PERFORMANCE

Term	Remarks	Staff Member's Signature
I	Complete	A-v
II	Complete	A-v

Exam Seat No.________________

THAKUR COLLEGE OF SCIENCE & COMMERCE

NAAC Accredited with Grade "A" (3rd Cycle)

THAKUR TRUSTS ®

ISO 9001 : 2015 Certified

Degree College

Computer Journal

CERTIFICATE

SEMESTER II UID No. ________________

Class Fy Bsc CS Roll No. 1709 Year 2019-20

This is to certify that the work entered in this journal is the work of ~~Mst.~~ / Ms. Sanjana Alwe

who has worked for the year 2019-20 in the Computer Laboratory.

A.Vral

Teacher In-Charge

Head of Department

Date : 27.02.20

Examiner

★★ INDEX ★★

PRACTICAL - 1

Topic: Basics of R Software

- R is a software for statistical analysis and data computing.
- It is an effective data handling software and outcome storage is possible.
- It is capable of graphical display.
- Its a free software

Q.1 Solve the following

1. $4+6+8 \div 2-5$

```
> 4+6+8/2-5
[1] 9
```

2. $2^2+|-3|+\sqrt{45}$

```
> 2^2+abs(-3)+sqrt(45)
[1] 13.7082
```

3. $5^3+7\times5\times8+46\div5$

```
> 5^3+7*5*8+46/5
[1] 414.2
```

4. $\sqrt{4^2+5\times3+7\div6}$

```
> 4^2+5*3+7/6
[1] 5.671567
```

5. round off
$46 \div 7 + 9 \times 8$
```
> round (46/7+9*8)
[1] 79
```

Q.2.

1.
```
>c(2,3,5,7)*2
[1] 4 6 10 14
```

2.
```
>c(2,3,5,7)* c(2,3)
[1] 4, 9, 10, 21
```

3.
```
>c(2,3,5,7)*c(2,3,6,2)
[1] 4 9 30 14
```

4.
```
>c(1,6,2,3)*c(-2,-3,-4,-1)
[1] -2 -18 -8 3
```

5.
```
>c(2,3,5,7)^2
[1] 4 9 25 49
```

6.
```
>c(4,6,8,9,4,5)^c(1,2,3)
[1] 4 36 512 9 16 125
```

7.
```
>c(6,2,7,5)/c(4,5)
[1] 1.50 0.40 1.75 1.00
```

Q.3.

$x = 20$

$y = 30$

$z = 2$

1. $x^2 + y^3 + z$

```
> x^2+y^3+z
[1] 27402
```

2. $\sqrt{x^2+y}$

```
> sqrt(x^2+y)
[1] 20.73644
```

3. x^2+y^2

```
[1] 1300
```

Q.4.

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

Q.5. Find $x+y$, $2x+3y$

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

```
> x=matrix(nrow=3, ncol=3, data=c(4,7,9,-2,0,-5,
      6,7,3))
> y=matrix(nrow=3, ncol)=3, data=c(10,12,15,-5,-4,-6,
      7,9,5))
> x+y
      [,1] [,2] [,3]
[1,]   14   -7   13
[2,]   19   -4   16
[3,]   24  -11    8

> 2x+3y

      [,1] [,2] [,3]
[1,]   38  -19   33
[2,]   50  -12   41
[3,]   63  -28   21
```

Scanned with CamScanner

Marks of statistic of CS student

59, 20, 35, 24, 46, 56, 55, 45, 27, 22, 47, 58, 54, 40, 50, 32, 36, 29, 35, 39

```
> x=c(59,20,35,24,46,56,55,45,27,22,47,58,
      54,40,50,32,36,29,35,39)
> length(x)
[1] 20
>breaks =seq(20,60,5)
>a=cut(x,breaks,right= FALSE)
>b= table(a)
>c = transform(b)
>c
```

	a	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

A.K. 28/11

PRACTICAL - 2

Topic : Probability Distribution

Q.1 Check whether the following are p.m.f or not

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

Since $P(2) = -0.5$, it can not be a p.m.f as in p.m.f $P(x) \geq 0 \ \forall x$

Q2

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

It cannot be a p.m.f. As in p.m.f $\Sigma P(x) = 1$

Q.3

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

Code:

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

Q.2

i. Find c.d.f for the following p.m.f and sketch the graph

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

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

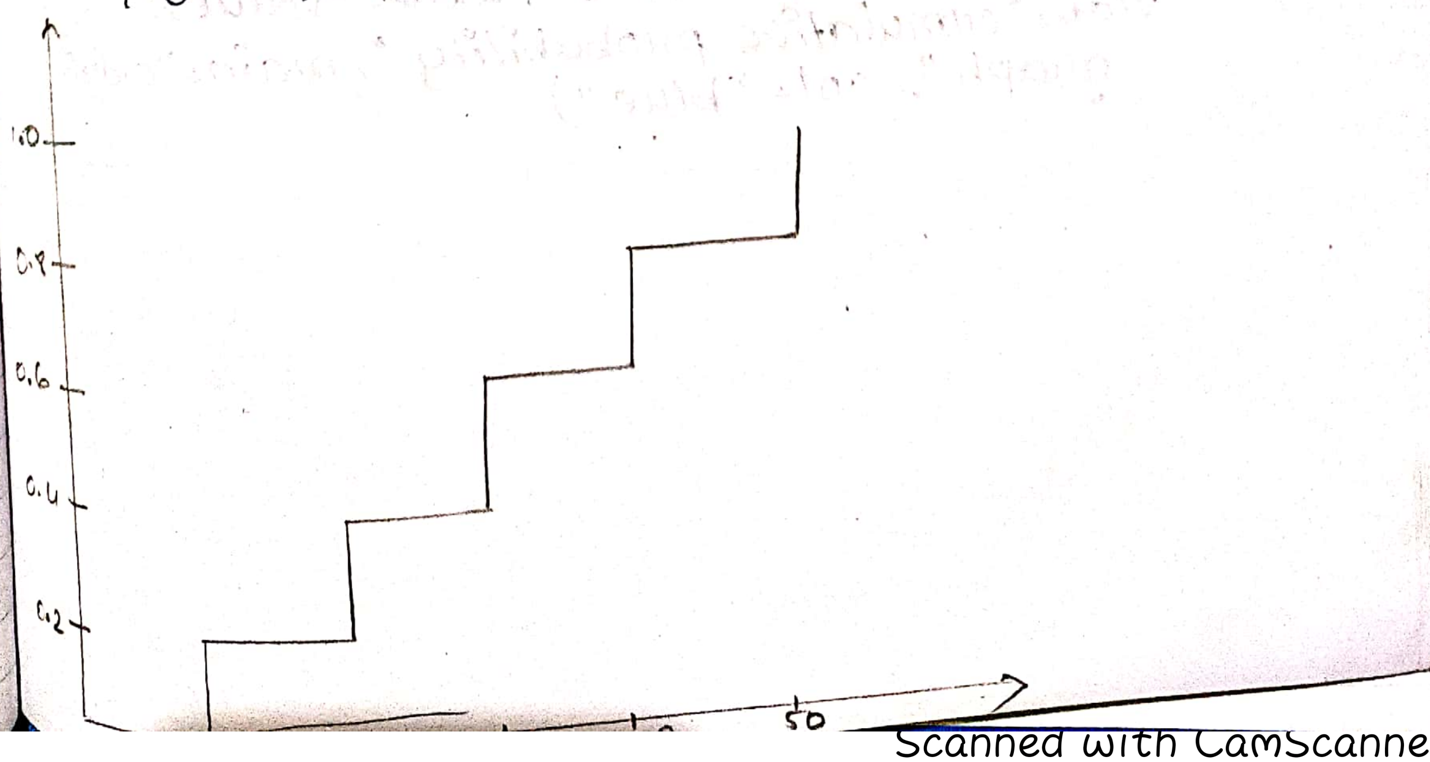
0.2 $10 \le x < 20$

0.4 $20 \le x < 30$

0.75 $30 \le x < 40$

0.95 $40 \le x < 50$

1.0 $x \ge 50$



8日

ii.

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

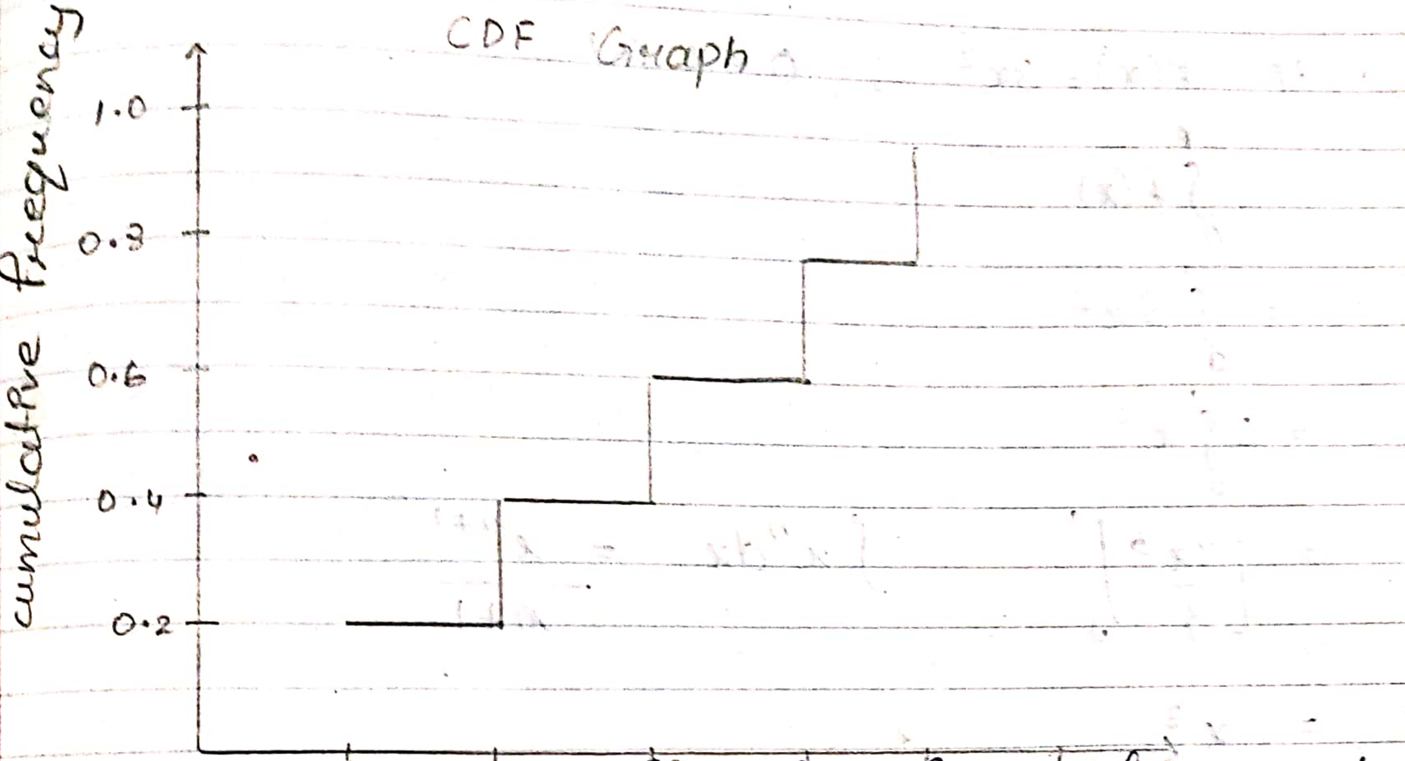
$$
\begin{aligned}
F(x) &= 0 && x < 1 \\
&= 0.15 && 1 \leq x < 2 \\
&= 0.25 && 2 \leq x < 3 \\
&= 0.1 && 3 \leq x < 4 \\
&= 0.2 && 4 \leq x < 5 \\
&= 0.2 && 5 \leq x < 6 \\
&= 0.1 && x \geq 6
\end{aligned}
$$

```
> 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
> x=c(1,2,3,4,5,6)
>plot(x,cumsum(prob),"s", xlab="values",
  ylab="cumulative probability",main="c.d.f
  graph", col="blue")
```

Q.3

i.

CDF Graph



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

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

$$\int_0^1 f(x)\,dx$$

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

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

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

$$= 2$$

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

$$\int_0^1 f(x)$$

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

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

$$= \left[\frac{3x^3}{3}\right]_0^1 \qquad \int x^n dx = \frac{x^{n+1}}{n+1}$$

$$= x^3$$

$$= 1.$$

Ans.

PRACTICAL - 3

Topic : Binomial Distribution

i. $P(X=x) = dbinom(x,n,p)$
ii. $P(X \leq x) = pbinom(x,n,p)$
iii. $P(X > x) = 1 - pbinom(x,n,p)$
iv. If x is unknown and P_1 is given as
$P(X \leq x) = qbinom(P_1, n, p)$

Q.1 Find the probability of exactly 10 success in 100 trials with $P = 0.1$

Q.2 Suppose there are 12 mcqs. Each question has 5 options out of which one is correct. Find the probability of having i. Exactly 4 correct answers ii. Atmost 4 correct answers iii. More than 5 correct answers $P = 1/5$

Q.3 Find the complete distribution when $n=5$, $p=0.1$

Q.4. $n=12$, $p=0.25$. Find the following probabilities
i. $P(X=5)$ ii $P(X \leq 5)$ iii. $P(X > 7)$ iv. $P(5 < x < 7)$

Q.5. A probability of a salesman making a sale to a customer is 0.15. Find the probability of i. No sells out of 10 customers ii. More than 3 sales out of 20 customers.

Q.6 A salesman has a 20% probability of making a sale to a customer out of 30 customers minimum number of sales he can make with 88% probability.

ns

Q.7. x follows binomial distribution with $n=10, p=0.$
plot the graph of p.m.f and c.d.f

Answers

Q.1

```
>dbinom(10,100,0.1)
[1] 0.1318653
```

Q.2

```
>dbinom(4,12,0.2)
[1] 0.1328756
>pbinom(4,12,0.2)
[1] 0.9274445
>1-pbinom(5,12,0.2)
[1] 0.01940528
```

Q.3

```
>dbinom(0:5,5,0.1)
[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.0000
```

Q.4

```
>dbinom(5,12,0.25)
[1] 0.1032414
>pbinom(5,12,0.25)
[1] 0.9455978
>1-pbinom(7,12,0.25)
[1] 0.00278151
>dbinom(6,12,0.25)
[1] 0.04014945
```

Q.5

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

Scanned with CamScanner

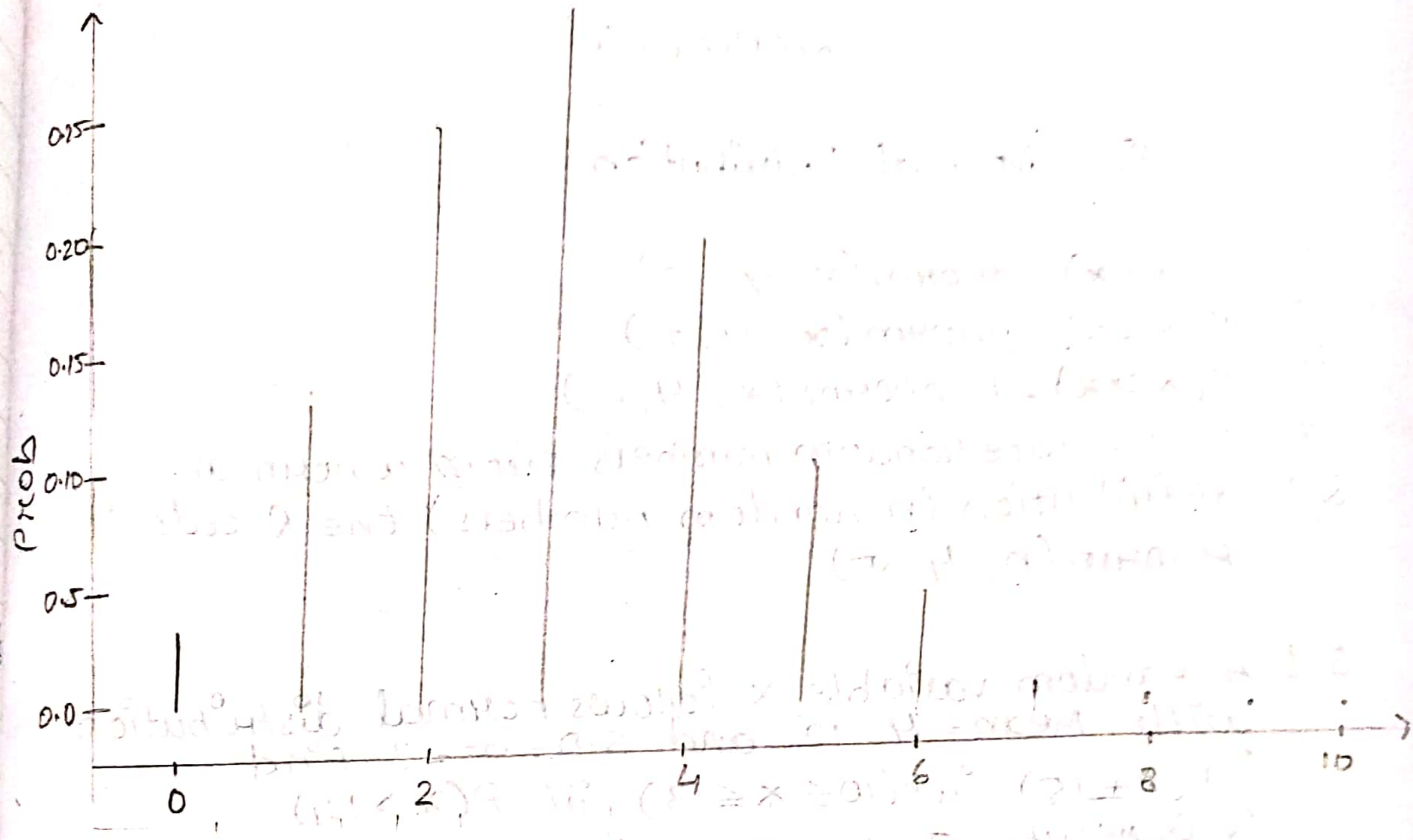
Q.6

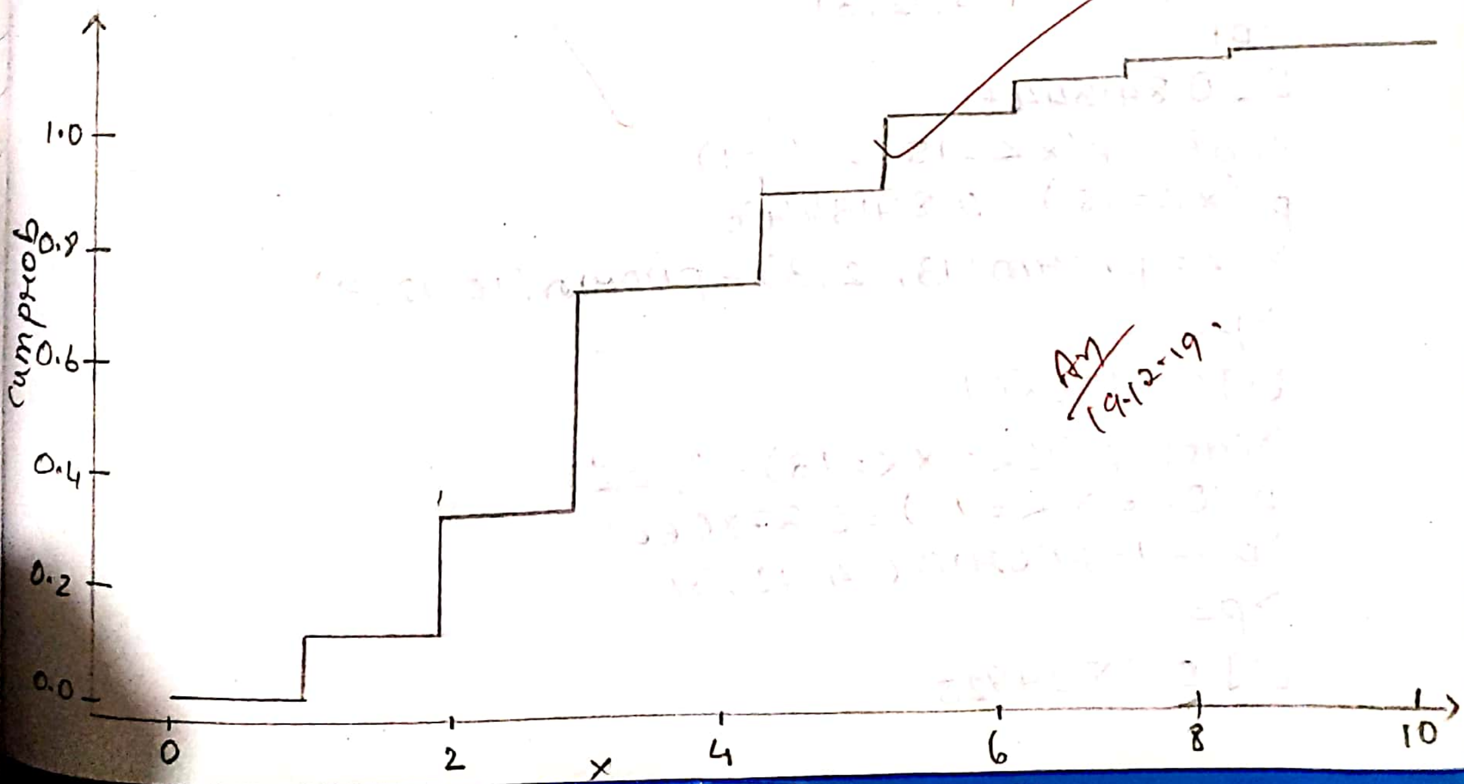
```
> qbinom(0.88,30,0.2)
[1] 9
```

Q.7

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

	xvalues	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	0.0000


prob
0.25
0.20
0.15
0.10
0.5
0.0
0
2
4
6
8
10


cumprob
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
x

PRACTICAL - 4

Aim : Normal Distribution

i. $P(X=x) = dnorm(x, \mu, \sigma)$
ii. $P(X \leq x) = pnorm(x, \mu, \sigma)$
iii. $P(X > x) = 1 - pnorm(x, \mu, \sigma)$
iv. To generate random numbers from a normal distribution (n random numbers) the R code is $rnorm(n, \mu, \sigma)$

Q.1 A random variable X follows normal distribution with Mean = $\mu = 12$ and S.D = $\sigma = 3$. Find
i. $P(X \leq 15)$ ii. $P(10 \leq X \leq 13)$ iii. $P(X > 14)$
iv. Generate 5 observations (random numbers)

```
>p1=pnorm(15,12,3)
>p1
[1] 0.8413447
>cat("P(X<=15) = ", p1)
P(X<=15) = 0.8413447
>p2 = pnorm(13,12,3) - pnorm(10,12,3)
>p2
[1] 0.3780661
>cat("p(10<=X<=13) = ", p2)
p(10<=X<=13) = 0.3780661
>p3=1-pnorm(14,12,3)
>p3
[1] 0.2524925
```

```
>cat("p(x>14) = ", p3)
p(x>14) = 0.2524925
>p4 = rnorm(5,12,3)
>p4
[1] 15.254723  16.548505  11.280515  6.419944  12.272460
```

Q.2 X follows normal distribution with $\mu = 10, \sigma = 2$
Find i. $P(X \leq 7)$ ii. $P(5 < X < 12)$ iii. $P(X > 12)$
iv. Generate 10 observations v. Find k such that
$P(X < k) = 0.4$

```
>a1 = pnorm(7,10,2)
>a1
[1] 0.0668072
>a2 = pnorm(5,10,2) - pnorm(12,10,2)
>a2
[1] -0.8351351
>a3 = 1 - pnorm(12,10,2)
>a3
[1] 0.1586553
>a4 = rnorm(10,10,2)
>a4
[1] 11.608931  9.920417  12.637741  8.073354
    8.721380  9.193726  9.366824  11.707106
    9.537584  10.715006
>a5 = qnorm(0.4,10,2)
>a5
[1] 9.493306
```

Q.3 Generate 5 random numbers from a normal distribution $\mu = 15$, $\sigma = 4$ find sample mean, median, S.D and print it.

```
x=rnorm(5,15,4)
[1] 10.7649  7.793249  9.953444  13.345904
    17.509668
> am=mean(x)
> am
[1] 11.87345
>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
> v
[1] 11.09965
~~>cat(~~
> sd= sqrt(v)
>sd
[1] 3.331613
>cat(" S.D is = ", sd)
S.D 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 $P(X < K) = 0.6$

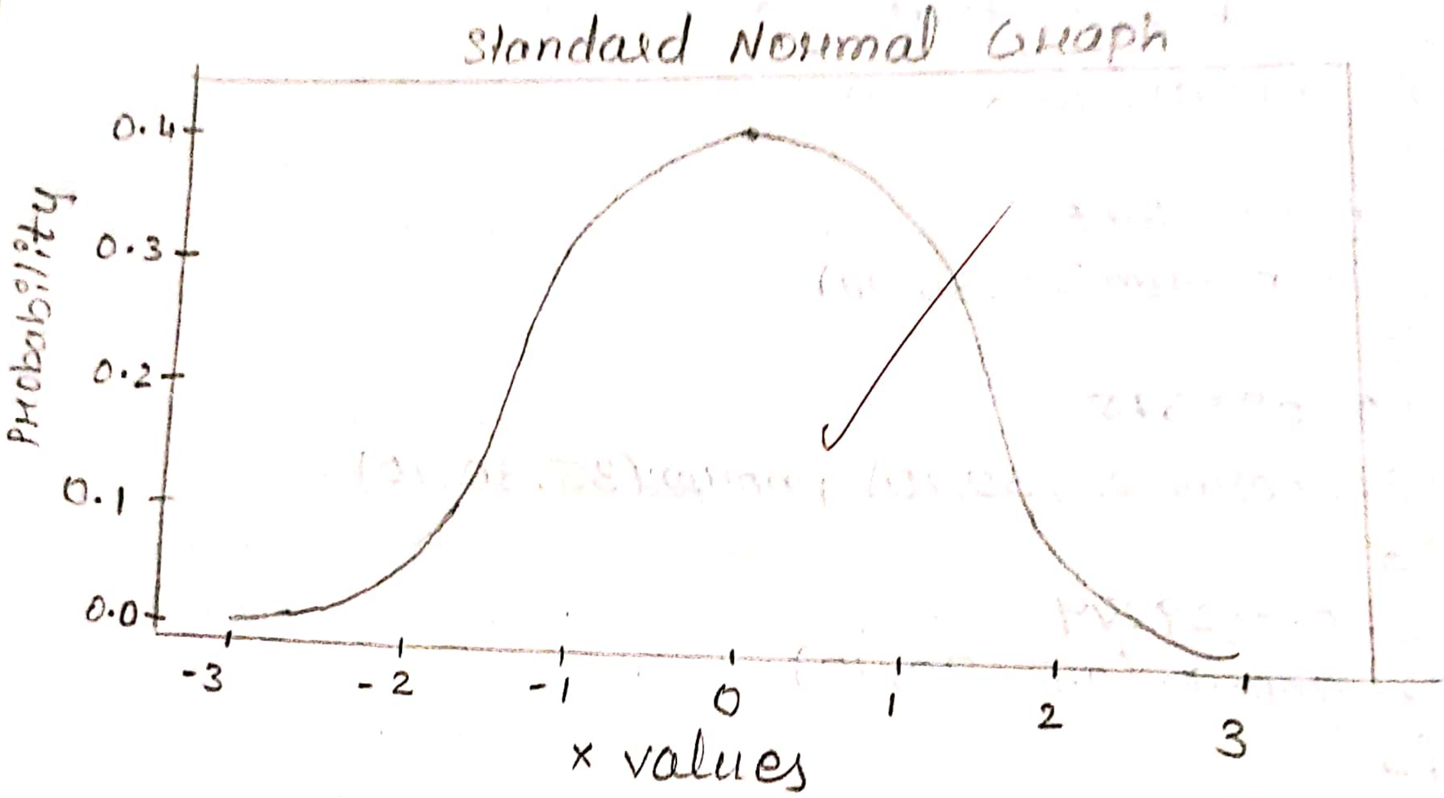
```
>f1=pnorm(40,30,10)
>f1
[1] 0.8413447
>f2=1-pnorm(35,30,10)
>f2
[1] 0.3085375
>f3=pnorm(25,30,10)-pnorm(35,30,10)
>f3
[1] -0.3829249
>f4=qnorm(0.6,30,10)
>f4
[1] 32.53347
```

Q.5. Plot the standard normal graph

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

Scanned with CamScanner

Ans:


Standard Normal Graph
0.4
0.3
0.2
0.1
0.0
Probability
-3
-2
-1
0
1
2
3
x values

PRACTICAL - 5

Aim: Normal and t-test

Q. Test the hypothesis $H_0 : \mu = 15$ $H_1 : \mu \neq 15$
Random sample of size 400 is drawn and it is calculated. Sample mean is 14 and S.D 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 of z is =", zcal)
calculated value of z is = -6.666667
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 2.616796e-11
```

Since pvalue is less than 0.05 we reject $H_0 : \mu = 15$

Q. Test the hypothesis $H_0 : \mu = 10$ $H_1 : \mu \neq 10$
Random sample of size 400 is drawn and its sample mean is 10.2 and S.D is 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 of z is = ", zcal)
calculated value of z is = 1.777778
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.07544036
```

since pvalue is more than 0.05 we accept $H_0 = \mu = 10$

Q.3. Test the hypothesis H_0: proportion of smokers in a college is 0.2. A sample is collected and it is calculated ~~the sample~~ 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] ~~4.74~~ -3.75
```

```
>pvalue= 2*(1-pnorm (abs(zcal))
>pvalue
[1] 0.000176 8346
```

Q. last year farmer's lost 20% of their crop. Random sample of 60 field are collected and it is found that a fied of crops is insect polluted. Test the hypothesis of 1% level of significance

```
>p=0.2
>p= 9/60
>n=60
>Q=1-P
>zcal=(p-P)/sqrt(P*Q/n))
>zcal
[1]-0.96842458
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.3329216
```

Q. Test the hypothesis $H_0: \mu = 12.5$ from the following sample of 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)
>v=(n-1)* var(x)/n
>sd= sqrt(v)
```

Scanned with CamScanner

```
>sd
[1] 0.1397176
>mo=12.5
>t=(mx-mo)/(sd/(sqrt(n)))
>t
[1]-8.894909
>pvalue=2*(1-pnorm(abs(t)))
>pvalue
[1]0
```

Ay/ 16-1-22

PRACTICAL - 6

Aim: Large sample test

Q.1 Let the population mean (the amount spent per customer in a restaurant) is 250. A sample of 100 customers selected. The sample mean is calculated as 275 and S.D is 30. Test the hypothesis that the population mean is 250 or not at 5% level of significance.

Q.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% of level of significance.

Q.1
```
> mx = 275
> m0 = 250
> n = 100
> sd = 30
> zcal = (mx-m0)/(sd/(sqrt(n)))
> zcal    [1] 8.333333
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0
```

Q.2
```
> P = 0.8
> Q = 1-P
> n = 1000
> p = 750/1000
```

```
>zcal=(p-P)/(sqrt(P*Q/n))
>zcal
[1] -3.952847
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 7.72268e-05
```

since value < 0.05 we reject H_0 at 5% level of significance

Q.3 Two random sample of size 1000 and 2000 are drawn from two population with same S.D 2.5. The sample means are 67.5 and 68 respectively. Test the hypothesis $H_0: \mu_1 = \mu_2$ os $H_1: \mu_1 \neq \mu_2$ at 5% level of significance

Q.4. A study of noise level in 2 hospital is given below. Test the claim that the hopitals have 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

Q.5. In a sample of 600 students in a college. 400 use blue ink in another college from a sample of 900 students 450 use blue ink. Test the hypothesis that the proportion of students using blue ink in two colleges are equal or not at 1% level of significance

Q.3
```
> n1= 1000
> n2= 2000
> mx1 = 67.5
> mx2= 68
> sd1= 2.5
> sd2= 2.5
> zcal = (mx1-mx2)/sqrt((sd1^2/n1)+(sd2^2/n2))
> zcal
[1] -5.163978
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1]2.417564e-07
```

Q4
```
> n1= 84
> n2= 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
> cat("calculated value of z is ", zcal)
calculated value of z is 1.162528
> pvalue = 2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.245021
```

Q.5

```
> n1=600
> n2=900
> p1=400/600
> p2=450/900
> p=(n1*p1+n2*p2)/(n1+n2)
> q=1-p
> p
[1] 0.5666667
> zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
> zcal
[1] 6.381534
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 1.753222e-10
```

Q.6 n1=200, n2=200, p1=44/200, p2=30/200

```
> n1=200
> n2=200
> p1=44/200
> p2=30/200
> p=(n1*p1+n2*p2)/(n1+n2)
> q=1-p
> zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
> pvalue=2*(1-pnorm(abs(zcal)))
> pvalue
[1] 0.07142888
> zcal
[1] 1.802741
```

2/3/20

∴ value > 0.05 mean H₀ is accepted

PRACTICAL-7

Aim : Small sample test

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

$H_0 : \mu = 66$

```
> t.test(x)

        One Sample t-test

data: x
t= 68.319, df= 9, p-value < 1.558e-13
alternative hypothesis : true mean is not equal to 0
95 percent confidence interval:
  65.65171  70.14829
sample estimates:
mean of x
     67.9
```

∵ value < 0.05 mean is rejected.

```
> pvalue = 1.558e-13
> if(pvalue > 0.05) (cat("accept no")) else (cat("
   reject no"))
reject no
```

RA

2. 2 groups of students score the following marks. Test the hypothesis that there is no significant difference between the two groups.

Group 1:- 18, 22, 21, 17, 20, 17, 23, 20, 22, 21

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

Ho : There is no difference betwen two groups.

```
>x=c(18,22,21,17,20,17,23,20,22,21)
>y=c(16,20,14,21,20,18,13,15,17,21)
>t.test(x,y)

        Welch Two Sample t-test

data: x and y
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
> if(0.03798>0.05)(cat("accept no")) else(cat(
reject no"))
reject no
```

Q. The sales data of six shops before and after a campaign is given below:
Before : 53, 28, 31, 48, 50, 42
After : 58, 29, 30, 55, 56, 45
Tes the hypothesis that the campaign is effective or not. Ho: There is no significant 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")
```

~~Welch Two~~
Paired t-test

data: x, y
t=-2.7815, df=5, p-value= 0.9806
alternative hypothesis: true difference is means is greater than 0
95 percent confidence interval:
-6.035547   Inf
sample estimates:
mean of the differences
-3.5

```
>if(0.9806 >0.05)(cat("accept no")) else(cat("reject no"))
```

~~rej~~ accept no.

4. Two medicines are applied to two groups of paitent respectively
Group1 - 10, 12, 13, 11, 14
Group2 - 8, 9, 12, 14, 15, 10, 9
Is there any significant difference between two medicines. Ho: There is no significant difference

```
>x = c(10, 12, 13, 11, 14)
>y = c(8, 9, 12, 14, 15, 10, 9)
>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
95 percent confidence interval:
 -1.781171        3.781171
sample estimates:
mean of x   mean of y
     12          11
```

```
>if (0.4406 > 0.05) (cat("accept no")) else(cat
 ("reject no"))
accept no.
```

The following are the weight before & after the diet program. Is the program effective
before : 120, 125, 115, 130, 123, 119
after : 100, 114, 95, 90, 115, 99
Ho: There is no significant difference

```
>x=c(120,125,115,130,123,119)
>y=c(100,114,95,90,115,99)
>t.test(x,y,paired=T, alternative="less")
```

Paired t-test

data: x and y
t=4.3458 , df=5, p-value=0.9963
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf 29.0295
sample estimates:
mean of the differences
19.83333

```
>if(0.9963>0.05)(cat("accept no"))(else(cat("reject no"))
accept no
```

AM
6.1.20

Answers

Q.1 Ho = 55

```
>mo= 55
>mx =52
>n = 100
>sd= 7
>zcal= (mx-mo)/(sd/(sqrt(n)))
>zcal
[1] -4.285714
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 1.82153e-05
```

∵ pvalue < 0.05 we reject Ho

Q.2 Ho = μ = 1/2

```
>n1 = 350
>n2= 700
>p= 350/700
>P=0.5
>Q = 1-P
>zcal= (p-P)/(sqrt(P*Q/n))
>zcal
[1] 0
>pvalue = 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 1
```

∵ pvalue > 0.05 we accept Ho

Q.3 Ho = P2

```
>n1=1000
>n2=1500
>p1= 20/1000
>p2= 15/1500
>p =(n1*p1+n2*p2)/(n1+n2)
>q=1-p
>p
[1] 0.014
>zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2)
>zcal
[1] 2.081283
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.03740801
```

∵ pvalue < 0.05 we reject Ho

Q.4 Ho = 100

```
> mx=99
> mo=100
>n= 400
> sd= 8
> zcal=(mx-mo)//(sd//(sqrt(n)))
>zcal
[1] -2.5
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.0124193
```

∵ pvalue < 0.05 we reject Ho

Q.5 ~~Ho:~~= Ho: M = 66

```
>x=c(63,63,68,69,71,71,72)
>t.test(x)

        One Sample t-test

data: x
t=47.94, df=6, p-value=5.522e-09
alternative hypothesis: true mean is not equal
to 0.
95 percent confidence interval:
64.66479    71.62092
sample estimates:
mean of x
 68.14286
```

∵ pvalue < 0.05 we reject Ho

Q.7 Ho = 1200

```
>m0=1200
>mx=1150
>sd=125
>n=100
>zcal=(mx-m0)/(sd/(sqrt(n)))
>zcal
>pvalue=2*(1-pnorm(abs(zcal)))
>pvalue
[1] 6.334248e-05
```

∵ pvalue < 0.05 we reject Ho

Q.6.

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

F test to compare two variances

data: x and y
F= 0.70686, num df = 8, denom df = 10, p-value= 0.6359
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.1833662 3.0360393
sample estimates:
ratio of variances
0.7068567

∵ pvalue > 0.05 we accept. Ho

Q.8 Ho = $P_1 \neq P_2$

```
>n1=200
>n2=300
>p1= 44/200
>p2= 56/300
>p(n1*p1+n2*p2)/(n1+n2)
>q=1-p
>p
[1]0.2
>zcal=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
>zcal
```

```
[1] 0.9128709
>pvalue= 2*(1-pnorm(abs(zcal)))
>pvalue
[1] 0.3613104
```

∴ pvalue > 0.05 we accept Ho

AM
13-2-20

AS

PRACTICAL - 9

Topic: Chi square texts & ANOVA

Q.1 Use the following data to test whether the condition of the home and the condition of the child is independent or not.

Condition of child	Condition of home: Clean	Dirty
Clean	70	50
Fairly clean	80	20
Dirty	35	45

Ho : condition of the home and child is 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
```

Person's Chi-squared test

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

∵ pv < 0.05 we reject Ho

Q2 Test hypothesis that vaccination and diseases are independent or not

	Vaccine	
	Affective	Not Affective
Affective	70	46
Not Affective	35	37

Ho : Disease and the 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-squared = 2.0275, df = 1 ,p-value = 0.1545

∵ pv > 0.05 we accept Ho

Q3 Perform a ANOVA for the following data

Type	Observation			
A	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"
```

```
> 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 = 0.00183

∴ pv < 0.05 we reject H_0

```
> 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	0.00183
Residuals	9	18.17	2.019		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Q4 The following data gives the life of the tyre of four ~~gm~~ brand

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 : The average life of tyres 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~ind, data=d, var.equal=T)
```

Oneway analysis of means

data: values and ind

F = 6.8445, num df=3, denom df=20,

p-value= 0.002849

∵ pv < 0.05 we reject Ho

Q.5.

```
>x=read.csv("C:/Users/Administrator/
        Desktop/marks.csv")
>print(x)

>am=mean(x$stats)
>am
[1] 37
>me=median(x$stats)
>me
[1] 38.5
```

```
>n= length (x $ stats)
>v= (n-1)* var (x $stats)/n
>v
[1] 160
> sd= sqrt (v)
>sd
[1] 12.64911
> am1= mean(x $maths)
>am1
[1] 39.4
>me1= median (x $maths)
>me1
[1] 37
>n1= length (x $ maths)
>v1=(n1-1)*var (x $maths)/n1
>v1
[1] 231.04
>sd1= sqrt (v1)
>sd1
[1] 15.2
> cor (x $ stats, x $ maths)
[1] 0.830618
```

AM
20.02

PRACTICAL-10

Topic: Non-parametric test

Q.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% los

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

Ho: Population median is 21.5

```
> x=c(.....)
> 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
```

∵ pv > 0.05 we ~~reject~~ accept Ho

Note: If the alternative (H_1) is ~~H~~ me < than pv=pbinom(sp,n,0.5)

If the alternative is me> than pv=pbinom(sn,n,0.5)

Q2 Following is the data of 10 observation. Apply sign test. Test the hypothesis that the population median is 625 against the alternative. it is more than 625

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

Ho: Population median is 625

```
>x=c(...)
>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
```

∴ pv > 0.05 we accept Ho

Q3 Following are the values of a sample. Test the hypothesis that the population median is 60 against the alternative it is more than 60 at 5% LOS 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(....)
> wilcox.test(x, alter="greater", mu=60)
```

Wilcoxon signed rank test with continuity correction

data: x

V = 145, p-value = 0.02298

alternative hypothesis: true location is greater than 60

∵ pv < 0.05 we reject Ho

Note: If the alternative is less than alter="less"

If the alternative is equal to than altee="two sided"

Q.4 Using WSR test. Test the population median is 12 or less than 12

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

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

Wilcoxon signed Rank test with continuity correction

data: x

V = 66, p-value = 0.9986

alternative hypothesis: true location is less than 12

∵ pv > 0.05 we accept Ho

The weights of students before or after they stop smoking are given below using WRT Test that there is no significant change

Before : 65, 75, 75, 62, 72

After : 72, 74, 72, 66, 73

Ho: Before and after there is no change

H1 : There is change

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

wilcoxon signed Rank test with continuity correction

data: d

V=4.5 p-value=0.4982

alternative hypothesis: true location is not equal to 0

∵ pV>0.05 we accept Ho

27.02.