28/11/19

Semester - II

Practical No. 1

Basics Of R-Software

→ R is a software for statistical analysis
computing.
→ It is an effective data handling
outcome storage is possible.
→ It is compable of graphical
→ It is a free software

Q1) Solve the following:

1.) > $4+6+8 \div 2-5$
[1] 7.3333

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

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

4.) $\sqrt{4^2+5\times3+7/6}$
[1] 5.671567.

5.) Round off $(46 \div 7 + 9 \times 8)$
[1] 79.

Scanned with CamScanner

Q2) Solve the following :-

1.) a = c (2,3,5,7) * 2

a

[1] 4 6 10 14

2.) b = c (2,3, 5,7) * c (2,3,[illegible])

b

[1] 4 9 10 21

3.) c = c (2,3,5,7) * c (2,3,6,4)

c

[1] 4 9 30 28.

4.) d = c(1, 6,2,3) * c (-2, -3, -4, -1)

d

[1] -2 . -18 -8 -3

5.) e = c (2,3,5,7) ^ 2

e

[1] 4, 9, 25 49

6.) f = c(4, 6, 8, 9, 4, 5) ^ c (1,2,3)

f

[1] 4 36 512 9 16 125

7.) g = c (6,2,7,5) / c (4,5)

g

[1] 1.50 0.40 1.75 1.00

:02

Q5) Find x+y and 2x+3y where,

$$x = \begin{bmatrix} 4 & -2 & 6 \\ 7 & 0 & 7 \\ 9 & -5 & 3 \end{bmatrix} \quad y = \begin{bmatrix} 10 & -5 & 7 \\ 12 & -4 & 9 \\ 15 & -6 & 5 \end{bmatrix}$$

```
→ x = matrix (nrow = 3, ncol = 3, data = c(4,7,9,-2,0,-5,67,3))
x

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

y = matrix (nrow = 3, ncol. = 3 , data = c(10, 12, 15, -5, -4, -6, 7, 9, 5)

> a = x + y
> a
      [,1] [,2] [,3]
[1,]   14   -7   13
[2,]   19   -4   16
[3,]   24   -11.  8

> b = 2x + 3y
> b
      [,1] [,2] [,3]
[1,]   38   -19   33
[2,]   50   -12   41
[3,]   63   -28   21
```

5/12/19

Practical No. 2

Probability Distribution

Q1) Check whether the following are p.m.f or not.

1)

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

2)

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

3)

x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1

Ans 1) $\because$ $P(2) = -0.5$, can not be a probability mass function.
$\because$ $P(x) \geq 0$ for all x.

Ans 2) It cannot be a probability mass function as p.m.f is not equals to 1.

```
> prob = c(0.2, 0.2, 0.3, 0.2, 0.2)
> sum(prob)
[1] 1.1
```

Ans 3)

```
> prob = c(0.2, 0.2, 0.35, 0.15, 0.1)
> sum(prob)
[1] 1.
```

Hence, it is a probability mass function.

Q2) 1) Find c.d.f. for the following p.m.f & sketch the graph.

x	10	20	30	40	50
$p(x)$	0.2	0.2	0.35	0.15	0.1

```
>prob=c(0.2,0.2,0.35,0.15,0.1)
>sum(prob)
[1] 1
>cumsum(prob)
[1] 0.20  0.40  0.75  0.90  1.00
```

$\rightarrow F(x) = 0 \quad x < 10$

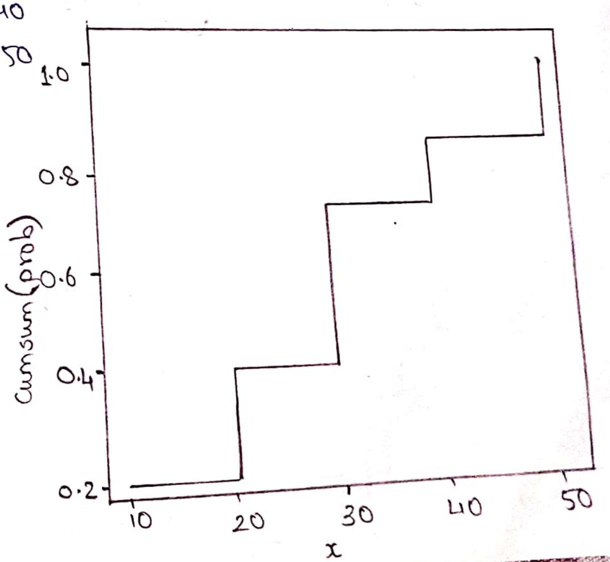
$0.2 \quad 10 \leq x < 20$

$0.4 \quad 20 \leq x < 30$

$0.75 \quad 30 \leq x < 40$

$0.90 \quad 40 \leq x < 50$

$1.0 \quad x \geq 50$.



Q. 2)

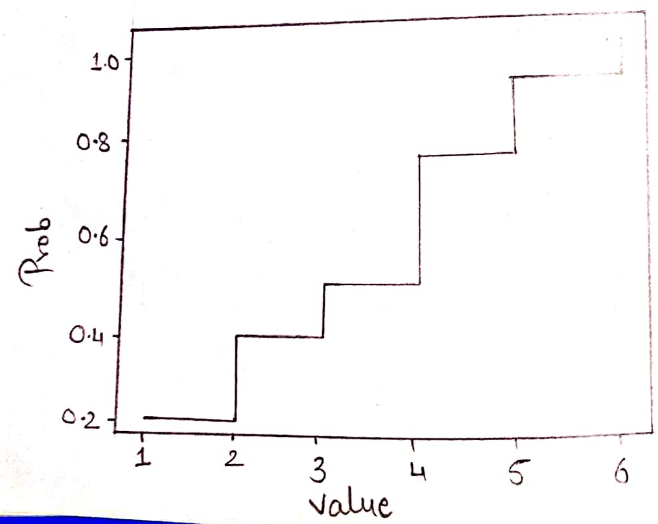
x	1	2	3	4	5	6
$P(x)$	0.15	0.25	0.1	0.2	0.2	0.1

→
```
> prob = c(0.15, 0.25, 0.1, 0.2, 0.2, 0.1)
> sum(prob)
[1] 1
> cumsum(prob)
[1] 0.15  0.40  0.50  0.70  0.90  1.00
```

$F(x) = 0 \quad x < 1$

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	$x \ge 6$

```
> x = c(1,2,3,4,5,6)
> plot(x, xlab = "value", ylab = "prob", main = "cdf", cumsum(prob),
"s")
```



Q3 Check whether the following is p.d.f or not.

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

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

Ans 1) $f(x) = \int_0^1 f(x)\,dx.$

$= \int_0^1 3 - 2x.$

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

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

$= [3x]_0^1 - \cancel{2}\left[\frac{x^2}{\cancel{2}}\right]_0^1$

$= 3(1-0) - (1-0)^2$

$= 3(1) - 1$

$= 3 - 1$

$= 2. \quad \therefore$ It is not a p.d.f.

Ans 2) $f(x) = \int_0^1 f(x)\,dx.$

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

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

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

$= (1-0)^3$

$= 1^3$

$= 1 \quad \therefore$ It is a p.d.f.

AM

Scanned with CamScanner

BINOMIAL DISTRIBUTION

Commands :-

1.) $P(X=x) = dbinom(x,n,p)$
2.) $P(X \leq x) = pbinom(x,n,p)$
3.) $P(X>x) = 1 - pbinom(x,n,p)$
4.) If x is unknown and $P_1 = P(X \leq x)$
then, $qbinom(P_1, n, \cancel{q}\, p)$

Q1.) Find the probability of exactly 10 success in 100 trials with $p = 0.1$

Q2.) Suppose there are 12 M.C.Q, each question has 5 option out of which one is correct. Find the probability of having ~~no.~~ 1.) exactly 4 correct answer
2.) at most 4 correct answer.
3.) More than 5 correct answer.

Q3.) Find the complete distribution when $n=5$ & $p=0.1$.

Q4.) $n=12$, $p=0.25$, find the following probability
1.) $P(X=5)$
2.) $P(X \leq 5)$
3.) $P(X>7)$
4.) $P(5<X<7) = P(X=6)$

Scanned with CamScanner

Ans i) > dbinom (10, 100, 0.1)
[1] 0.1318653.

Ans 2)
i) > dbinom (4, 12, 0.2)
[1] 0.1328756
ii) > pbinom(4, 12, 0.2)
[1] 0.9274445
iii) > 1- pbinom (5, 12, 0.2)
[1] 0.01940528

Ans 3) > dbinom (0:5, 5, 0.1)
[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001

Ans 4) > dbinom (5, 12, 0.25)
[1] 0.1032414.
ii) > pbinom (5, 12, 0.25)
[1] 0.9455978.
iii) > 1- pbinom (7, 12, 0.25)
[1] 0.00278151
iv.) > dbinom (6, 12, 0.25)
[1] 0.04014945

Ans 5) i.) > dbinom (0, 10, 0.15)
[1] 0.1968744

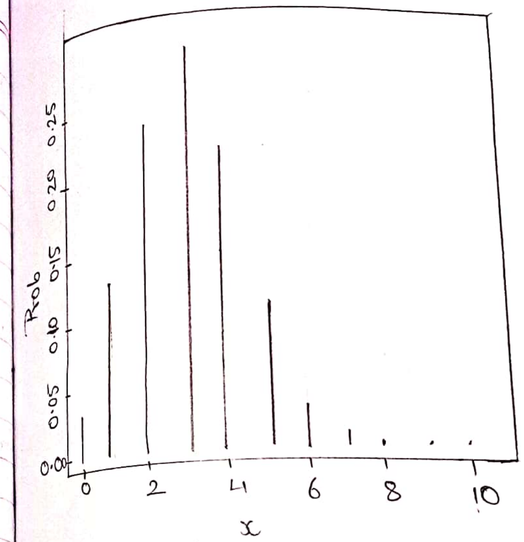
ii) 1- pbinom (3, 20, 0.15)
[1] 0.3522748

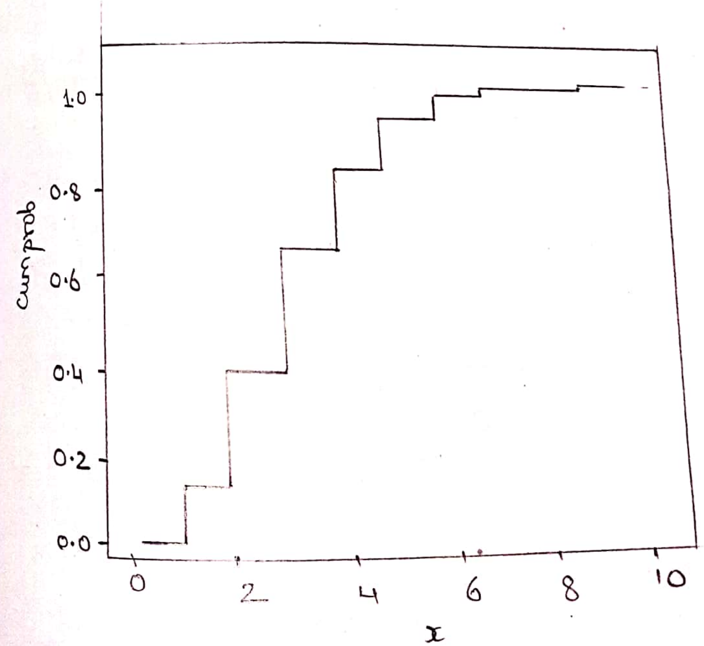
Ans 6) > qbinom (0.88, 30, 0.2)
[1] 9.

Ans 7) > n=10, p=0.3
> x=0:n
> prob = dbinom (x, n, p)
> cumprob = pbinom (x, n, p)
> d= data.frame ("xvalue" = x , "probability" = prob)
> print (d)

	xvalue	probability
1	0	0.0282
2	1	0.1210
3	2	0. 2334
4	3	0. 2668
5	4	0. 2001
6	5	0. 1029
7	6	0.0367
8	7	0. 0090
9	8	0. 0014
10	9	0. 0001
11	10	0. 0000059

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



19-12-19

1/1/20

PRACTICAL NO. 4

NORMAL DISTRIBUTION

1.) $P(X=x) =$ dnom (x, μ, σ)
2.) $P(X \le x) =$ pnom (x, μ, σ)
3.) $P(X > x) = 1-$ pnom (x, μ, σ)
where, μ = Mean
σ = Standard Deviation.

4.) To generate random numbers from a Normal Distribution (in random numbers), the R code is
= rnom (n, μ, σ)

) A random variable x follows Normal Distribution with mean $\mu = 12$, $\sigma = 3$, find i) $P(x \le 15)$ 2.) $P(10 \le x \le 13)$ iii) $P(x > 14)$ 4) Generate 5 observations (random numbers).

$\mu = 12$, $\sigma = 3$.

```
> pl = pnorm (15, 12,13)
> cat ('P(x <= 15):', pl)
  P(x<=15):. 0.8413447.
> P2 = pnorm (13,12,3) - pnorm(10,12,3)
> cat ('P(10<= x <=13):', p2)
  P(10<=x<=13): 0.3780661
> P3 = 1- pnorm (14,12,3)
> cat ('P(x>14):', P3)
  P(x>14): 0.2524425.
> rnorm (6,12, 3)
[1] 9.007282  14.034168  11.837712  14.460544  11.1125136
```

Q2) x follows Normal Distribution with $\mu=10$, $\sigma=2$. Find 1.) $P(x \le 7)$ 2.) $P(5 \le x \le 12)$ 3.) $P(x>12)$ 4.) Generate 10 observations, v.) Find k such that $P(x \le k) = 0.4$

```
> P1 = pnorm (7,10,2)
> cat ('P(x <=7):', P1)
  P(x<=7) : 0.0668072.
> P2 = pnorm (12,10,2) - pnorm (5,12,2)
> cat ('P(5<=x<=12):', P2)
  P(5<=x<=12) : 0.8551351
> P3 = 1- pnorm (12, 12,2)
> cat ('p(x>12):', p3)
  P(x>12): 0.1586553.
> rnorm
[1] 13.082059  13.997909  14.498445  9.950494  11.048673
[5] 13.554411  11.736256  13.776210  6.919454  11.062674
> qnorm (0.4, 10,2)
  [1] 9.493306.
```

Q3) Generate 5 random numbers from a Normal Distribution with $\mu = 15$ & $\sigma = 4$. Find Sample mean, median, S.D & print it.

```
> x = rnorm (5,15,4)
> x
[1] 17.94827  15.54403  16.86062  15.32839  24.94837
> am = mean (x)
> cat ('Arithmetic Mean is', am)
  Arithmetic Mean is 18.12612.
>med = median (x)
>cat ('Median is', med)
  Median is 16.86062
>n=5
>var = (n-2) * var(x)/n
[1] 9.399908
>sd = sqrt (var)
>cat ('S.D is', sd)
  S.D is 3.065927.
```


Q4.) ~~Fo~~ x follows Normal Distribution, $X \sim N(30,100)$, $\mu = 30$, $\sigma = 10$, Find 1.) $P(x \le 40)$ 2.) $P(x>35)$ 3.) $P(5<x<35)$ 4.) Find K such that $P(x<K) = 0.6$

```
> P1 = pnorm (40, 30,10)
> P1
[1] 0.8413447.
>P2 = 1 - pnorm (35, 30,10)
>P2
[1] 0.3085375
>P3 = pnorm (35,30,10) - pnorm (25,30,10)
>P3
[1] 0.3829249
> qnorm (0.6, 30,10)
[1] 32.53347.
```

AM

16/1/20

PRACTICAL No. 5

NORMAL AND t-TEST

1) Test the hypothesis $H_0: \mu = 15$ against $H_1: \mu \neq 15$. Random sample of size 400 is drop & it is calculated the sample mean is 14 & standard deviation is 3. Test the hypothesis at 5% level of significance

```
>mo = 15; mx = 14; sd = 3; n = 400
>zcal = (mx-mo)/(sd/(sqrt(n)))
>zcal
[1] -6.666667
>cat ("calculated value at 2 is = ", zcal)
calculated value at 2 is = -6.666667
>pvalue = 2 * (1-pnorm(abs(zcal)))
>pvalue
[1] 2.616796e-11
```

Since, pvalue is less than 0.05, we ~~reg~~ reject $H_0: \mu = 15$.

2.) Test the hypothesis $H_0: \mu = 10$ against $H_1: \mu \neq 10$. Random sample of size 400 is drawn with sample mean 10.2 & standard deviation ~~25~~ 2.25, test the hypothesis at 5% level of significance.

```
>mo = 10 ; mx = 10.2 ; sd = 2.25 ; n = 400
> zcal = (mx-mo)/(sd/(sqrt(n)))
> zcal
[1] 1.777778
> cat ("calculated value at 2 is = ", zcal)
Calculated value at 2 is = 1.777778
> pvalue = 2 * (1-pnorm (abs (zcal)))
> pvalue
[1] 0.07544036
```

Since, pvalue is more than 0.05, we accept $H_0: \mu = 10$

3.) Test the hypothesis (H_0), proportion of smokers in our college is 0.2. A sample is collected & ~~it~~ sample proportion is calcuted ~~at~~ as 0.125. Test the hypothesis at 5% level of significance (sample size is 400).

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

∵ pvalue is less than 0.05, we reject H_0 ~~:~~ = 0.2

4) Last year farmers lost 20% of their crops. A random sample of 60 fields are collected & it is found that 9 fields crops are insect polluted. Guess the hypothesis at 1% level of significance.

```
>P = 0.2, p = 9/60 ; n = 60 ; Q = 1-P
> zcal = (p-P) / (sqrt (P*Q/n))
> zcal
[1] -0.9682458
>cat ("calculated value at 2 is = ", zcal)
Calculated value at 2 is = -0.9682458.
> pvalue = 2* (1-pnorm (abs (zcal)))
> pvalue
[1] 0.3329216
```

∵ pvalue is greater than 0.05, we accept H_0.

03

5) ~~Use~~ Test the hypothesis $H_0: \mu_0 = 12.5$ from the following sample at 5% level of significance.

```
→>x=c(12.25, 11.97, 12.15, 12.08, 12.31, 12.28, 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 (var)
>sd
[1] 0.1397176
>zcal = (mx-mo)/(sd/(sqrt(n)))
>zcal
[1] -8.894909.
>cat ("calculated value is", zcal)
  Calculated value is -8.894909.
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0
```

∵ pvalue is less than 0.05, we reject Ho.

AM

PRACTICAL No. 6

Topic :- Large Sample Test.

Q1) Let the population mean (the amt. spent per customer in a restaurant) is 250. A sample of 100 customer selected. The sample mean is calculated as 275 & S.D is 30. Test the hypothesis at the population mean is 250 or not at 5% level of significance.

Q2) In a random sample of 1000 student, it is found that 750 use blue pen. Test the hypothesis that the population proportion is 0.8 at 1% level of significance.

→1) $H_0: \mu = 275$ against $H_1: \mu \neq 275$

```
> mo = 250 ; mx = 275 ; n=100 ; sd = 30
> zcal = (mx-mo) / (sd / (sqrt(n)))
>zcal
[1] 8.333333
>cat ("calculated value at 2 is =", zcal)
  calculated value at 2 is = 8.333333
> pvalue = 2*(1-pnorm (abs (zcal)))
>pvalue
[1] 0
```

∵ pvalue is less than 0.05, we reject Ho at 5% level of significance.

→2) Ho: $P = 0.8$ against $H_1 : P \neq 0.8$

$p = 0.8$,

$Q = 1 - P$

Q

[1] 0.2

```
> p = 750/1000 = 0.75
> n = 1000
> zcal = (p-P)/(sqrt(P*Q/n))
> zcal
[1] -3.952847.
> cat("calculated value at z is =", zcal)
calculated value at z is -3.952847.
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 7.72268e-05
```

∵ pvalue is less than 0.05, ∴ we reject Ho at 1% level of significance.

Q3) Two random sample of size 1000 & 2000 are drawn from 2 population with same S.D i.e. 2.5. The sample means are 67.5 & 68 respectively, Test the hypothesis $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ at 5% level of significance.

Q4) A study of noise level in 2 hospital given below. Test the claim that the 2 hospital has same level of noise at 1% level of significance.

	Hospital A	Hospital B.
Size	84	34
Mean	61.2	59.4
S.D	7.9	7.5

Q5) In a sample of 600 student in a clg, 400 use blue ink another college from a sample of 900 student 450 use blue ink. Test the hypothesis that the proportion of student using blue ink in 2 colleges are equal or not at 1% level of significance.

→ 3.) $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

$n_1 = 1000$; $n_2 = 2000$, $mx_1 = 67.5$; $mx_2 = 68$; $sd1 = sd_2 = 2.5$.

```
>zcal =(mx1-mx2)/ sqrt ((sd1^2/n1) + (sd2^2/n2))
>zcal
[1] -5.163978.
>cal ("zcalculated is =",zcal)
 calculated is = -5.163978.
>pvalue = 2* (1-pnorm(abs(zcal)))
>pvalue
[1] 2.417564e-07
```

∵ pvalue is less than 0.05, we reject Ho at 5% level of significance.

→ 4.) $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$

$n_1 = 84$; $n_2 = 34$; $mx1 = 61.2$; $mx2 = 59.4$; $sd1 = 7.9$; $sd2 = 7.5$

```
>zcal = (mx1-mx2)/sqrt((sd1^2/n1)+ (sd2^2/n2))
>zcal
[1] 1.162528
>cal ("zcalculated is =",zcal)
 calculated is = 1.162528.
>pvalue = 2* (1-pnorm (abs (zcal)))
>pvalue
[1] 0.2450211
```

∵ pvalue is greater than 0.05, ∴ we accept Ho. at ~~18~~1% level of significance

5) $H_0: P_1 = P_2$ against $H_1: \mu_1 \neq \mu_2$

$n_1 = 600$, $n_2 = 900$

$p1 = 400/600 = 0.666667$

$p2 = 450/900 = 0.5$

$p = (n_1 * p_1 + n_2 * p_2)/(n_1+n_2) = 0.56666667$

$q = 1-p = 0.4333333$

```
>zcal = (p1-p2) / sqrt (p*q*(1/n1 + 1/n2))
>zcal
[1] 6.381534.
>cat ("calculated is =", zcal)
 calculated is 6.381534
>pvalue = 2*(1-pnorm (abs(zcal)))
>pvalue
[1] 1.753222e-10
```

∵ pvalue is less than 0.05, we reject Ho. at 1% level of significance

6) ~~A~~= First, $n_1 = n_2 = 200$, $P_1 = 44/200$, $P_2 = 30/200$. Test hypothesis at 5% level of significance.

→) $n_1 = n_2 = 200$; $p1 = 44/200$; $p_2 = 30/200$

```
>p= (n1 * p1 +n2*p2)/(n1+n2)
>p
[1] 0.185
>q=1-p
>q
```

[1] 0.815

```
> zcal = (p1-p2)/sqrt (p*q*(1/n1 + 1/n2))
> zcal
[1] 1.802741
> cat ("calculated . is =", zcal)
  calculated is = 1.802741
> pvalue = 2*(1-pnorm (abs(zcal)))
> pvalue
[1] 0.07142888
```

∵ pvalue is greater than 0.05, ∴ we accept H_0 at 5 % level of significance.

27-01-20

PRACTICAL No. 7

Topic :- Small sample Test.

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

$H_0 : \mu = 66$ > x = c (63, 63, ..., 71, 72)

```
> n = length (x)
> n
[1] 10
> t.test (x),
        One sample t-test
```

Q.2.) Two groups of student score the following marks, Test the hypothesis that there is no level of significance difference between 2 groups.

Group A : 18, 22, 21, 17, 20, 17, 23, 20, 22, 21
Group B : 16, 20, 14, 21, 20, 18, 13, 15, 17, 21.

→ Ho: There is no difference between two groups
x = c(18, 22, 21, 17, ... 22, 21).
x = c(16, 20, ..., 17, 21)
t.test (x, y)
Welch Two sample t-test.
data : x and y
t = 2.2573, df = 16.376, pvalue = 0.03198.
alternative hypothesis : true difference in means is not equal to 0.
0.1628205 5.0371795
Sample estimates :
mean of x ~~xx~~ mean of y
20.1 17.5
> If (pvalue > 0.05) {cat ("Accept Ho")} else {cat ("Reject Ho")}
Reject Ho.

Q3.) Sales data of 6 shop before and after are special campaign given below:-
Before : 53, 28, 31, 48, 50, 42.
After : 58, 29, 30, 55, 56, 45.
Test the hypothesis the campaign is effective or not.

Ho: There is no significance difference of sales before and after the campaign.
> x = c(53, 28, 31, 48, 50, 42)
> y = c(58, 29, 30, 55, 56, 45)
> 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.
-6.035547 Inf.
Sample estimates :
mean of the differences
-3.5
> pvalue = 0.9806
> if (pvalue > 0.05) {cat ("Accept Ho")} else {cat (" Reject Ho")}
Accept Ho.

Q4.) Two medicines are applied to 2 group of patient
Group 1 : 10, 12, 13, 11, 14
Group 2 : 8, 9, 12, 14, 15, 10, 9
is there any significance difference between two groups.

Q5.) The following weights before and after a diet program, is diet program effective
Before : 120, 125, 115, 130, 123, 119.
After : 100, 114, 95, 90, 115, 99

→ 4.) Ho = There is no signifance difference between medicines

```
>a = c(10,12,13,11,14)
>b = c(8,9,12,14,15,10,9)
>t. test (a,b)
        Welch Two sample t-test.
data : a and b.
t= 0.65591 , df = 9.567, pvalue = 0.5273.
alternative hypothesis: true difference in mean is not equal to 0.
95 percent confidence interval:
  -1.934382    3.534382
Sample estimates:
mean of x    mean of y
  11.8         11.0.
>pvalue = 0.5273
>if (pvalue >0.05){cat ("Accept Ho")} else {cat("reject Ho")}
Accept Ho.
```

→Q5.) Ho = There is no difference between before and after.

```
>A = c(120,125,115,130,123,119)
>B= c(100,114,95,90,115,99)
>t.test (A,B, paired =T, alternative ="less")
            Paired t-test.
data: A and B.
t = 4.3458 , df = 5 , pvalue = 0.9963.
alternative hypothesis : true difference in mean is less than 0,
95 percent confidence interval :
      -Inf    29.0295
Sample estimates:
mean of the differences
       19.83333
>pvalue = 0.9963.
> if (pvalue >0.05) {cat ("Accept Ho")} else {cat ("Reject Ho")}
Accept Ho.
```

M

* Practical No. 8 *

Topic :- Large and small sample tests.

1) Ho : μ = 55 against H_1 : μ ≠ 55

```
>n= 100 , mx = 52, mo = 55, sd= 7.
>zcal = (mx - mo) / (sd / sqrt (n))
>zcal
[1] -4.285714
> cat (" calculated value at 2 is =",zcal)
calculated value at 2 is = -4.285714.
>pvalue = 2* (1- pnorm (abs (zcal)))
>pvalue.
[1] 1.82153 e - 05
```

∴ pvalue is less than 0.05, we reject Ho at 5% level of significance.

2) Ho : P = 0.5 against H_1 : P ≠ 0.5

```
>P= 0.5
> Q=1-P
>Q
[1] 0.5
>p= 350/700
>p
[1] 0.5
> n= 700
>zcal = (p-P)/ sqrt (P*Q/n)
> zcal
```

[1] 0
>cat ("calculated value at 2 is =", zcal).
calculated value at 2 is = 0 >
>pvalue = 2* (1-pnorm (abs (zcal)))
>pvalue
[1] 1

∵ pvalue is greater than 0.5, we accept Ho. at 1% level of significance.

4) n= 400 , $H_0: \mu = 99$ against $H_1: \mu \neq 99$
mo = ~~99~~ 100
mx = ~~100~~ 99
var = 64
>sd = sqrt (64)
>sd
[1] 8
zcal = (mx-mo) / sd /sqrt (n)
>zcal
[1] -2.5
> cat ("calculated value at 2 is =", zcal)
calculated value at 2 is = 2.5 >
> pvalue = 2* (1- pnorm (abs(zcal)))
>pvalue
[1] 0.01241933

∵ pvalue is less than 0.5, we reject Ho at 5% level of significance.

5) $H_0: \mu = 66$
>x=c (63,63,68,69,71,71,72)
>n= length (x)
>n
[1] 7
>t.test (x)
One Sample t-test.

data : x
t= 47.94, df=6, pvalue = ~~5.22~~ 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~~ ~~is~~ If (pvalue > 0.05) {cat ("Accept Ho")} else.
{cat ("Reject Ho")}.

Reject Ho.

7) $H_0: \mu = 1150$ against $H_1: \mu \neq 1150$.
n= 100
>mx = 1150
>mo = 1200
>sd = 125
>zcal = (mx-mo) /(sd /sqrt(n))
>zcal
[1] -4.
>cat (" calculated value at 2 is=", zcal)
calculated value at 2 is = -4.
>pvalue = 2* ((1- pnorm (abs (zcal))).
>pvalue
[1] 6.334248e-05

∵ pvalue is less than 0.5, we reject Ho.

6) Ho : ~~Ɛ~~ $\sigma_1 = \sigma_2$ against $H_1 : \sigma_1 \neq \sigma_2$.

```
> x = c (66,67,75,76,82,84,88,90,92)
> x = c (64,66,74,78,82,85,87,92,93,95,97)
> Var.test (x,y)
```

F test to compare two variance

data : x and y

F = 0.78803, num df = 7, denom df = 10, pvalue = 0.7737

alternative hypothesis : true ratio of variances if not equal to

1 . 95 percent confidence interval :

0.199509 3.751881

Sample estimates :

ratio of variance

0.7880255

∵ pvalue is greater than 0.05, Hence we accept Ho.

8) Ho: $P_1 = P_2$ against $H_1 : P_1 \neq P_2$.

```
> n1 = 200 , n2 = 300
> p1 = 44/200, p2 = 50/300
> P = (n1 p1 + n2 p2)/(n1+n2)
> p
[1] 0.2.
> q = 1-p.
[1] 0.8
> zcal = (p1 - p2)/sqrt (p*q* (1/n1 +1/n2))
> zcal
[1] 0.9128709.
```

Scanned with CamScanner

Practical No. 9

Topic :- Chi - square tests and ANOVA.

1.) Use the following data to test whether the condition of ~~hope~~ home & the condition & the child are independent or not.

		Condition of home clean	Dirty
C. of child	clean	70	50
	fairly c	80	20
	Dirty	35	45

→ Ho: Condition of 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.
data : y
X-squared = 25.646, df = 2, pvalue = 2.698e-0.6
```

∴ pvalue is less than 0.05, we reject Ho at 5% Level of significance.

2.) Test the hypothesis that the vaccination & the diseases are independent or not.

		Vaccine Aff.	Not Aff
Disease	Aff	70	46
	N.A.	35	37

→ Ho: Vaccination and Diseases are independent or not.

```
>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-squared = 2.0275, df = 1, p-value = 0.1545.
```

∴ pvalue is greater than 0.05, we accept Ho at 5% L.O.S

3) Perform a ANOVA for the following data.

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

→ Ho: The means are equal for A, B, C, D

```
> 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"
> one-way.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, pvalue = 0.00133

> 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     0.00133 **
Residuals   9   18.17    2.019
```

∵ pvalue is less than 0.05, ∴ we reject Ho at 5% LOS

4) Following data gives the lifes of the ~~typ~~ types of 4 brands.

Type	Life
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.

→ Ho: Average life of A, B, C, D are equal.

```
> x1 = c(20, 23, 18, 17, 18, 22, 24)
> x2 = c(19, 15, 17, 20, 16, 17)
> x3 = c(21, 19, 22, 17, 20)
> x4 = c(15, 14, 16, 18, 14, 16)
> d = stack(list(b1 = x1, b2 = x2, b3 = x3, b4 = x4))
> {names(d)
[1] "values" "ind"
> oneway.test(values ~ test, data = d, var.equal = T)

        Oneway analysis of means
data : values and ind
F = 6.8445, num df = 3, denom df = 20, pvalue = 0.002349
```

∵ pvalue is less than ~~0.05~~, we reject Ho at 5% L.O.S.

5) >x= read. csv ("C:/Users/admin/Desktop/marks. csv")

>x

	stats	maths.
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

>am = mean (x $ stats).
>am.
[1] 37.
>median = median (x $ stats)
> median.
[1] 38.5.
>n= length (x $ stats)
>n
[1] 10
>sd = sqrt ((n-1) * var (x $ stats) /n)
>sd
[1] 12.64911
> aml = mean (x $ maths)
> aml
[1] 39.4

>median= median (x $ maths)
>median.
[1] 37.
>n = length (x $ ~~st~~ maths)
>n
[1] 10
>sd= sqrt ((n-1) * var (x $ maths) /n)
>sd
[1] 12.64911

AM 20.3.2

PRACTICAL No. 10

TOPIC :- Non - Parametric test.

1) Following are the amount of sulphur oxide emitted by some industry in 20 days. Apply sign test to test the hypothesis that the population median is 21.5 at 5% L.O.S.

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

→ Ho: Population ~~meadi~~ median is 21.5

```
> x = c( 17, 15 ----- 24, 26)
> me = 21.5
> sp = length (x [x > me])
> sn = length (x [x < me])
> n = sp + sn.
> n
[1] 20
> pv = pbinom (sp, n, 0.5)
> pv
[1] 0.4119015.
```

∴ pvalue is ~~less~~ greater than 0.05, we accept Ho at 5% L.O.S

Note :- If a alternative is (H_1) not equal to me or me< then pv = pbinom (sp, n, 0.5) & H_1 : me> then pv = pbinom (sn, n, 0.5)

2.) Following is the data of 10 observations. Apply sign test, Test the hypothesis that the population median is 625 against the alternative which is more than 625.

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

→ Ho: population median is 625

```
>x = c(612, 619 ...... 663)
> me = 625
>sp = length (x[x>me])
>sn = length (x[x<me])
>n = sp+sn
>n
[1] 10
>pv = pbinom (sn,n,0.5)
>pv
[1] 0.0546875
```

∵ pvalue is greater than 0.05, ∴ we accept Ho.

3.) Following ~~are~~ are ~~of~~ a sample. Test the hypothesis, that the population median is 60 against the alternative which is ~~more~~ more than 60 at 5% L.O.S using Wilcoxon signed Rank Test.

63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 69, 48, 66, 72, 63, 87, 69

→ Ho: population median = 60

H1: population median > 60

```
>x = c(63, 65, ...., 69)
>wilcox.test (x, alter="greater", mu=60)
```

Wilcoxon signed rank test with continuity correction

data: x.

V = 128, pvalue = 0.03338

alternative hypothesis: true location is greater than 60.

∴ pvalue is less than 0.05, we reject Ho at 5% L.O.S.

Note: If alternative is equal to then, alter="less"
if not equal to then, alter="two.sided".

4.) Using Wilcoxon test, Test the hypothesis, population median is 12 or less than 12.

15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26

→ Ho: population median is 12.

```
>x = c(15, 17, ....., 26)
>wilcox.test (x, alter="less", mu=12)
```

Wilcoxon signed Rank test with continuity correction

data: x

V = ~~145~~ 66, pvalue = 0.9986

alternative hypothesis: ~~true~~ location is less than 12.

∵ pvalue is greater than 0.05, ∴ we accept Ho at 5% L.O.S.

5.) The weights of students before and after they stop smoking are given below. Using WSRT test, there is no significance change.

Weight before	Weight after.
65	72
75	74
62	72
72	66
	73

→ H_0: Before & after, there is no change.
H_1: There is change.

```
> x = c(65, ..., 72)
> y = c (72, ... 73)
> d = x-y
> wilcox.test (d, alter = "two.sided", mu = 0)
```

Wilcoxon signed rank test with continuity correction.

data: d

V = 4.5 , pvalue = 0.4982.

alternative hypothesis: true location is not equal to 0

∵ pvalue is greater than 0.05, ∴ we accept Ho at 5% L.O.S.

AM
27/2