0.2 + 0.2$

$= 1.1$

$\therefore$ The given data is not a pmf because the $P(x) \neq 1$.

2) Find the c.d.f for the foll pmf and sketch the graph

$$f(x) = \begin{cases} 0 & x < 10 \\ 0.2 & 10 \le x < 20 \\ 0.4 & 20 \le x < 3 \\ 0.75 & 30 \le x < 4 \\ 0.90 & 40 \le x < 5 \\ 1.0 & x \ge 50 \end{cases}$$

iii)

x	1	2	3	4	5	6
$P(x)$	0.15	0.25	0.1	0.2	0.2	0.1

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.15 & 1 \le x < 2 \\ 0.40 & 2 \le x < 3 \\ 0.50 & 3 \le x < 4 \\ 0.70 & 4 \le x < 5 \\ 0.90 & 5 \le x < 6 \\ 1.00 & x \ge 6 \end{cases}$$

... whether the following is pdf or not

i) $f(x) = 3 - 2x; \quad 0 \leq x \leq 1$

ii) $f(x) = 3x^2; \quad 0 < x < 1$

i) $f(x) = 3 - 2x$

$= \int_0^1 f(x) = \int_0^1 (3-2x)dx = \int_0^1 2x\,dx$

$\therefore [3x - x^2]_0^1 = 2$

$\because$ The $\int_0^1 f(x) = 1$ $\therefore$ It is not a pdf.

ii) $f(x) = 3x^2; \; 0 < x < 1$

$\int_0^1 f(x)$

$= \int_0^1 3x^2 \quad \therefore 3\int_0^1 x^2$

$\therefore \left[3\frac{x^3}{3}\right]_0^1 \quad \because x^n, \frac{x^{n+1}}{n+1}$

$\therefore x^3$

$\therefore 1$

$\therefore$ The $\int_0^1 f(x) = 1$ $\therefore$ It is a pdf.

Scanned with CamScanner

Q.3) Check that whether the following is pdf or not

i) $f(x) = 3 - 2x; \quad 0 \le x \le 1$

ii) $f(x) = 3x^2; \quad 0 < x < 1$

i) $f(x) = 3 - 2x$

$= \int_0^1 f(x) = \int_0^1 (3-2x)\,dx = \int_0^1 2x\,dx$

$\therefore [3x - x^2]_0^1 = 2$

$\because$ The $\int_0^1 f(x) = 1$ $\therefore$ It is not a pdf.

ii) $f(x) = 3x^2; \quad 0 < x < 1$

$\int_0^1 f(x)$

$= \int_0^1 3x^2 \quad \therefore 3\int_0^1 x^2$

$\therefore \left[3\frac{x^3}{3}\right]_0^1 \quad \because x^n = \frac{x^{n+1}}{n+1}$

$\therefore x^3$

$\therefore$ The $\int_0^1 f(x) = 1$ $\therefore$ It is a pdf.

01

PRACTICAL - 3

Topic :- Binomial distribution.

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

P(X ≤ x) = pbinom(x, n, p)

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

If x is unknown

p1 = P(X ≤ x) = qbinom(p1, n, p)

1) Find the probability of exactly 10 success in hundred trials with p = 0.4

2) Suppose there are 12 MCQ each question has 5 option out of which 1 is correct find the probability of having
i) exactly 4 answer.
ii) at most 4 correct answer.
iii) More than 5 correct answer

3) Find the complete distribution when n = 5 & p = 0.1

4) x = 12, p = 0.25, find all probability
i) P(x = 5) iii) P(x > 7)
ii) P(x ≤ 5) iv) P(5 < x < 7)

5) The probability of salesman of salesman making a sale to customer 0.15. Find probability of
i) No sales out of 10 customer.
ii) More than 3 sales out of 20 custom

6) A salesman has 20% probability of making a sale to customer out of 30 customer what minimum no of sales he can make with 88% of probabi

7) x follow binomial distribution with n=10, P=0.3 plot the graph.

Answers.

5) dbinom(0, 10, 0.15)
[1] 0.1968744
> 1-dbinom(3,20,0.15)
[1] 0.3522748

6) qbinom (0.88, 30, 0.2)
[1] 9

7) n=10
p = 0.3
~~x = 0-n~~
~~Prob~~ = dbinom(x, n, p)
cumprob = pbinom(x, n, p)
d = data.frame("x value", =x,
"Probability", =prob)
> print(d)

	x values	probabilities
1	0	0.0282
2	1	0.1210
3	2	0.2334
4	3	0.2668
5	4	0.2001
6	6	0.1029
7	6	0.0367
8	7	0.0090
9	8	0.0014
10	9	0.0001
11	10	0.0000

PRACTICAL : 4

* Aim :- Normal distribution. 06

i) P(x ≤ x) - dnorm (x, u, 6)
ii) p(x ≤ x) = pnorm (x, u, 6)
iii) P(x > x) = 1 - pnorm (x, u, 6)
iv) To generate random no from a normal distribution (n random variable) the R code is rnow (n, u, 6)

code

```
> P1 = pron(15, 12, 5)
> P1
 [1] 0.8415447
> cat("P(x<=15) = P,]
   P(x<=15) = 0.841544
 P2 = Pnorm (15, 12, 15) - Pnorm(10,12,5)
>P2
[1] 0.578066
> cat (" P(10 < x ≤ 13) = ", P2)
   P(10 ≤ x ≤ 15) = 0.3780661
> P3 = 1 - pnorm (14, 12, 5)
> P3
> cat("P(x > 14) = "P3)
   P(x>14) - 0.2524925
 P4 = rnorm (5, 12, 3)
 P4
[1]15254723      16.25505   11.28051S
                 12.27690
```

PRACTICAL : 4

06

* Aim :- Normal distribution.

i) P(x ≤ x) - dnorm (x, u, 6)

ii) P(x ≤ x) = pnorm (x, u, 6)

iii) P(x > x) = 1 - pnorm (x, , u, , 6)

iv) To generate random no from a normal distribution (n random variable) the R code is rnow (n, 4, 6)

Code

```
> P1 = pron(15, 12, 5)
> P1
[1] 0.8415447
> cat("P(x <= 15) = P, ]
P(x <= 15) = 0.841544)
P2 = Pnorm (15, 12, 15) - Pnorm (10, 12, 5)
> P2
[1] 0.5780661
> cat(" P(10 < x ≤ 13) = ", P2)
P(10 ≤ x ≤ 15) = 0.3780661
> P3 = 1. pnorm (14, 12, 5)
> P3
> cat("P(x > 14) = "P3)
P(x>14) - 0.2524925
P4 = rnorm (5, 12, 3)
P4
[1] 15254723    16.25505   11.280515
12.27690
```

2) x follows normal distribution with
$\mu = 10$, $\sigma = 2$
find i) $P(x \le 7)$ ii) $(5 < x < 12)$ iii) $P(x$
iv) generate 10 observation.

code :-

```
> a1= pnorm(7,10,2)
> a1
> a2 - pnorm (5,10,2)-pnorm(12,10,2
> a2
[1]- 0.835151
> a3=1-pnorm(12,10,2)
> a3
[1] 0.1586553
> a4
[1] 11.608931   9.92041   12.63741  8.07
    8.723100    9.193762  9.366861  11.70
    9.53784     10.715006
> a5
[1] 9.4993306
```

Q.3)
code:-

```
> rnorm (15, 15, 4)
[1]  10.7649   7.79324   9.953444
                         13.345904
   17.50998
> am = mean (x)
 [1]  11.8735
> cat(" Sample mean is =" , am)
  Sample mean is = 11.87345
> me = median (x)
> me
 [1] 10.76499
> cat ("Median is =", me)
  median is =10.76499
> n =5
> V= (n-1) * var (x)/n
> SD=sqrt(V)
> SD
 [1] 3.3163
> cat ("S.D is =, 30)
  SD is= 3.331613
```

Q.4) $x \sim N(30, 100)$, $\sigma = 10$

i) $P(x \leq 40)$

ii) $P(x > 35)$

iii) $P(25 < x < 35)$

iv) find k such that $P(x < k) = 0.6$

```
> f1 = pnorm (40,30, 10)
> f1
[1] 0.8413447
> f2 = 1- pnorm (35,30,10)
f2
[1] 0.3085375
> F3 = pnorm (25,30,10)- pnorm
                (35,30,10)
> f3
[1]- 0.382949
> f4 = qnorm (0.6, 30, 10)
  [1] 32.53347
```

Q.5) Plot the standard graph

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

PRACTICAL : 5

Topic : Normal and t-test

1) $H_0 : \mu = 15$ $H_1 : \mu \neq 15$

Test the hypothesis

Random sample of size 400 is drawn and it is calculate. The sample mean is 14 & S.D is 3 test the hypothe at 5% level of significance.

0.05 > accept the value

0.05 < less the reject

```
> mo = 15
> mx = 14
> sd = 3
> n = 400
> zcal = (mx - mo)/(sd/(sqrt(n)))
> zcal
[1] -6.6667
> cat ("calculated value of Z is = ", zc
calculated value of Z is = -6.6667
> pvalue = 2 * [1 - pnorm (abs(zcal
> pvalue
[1] 2.616769e-11
```

∴ The value is less than 0.05 we will reject the value of $H_0 = \mu = 15$

2)

```
> m= 10
> n = 400
> mx = 10.2
> Sd = 2.25
> zcal = (mx - mo) /(sd/ sqrt(n)))
>zcal
[1] 1.7778
> pvalue = 2 * (1 - pnorm (abs (n)))
> pvalue
[1] 0.07544036
```

∴ The value is accepted.

3) Test the hypothesis: in college is 0.2
A sample is collected & calculated the sample proportion al as 0.125.
Test the hypothesis at 5%.

```
> P= 0.2
> p= 0.125
> n= 400
> Q= 1-d
> zcal= (p-P) /(sqrt (P * Q/n))
> cat ("calculated value of z is =", zcal)
[1] calculated value of z is =-3.75
>pvalue = 2 * (1- pnorm (abs (zcal)))
> P value.
```

5) Test the hypothesis H0: μ=12.5 the foll sample at 5% level of significance.

```
> x = c(12.25, 11.97, 12.5, 12.08, 12.3,
        12.28, 11.94, 11.94, 11.89, 12.16,
        12.04)
> n=length(x)
> n
[1] 10
> mx = mean (x)
> mx
[1] 12.107
> variance = (n-1) * var (x) / n
> variance
[1] 0.019521
> sd = sqrt (variance)
> sd
[1]= 0.1397176
> m0 = 12.5
> t = (mx - m0) / (sd / sqrt (n))
> t
[1] -8.89409
> pvalue = 2 * (1- pnorm (abs
                           (t)))
> pvalue
[1] 0
```

PRACTICAL - 6

• Aim : Large Sample Test

1) let the population mean (the amount spent per customer in a restaurant) is 250. sample of 100 customers selected is calculated as 275 and 30 test the not 5%. level of significance.

2) In a random sample of 1000 students it is found that 750 use blue pen test the hypothesis that the population proportion is 0.8 at 1% level of significance.

Solⁿ) 1)

```
> mo = 250
mx = 275
> sd = 30
> n = 100
> zcal = (mx - mo)/(sd / (sqrt(n)))
> cat("calculated value of z is =", zcal)
[1] calculated value is = 8.3333
> pvalue = 2 * (1 - pnorm(abs(zcal)))
> pvalue
[1] 0
```

∴ The value is less than 0.05 we will reject the value of $\mu_0 = \mu = 250$

Scanned with CamScanner

if (pvalue > 0.05) { cat("accept
else { cat ("reject Ho") }
reject Ho>

Q.2) Sol^n
Ho : There is no difference between two groups.
> x = c(18, 22, 21, 17, 20, 17, 22, 21)
> y = c(16, 20, 14, 20, 18, 13, 21)
> t.test (x, y)
Welch two sample t-test
data : x and y
t = 2.2513, df = 16.316, p value 0.1628205
95 percent confidence interval
Sample estimates

Since p-value is less than 0.5

we reject hypothesis at 5% level of significance.

3) The sales data of 6 shoes before & after a special campaign are given below.
12
Before : 53, 28, 31, 48, 50, 42
After : 58, 29, 30, 55, 56, 45
Test the hypothesis that the campaign is affective or not
Ho - There is no significance difference of before & after campaign.

x = c(Before)
y = c(After)
> t.test(x, y, paired = T, alternative = "greater")
paired t-test
data : x & y
t = -2.7815 df = 5 pvalue = 0.9806
alternative hypothesis.
True difference in means is greater than 0.
-6.0334, + inf
Sample estimated mean of difference -3.5
p value is greater than 0.05, we accept the hypothesis at 5% level of significance.

4) Following the program is before & after the program is off

Before : 120, 125, 115, 130, 123, 115

After : 100, 114, 95, 90, 115, 99

soln: Ho there is also bio signific difference.

>x = c(Before)

>y = 6 (After)

t-test (x, y, period = T, alternative

data : x & y

t = 4.3458, df = 5. Pvalue = 0.99

alternative hypothesis : true in is less than 0.

95 percent confidence interval inf 29.0295

sample estimates.

19.83833

∴ Pvalue is greater than 0.05 we accept the hypothesis 5% level of significance

2) So 1"

> P = 0.8

> q = 1-P

> n = 1000

> p = 750/1000

> z (a) (p-P)/(sqrt(P* q/n))

> cat("calculated value of z is

: -3.952847

> pvalue

(1) 7.72868e-05

3) n1 = 1000

> n2 = 2000

> mx1 = 67.5

> mx2 = 68

> sd1 = 2.5

> sd2 = 2.5

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

> z cal

[1] -5.163928

> pvalue

> pvalue

[1] 2.41756e-6.7

∴ (Rejected

PRACTICAL - 7

Topic: Small sample test

Q.1 The marks of 10 Students are given by 63, 63, 66, 67, 68, 69, 70, 72, 71. Test the hypothesis that sample comes: Test the hypothesis that the sample comes from a population with average marks 66.

Soln: $H_0 : \mu = 66$

```
> x = c(63, 63, 66, 67, 68, 69, 70, 70,
        71, 72)
> t.test(x)
        one sample t-test
data : x
t = 68.319 , df = 9 , p-value = 1.558
  alternative   hypothesis:       e-13
true mean is not equal to 95 percent
confidence interval 65.65171
  70.146829
sample estimates:
  mean x
     67.9
```

Since p-value is less than 0.05 we reject hypothesis at 5% level of significance.

```
> Los = 0.05
> pvalue = 1.558e-13
```

if (pvalue >0.05) { cat("accept
else { cat ("reject Ho") }
reject Ho>

Q.2) Solⁿ
Ho : There is no difference between two groups.
>x = c(18, 22, 21, 17, 20, 17, 22, 21)
>y = c(16, 20, 14, 20, 18, 13, 21)
> t.test (x, y)
welch two sample t-test
data : x and y
t = 2.2513, df = 16.316, p value
95 percent confidence interv
0.1628205
Sample estimates

Since p-value is less than 0.5

we reject hypothesis at 5% level of significance.

3) The sales data of 6 shoes before & after a special campaign are given below. 12

Before : 53, 28, 31, 48, 50, 42

After : 58, 29, 30, 55, 56, 45

Test the hypothesis that the campaign is affictive or not

Ho. There is no significance difference of before & after campaign.

x = c(Before)

y = c(After)

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

pared t test

data : x & y

t = -2.7815 df = 5 p value = 0.9806

alternative hypothesis.

True difference in means is greater than 0

-6.0334, + inf

Sample estimated mean of difference -3.5

P value is greater than 0.05, we accept the hypothesis at 5% level of significance.

4) Following are weights before & after the program is effec

Before : 120, 125, 115, 130, 123, 116

After : 100, 114, 95, 90, 115, 199

Soln: Ho there is also bio significa difference.

> x = c(Before)

> y = c(After)

t-test (x, y, period = T, alternative = "

data : x & y

t = 4.3458, df = 5, pvalue = 0.9963

alternative hypothesis : true diffe in is less than 0.

95 percent confidence interval

inb 29.0295

sample estimates.

19.83833

∴ Pvalue is greater than 0.05 we accept the hypothesis at 5% level of significance.

PRACTICAL-8

Topic : Large & Small Test

1) $H_0 = \mu = 55$, $H_1 = \mu \neq 55$
> n = 100
> 1x = 52
> mo = 55
> sd = 7
> z cal = (nx - mo) c(sd/sqrt(n)))
> [1] -4.28714
> pvalue = 2 * (1. pnorm (abs (zcal)))
> pvalue
[2] 1. 82153 e^-0.5

As pvalue is less than 0.05 we reject H at 9% level of significance.

2) $H_0 = P = 0.5$ against $H_0 : P \neq 0.5$
> P = 0.5
> q = 1 - P
> n = 700
> z cal = (P - P) / (sqrt (P * q/n))
> z cal
[1] 0
> Pvalue = 2 * (1 - pnorm (abs (zcal))
> pvalue
[2] 1

As pvalue is greater 0.05 we accept Ho at 1% level.

* >H1=1000
>n2=1500
>P1=2/1000
>P2=1/1500
>P=(n1*P1+n2*P2)/n1+n2
>P
[1] 0.0012
>q=1-P
[1] 0.988
>zcal (P1-P2)/sqrt(p*q*(1/n1+n/n2)
>zcal
[1] 0.9433752
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.345489

∴ Pvalue is greater than 0.05 we accept Ho and 5% level of significance.

*
Soln: Ho = μ = 100 against H_1 $\mu \neq 100$

> var = 64
> n = 400
> mo = 100
> mx = 90
> sd = sqrt (var)
[1] 8
> zcal = (mx - mo)/ sd/ sqrt (n))))
> zcal
> zcal
[1] 2.5
pvalue : 2 * (1 - pnorm (abs (zcal)))
> pvalue
> 0.01924193
since pvalue is less than 0.05
we 5% level of significance.
~~Since~~ . Sample estimates
mean of x
68.14286
since pvalue is less than 0.05
will reject Ho at 1%.

PRACTICAL - 9

Topic - Non parametric testing of hypothesis using R-environment

1) The following data represents earnings (in dollars) for a random sample of five common stacks listed

Data :- 168 , 3.35, 2.50 , 6.25 , 3.2

```
>x <- c(1.68,3.35,2.50,6.23,3.24)
> x< - length (x);
> n
> x > 4
[1] FALSE FALSE TRUE TRUE
> s<- sum (x>4); s
[1] 1
>binom.Test (5, n, P= 0.5, alternative
        =" greater"); exact
    binomial test.
```

data 95 confidence interval

0.0102062 2 1.000000

sample estimates

probability of success 0.2

The scores of 8 students in reading before & after lesson are as follows
Test whether there is effect of reading.

Inspector	1	2	3	4	5	
caliper 1	0.265	0.268	0.266	0.269	0.264	
caliper 2	0.263	0.262	0.270	0.271	0.260	

caliper 1 & caliper 2 are are

Code:-

```
>x <- c (0.265, 0.268, 0.266, 0.267,
  0.269, 0.264)

>y <- c (0.263, 0.262, 0.261, 0.271,
  0.260)

>wilcox-test (x,y, alternative=
  "greater")
        wilcoxon rank test.
 data : x and y
  w = 24, p=0.197
```

alternative hypothesis
: three location shift is
than 0'

∴ p value is greater than
0.05 we accept it

) The scores of 8 students in reading before and after lessons are as follows

16

tudent No	1	2	3	4	5	6
re before	10	15	16	12	09	07
re after	13	16	15	13	09	10

DE :-

```
 <- c( 10, 15, 16, 12, 09, 07, 11, 12);
 <- c( 13, 16, 15, 13, 08, 13, 10))
D <- b - a ;
wilcox-test(D, alternative "greater")
```

wilcox-or signed rank list with continuity.

In wilcox-test default (n, alternative - "greater") cannot complete exact p- value value with this

∴ p- value is greater than 0.05 we accept it

PRACTICAL 10

Aim :- chi square test & ANOVA (Analysis of variances)

Q.1) Use the following data to test whether the condition of home & condition of child are not.

Child	Home Clean	Dirty
clean	70	50
fairly clean	80	20
Dirty	35	45

H0 :- Condition of Home & Child are independent

> x = c(70, 80, 35, 50, 20, 45)
> m = 3
> n = 2
> matrix(x, nrow=m, ncol=n)
> y

	[,1]	[,2]
[1,]	70	50
[2,]	80	20
[3,]	35	45

```
> pu = chisq.test[y]
> pu
```

Pearson'. chi-squared test

data : y
x-square = 25.646
df = 2
p-value = 2.6986e06

∴ They are dependent

Test the hypothesis that vacation & disease are indepent or Not.

Q) Test the hypothesis that vaccation & disease are independent or not

vaccine

	Affected	Not
Disease Affect	70	46
Non-affected	35	37

Ho: Disease & vaccine are independent.

```
x = c(70, 35, 46, 37)
> m = 2
> n = 2
> y = matrix(x, nrow = m,
        ncol = n)
> y
        [,1]    [,2]
[1]      70      46
[2]      35      37

> PV = chisq.test(y)
> PV
  pearson's chi squared
    test with Yates
      continuity correction
```

data : y

x-square = 2.0275

df = 1

p-value = 0.1545

P-value is more than 0.05

we accept the hypothes

// They are Independent