Basic of R software:

1) R is a software for data analysis and statistical computing.

2) This software is used for effective data handling and output storage is possible.

3) It is possible of graphical display.

4) It is a free software.

Ques

1). $2^2 + \sqrt{25} + 35$

Ans:- 2^2 + sqrt(25) + 35

[1] 44.

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

Ans:- 2*5*3 + 62/5 + sqrt(49)

[1] 49.4

3). $\sqrt{76 + 4 \times 2 + 9 \div 5}$

Ans:- sqrt(76 + 4*2 + 9/5)

[1] 128.

4) $42 + |-10| + 7^2 + 9 \times \cancel{5}$

Ans 42 + abs(-10) + 7**2 + 3*9

[1] 128

18

5) $x=20$ find x^2+y^2, $\sqrt{y^3-x^3}$, $x+y$, $|x-y|$

$y=30$

x^2+y^2

x^2 + y^2

[1] 1300

$x+y$

x + y

x+y = [1] 50

$|x-y|$

x-y

[1] 10

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

sqrt(y^3 - x^3)

[1] 137.84

6) c[2, 3, 4, 5] ^2

c[2.3.4.5] ^2

4, 6, 9, 25

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

[1] 12. 15. 18. 24

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

[1] -4, -9, -25, -28

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

[1] 16 27 40 63

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

[1] 1 8 9 64 25 216

Find the sum product maximum, minimum of the values

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

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

length(x)
[1] 8

sum(x)
[1] 65

prod(x)
[1] 11340000

max(x)
[1] 15

min(x)
[1] 5

Q => for matrix. $\begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$

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

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

Q -> $x = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ $y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ 10 & 6 & 12 \end{bmatrix}$ find x+y, x*y, 2x+3y.

Sol^n :- 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)).

x+y

	[,1]	[,2]	[,3]
[1,]	3	8	17
[2,]	0	13	-3
[3,]	13	12	21

x*y

	[,1]	[,2]	[,3]
[1,]	2	16	70
[2,]	-4	40	-88
[3,]	30	36	108

2*x+3*y

8	20	44
-2	34	-17
36	30	54

6) > x = c(2, 4, 6, 1, 3, 5, 7, 13, 16, 14, 17, 19, 3, 3, 2, 5, 0, 15, 9, 14, 18, 10, 12)

```
length(x)
[1] 23
a = table(x)
transform(a).
```

x	Freq
0	1
1	1
2	2
3	3
4	1
5	2
6	1
7	1
9	1
10	1
12	1
14	2
15	1
16	1
17	1
18	2
19	1

```
> breaks = seq(0, 20, 5)
> b = cut(x, breaks, right = F)
> c = table(b).
transform(c).
```

	b	Freq
1	[0,5)	8
2	[5,10)	5
3	[10,15)	4
4	[15,20)	6

4-12-19

Practical - 2

Problem on P.d.f and C.d.f

1) Can the followings be c.d.f;

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

(iii) $f(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & o.w \end{cases}$ (iv) $f(x) = \begin{cases} \frac{3x}{2}\left(1-\frac{x}{2}\right) & 0 \le x \le 2 \\ 0 & o.w \end{cases}$

(1) soln:- $\int f(x)\,dx = 1$ is condition for pdf

(1) $\int_1^2 (2-x)\,dx \Rightarrow \int_1^2 2\,dx - \int_1^2 x\,dx \Rightarrow [2x]_1^2 - \left[\frac{x^2}{2}\right]_1^2$

$(4-2)-(2-0.5) \neq 1$ so not P.d.f

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

soln:- $\int_0^1 3x^2\,dx = \left[\frac{\cancel{3}x^3}{\cancel{3}}\right]_0^1 = \cancel{\left[\frac{1}{3} - 0\right] \neq} \quad x^3\Big|_0^1 = 1$

so ~~not~~ P.d.f

(3) $f(x) = \begin{cases} \frac{3x}{2}\left(1-\frac{x}{2}\right) & 0 \le x \le 2 \\ 0 & o.w \end{cases}$

soln:- $f(x) = \int_0^2 \frac{3x}{2} - \frac{3x^2}{4}\,dx \Rightarrow \int_0^2 \frac{3x}{2}\,dx - \int_0^2 \frac{3x^2}{4}\,dx$

$\left[\frac{3x^2}{4}\right]_0^2 - \left[\frac{x^3}{4}\right]_0^2 \Rightarrow \left[\frac{3 \times \cancel{4}}{\cancel{4}} - 0\right] - [2] =$

$3 - 2 = 1$ so it p.d.f

2 Can the following be p.m.f

①

x	1	2	3	4	5
$p(x)$	0.2	0.3	-0.1	0.5	0.1

②

x	0	1	2	3	4	5
$p(x)$	0.1	0.3	0.2	0.2	0.1	0.1

③

x	10	20	30	40	50
$p(x)$	0.2	0.3	0.3	0.2	0.2

Solⁿ-1→ ~~Hence~~ since 1 valid is -vely so it is not p.m.f.

Solⁿ 2→ Since $p(x) \geq 0 \;\forall x$
and $\Sigma p(x) = 1$ it is p.m.f.

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

3→
```
prob1 = c(0.2, 0.3, 0.3, 0.2, 0.2)
sum(prob1)
[1] = 1.2
```
so it is not p.m.f.

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

31

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

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

$P(2 \le x \le 4) = P(2) + P(3) + \cancel{P(4)} = 0.2 + 0.2 \cancel{+ 0.1} = 0.4$

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

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

Q.4 find c.d.f

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

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.

$$
f(x) = \begin{cases} 0 & \text{if } x < 0 \\ .1 & \text{if } 0 \le x < 1 \\ .2 & \text{if } 1 \le x < 2 \\ .4 & \text{if } 2 \le x < 3 \\ .6 & \text{if } 3 \le x < 4 \\ .7 & \text{if } 4 \le x < 5 \\ .9 & \text{if } 5 \le x < 6 \\ 1 & \text{if } 6 < x \end{cases}
$$

Scanned with CamScanner

(2)

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

$$
F(x) = \begin{cases} 0 & \text{if } x < 0 \\ .2 & \text{if } 0 \le x < 12 \\ .55 & \text{if } 12 \le x < 14 \\ .7 & \text{if } 14 \le x < 16 \\ .9 & \text{if } 16 \le x < 18 \\ 1 & \text{if } 18 < x \end{cases}
$$

AW
4.12.19

Practical - 3

Probability distribution and binomial distribution

i) Find the c.d.f of the following p.m.f and draw the graph

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

Ans :-

```
x = c(10,20,30,40,50)
prob = c(0.15,0.25,0.3,0.2,0.1)
cumsum(prod)
[1] 0.15 0.40 0.7 0.9 1.00
```

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

$0.15 \quad 10 \leq x < 20$

$0.40 \quad 20 \leq x < 30$

$0.70 \quad 30 \leq x < 40$

$0.90 \quad 40 \leq x < 50$

$1 \quad x \geq 50$

```
plot (x, prob, xlab="values", ylab="Probability", main="graph of pmf
    "s").
```

Binomial distribution

Suppose there are 12 MCQ in a test. Each Question has 5 option only one of them correct. Find the probability of having (1) 5 correct answer (2) atmost 4 correct answer

Sol^n

It is given that n = 12, p = 1/5, q = 4/5

X = total no of correct answer

X ~ B(n, p)

n = 12

p = 1/5

> dbinom(5, 12, 1/5)

[1] 0.05315022

(2) atmost 4

pbinom(4, 12, 1/5)

[1] 0.9274445

Q.3) There are 10 member in commitee the probability of any member attening a commeeting 0.9 find probability (1) 7 member attened (2) atleast 5 member attened (3) atmost 6 member attened

Sol^n

It is given n=12, p=.9, q=.1 x=7

7 member
> dbinom(7,10,0.9)

[1] 0.05739563

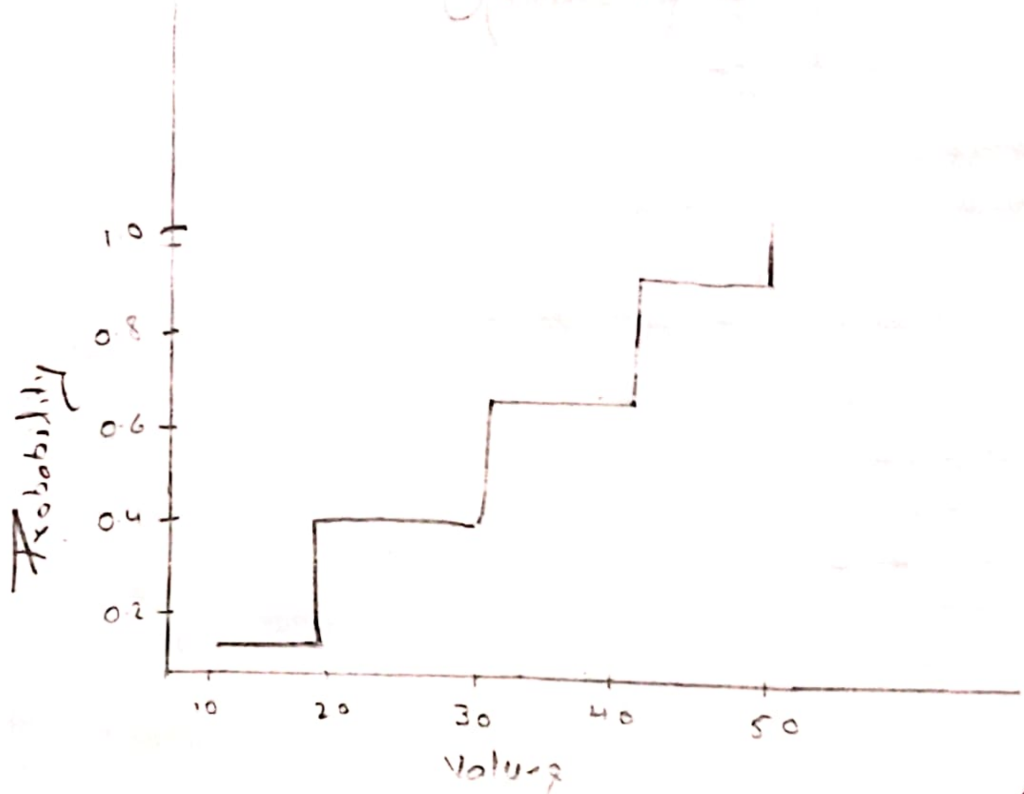
(2) atleast 5 member

1-pbinom(4,10,0.9)

~~[1] 0.9983651~~

[1] 0.9998531

①



③ atmost 6 member

pbinom (6, 10, 0.9)

[1] 0.0127952

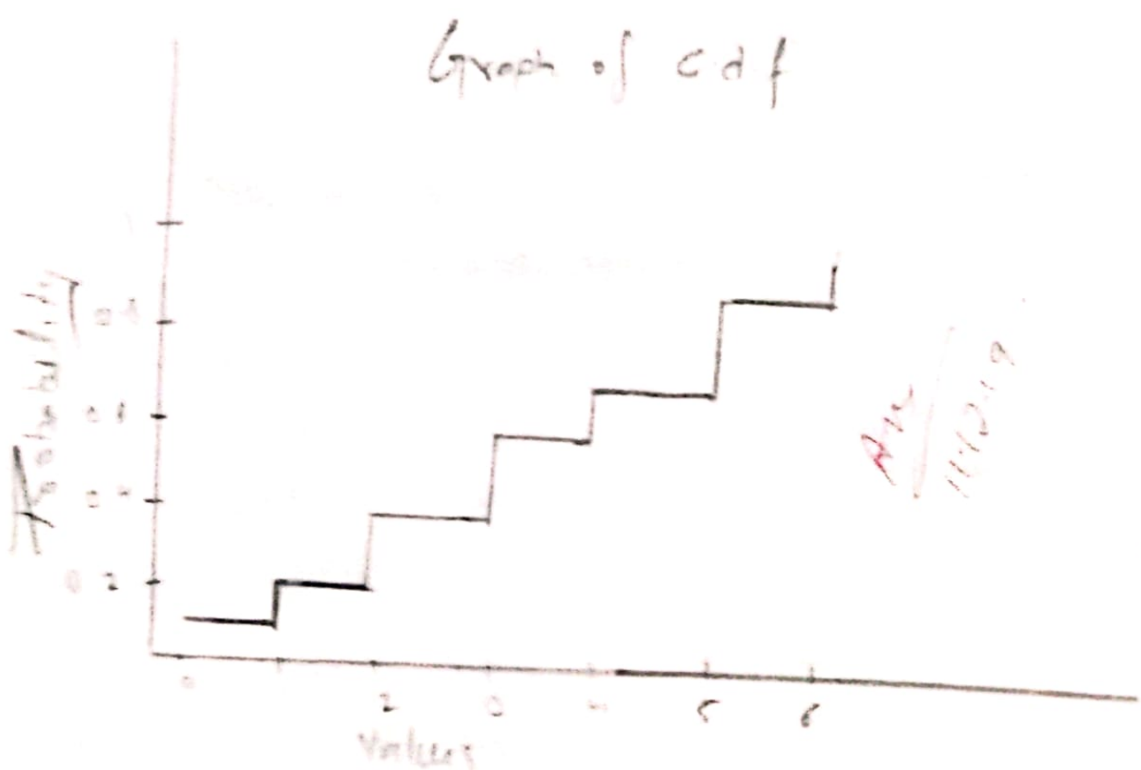
Q.4 Find the C.D.F and Draw the graph

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

```
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 1.00
```

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.1 & 0 \le x < 1 \\ 0.2 & 1 \le x < 2 \\ 0.4 & 2 \le x < 3 \\ 0.6 & 3 \le x < 4 \\ 0.7 & 4 \le x < 5 \\ 0.9 & 5 \le x < 6 \\ 1 & 6 \le x \end{cases}$$

```
plot(x, prob, xlab="Values", ylab="Probability", main="graph of
cdf", "s").
```



Practical-4

Binomial Distribution

Find the complete binomial distrubution when $n=5$ and $p=0.1$

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

X follows binomial distribution with $n=12$, $p=0.25$, find

$P(x=5)$

$P(x\leq5)$

$P(x>7)$

$P(5<x<7)$

The probability of salesman make a sale to a customer is 0.15

Find the probability

① No sale for 10 customer, more than 3 sale in 20 customer.

A student write 5 MCQs, each question has 4 options out of which 1 correct calculate the probability atleast 3 correct answer.

Q.1

Note :- $P(X=x)$ = dbinom (x, n, p)

$P(X \le x)$ = pbinom (x, n, p)

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

To find the value of x for which the probability given is p_1 command is qbinom (p_1, n, p)

1) n=5 p=0.1

dbinom (0:5, 5, 0.1)

[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001

2) dbinom (10, 100, 0.1) n=100

[1] 0.1318653

3) n=12 p=0.25

(I) dbinom (5, 12, 0.25) $P(X=5)$

[1] 0.1032414

(II) $P(X \le 5)$

pbinom (5, 12, 0.25)

[1] 0.9455

(III) $P(X > 7) = 1 - P(X \le 7)$ = 1- pbinom (7, 12, 0.25)

[1] 0.0027815 1

(IV) $P(5 < X < 7)$

dbinom (6, 12, 0.25)

0.04014945

4) (I) $n = 10$ $p = 0.15$ $x = 0$

dbinom (0,10,0.15)

[1] 0.1968744.

(II) $n = 20$ $p = 0.15$

$P(x > 3) = 1 - P(x \leq 3) = 1 -$ pbinom (3,20,0.15)

[1] 0.3522748

5) $n = 5$, $p = 0.25$, $x = 3$

$P(x \geq 3)$

$1 - P(x \leq 2)$

1 - pbinom (2,5,0.25)

[1] 0.1035156

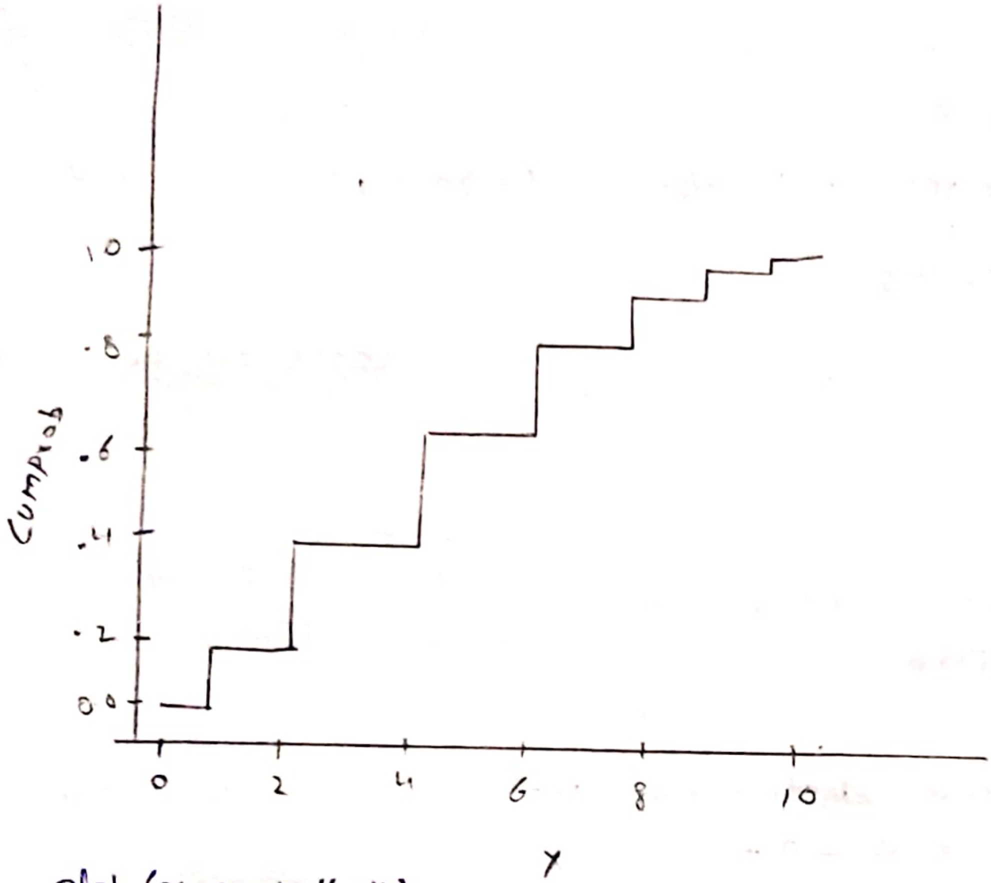
Q.6 X follows binomial distribution $n = 10$, $P = 0.4$ plot the graph of p.d.f and c.d.f

Soln $n = 10$, $p = 0.4$ $x = 0:n$

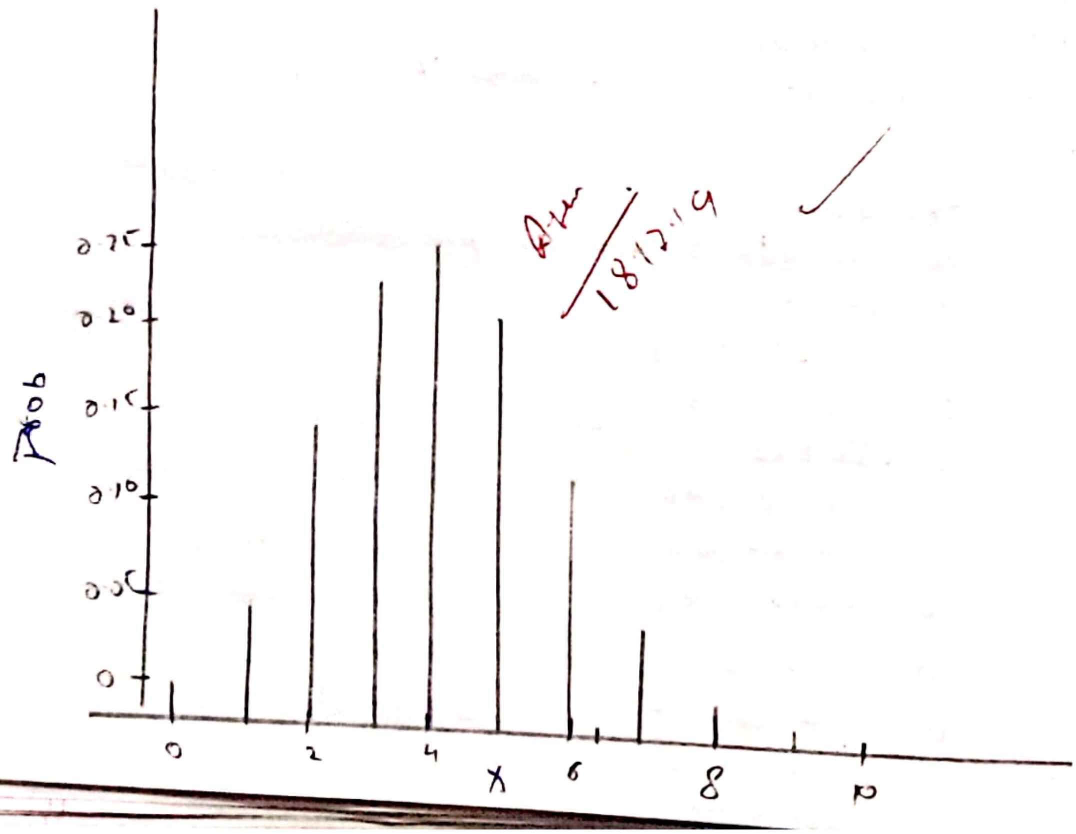
```
prob= dbinom(x,n,p)
cumprob = pbinom(x,n,p)
d =dataframe("xvalues"=x, "probability"=prob)
print(d)
```

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

plot (x, cumprob, "s")



plot (x, prob, "h")



18.12.19

Practical-5

8/01/2020

Normal Distribution

1) $P[X=x]$ = dnorm(x, μ, σ)

2) $P(X \le x)$ = pnorm (x, μ, σ)

3) $P(X > x)$ = 1- pnorm (x, μ, σ)

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

To find the value of k as shown that

5) $P(X \le k) = p_1$ qnorm (p, μ, σ)

To generate ~~n~~ random numbers command : x

rnorm(n, μ, σ)

Q.1 $x \sim N(\mu = 50, \sigma^2 = 100)$. find.

① $P(x \le 40)$ ② $P(x > 55)$ ③ $P(42 \le x \le 60)$

④ $P(x \le k) = 0.7$ $k = ?$

Q.2 ⇒ $x \sim N(\mu = 100, \sigma^2 = 36)$

① $P(x \le 110)$ ② $P(x \le 95)$ ③ $P(x > 115)$

④ $P(95 \le x \le 105)$ ⑤ $P(x \le k) = 0.4$ $k = ?$

Q.3 Generate 10 random numbers from a Normal distribution with mean = 60 S.D = 5. Also calculate sample mean, median, ~~S.D~~. Variance, S.D

⑤ Draw the graph of standard normal distr
when

Ans

1)
```
a = pnorm(40,50,10)
cat("P(x<=40)=", a)
```
P(x<=40 = 0.1586553.

②
```
b = 1-pnorm(55,50,10)
cat("p(x>55) = ",b)
```
p(x>55) = 0.3085375

③
```
c = pnorm(60,50,10) - pnorm(42,50,10)
cat("p(42 <= x <= 60) = ",c)
```
p(42 ≤ x ≤ 60) = 0.6294893

④
```
d = qnorm(0.7,50,10)
cat("p(x<=k) = 0.7, K is = ", d)
```
p(x ≤ k) = 0.7   K is = 55.24401

(7). (1) a = pnorm (110, 100, 6)
cat ("p(x ≤ 110) = ", a)
p(x ≤ 110 = 0.9522096.

(2) b = pnorm (95, 100, 6)
cat ("p(x<=95) = ", b)
p(x<=95) = 0.2023284.

(3) c = 1 - pnorm (115, 100, 6)
cat ("p(x > 115) = ", c)
p(x >115) = 0.00620966r

(4) d = pnorm (105, 100, 6) - pnorm (95, 100, 6)
cat ("p(95 <= x <= 105) = ", b)
p(95 <= x <= 105) = 0.5953432.

(5) e = qnorm (0.4, 100, 6)
cat ("p(x <= k] = 0.4 k = k = 98.47002.

(3). a = rnorm (10, 60, 5).   n = 10   μ = 60   σ = 5
[1] 58.45543 55.02759 60.59413 58.66145 60.02250
60.54647 61.27573 71.97780 55.01164 68.01424

mean(a)
[1] 61.18175

median (a)
[1] 60.27534

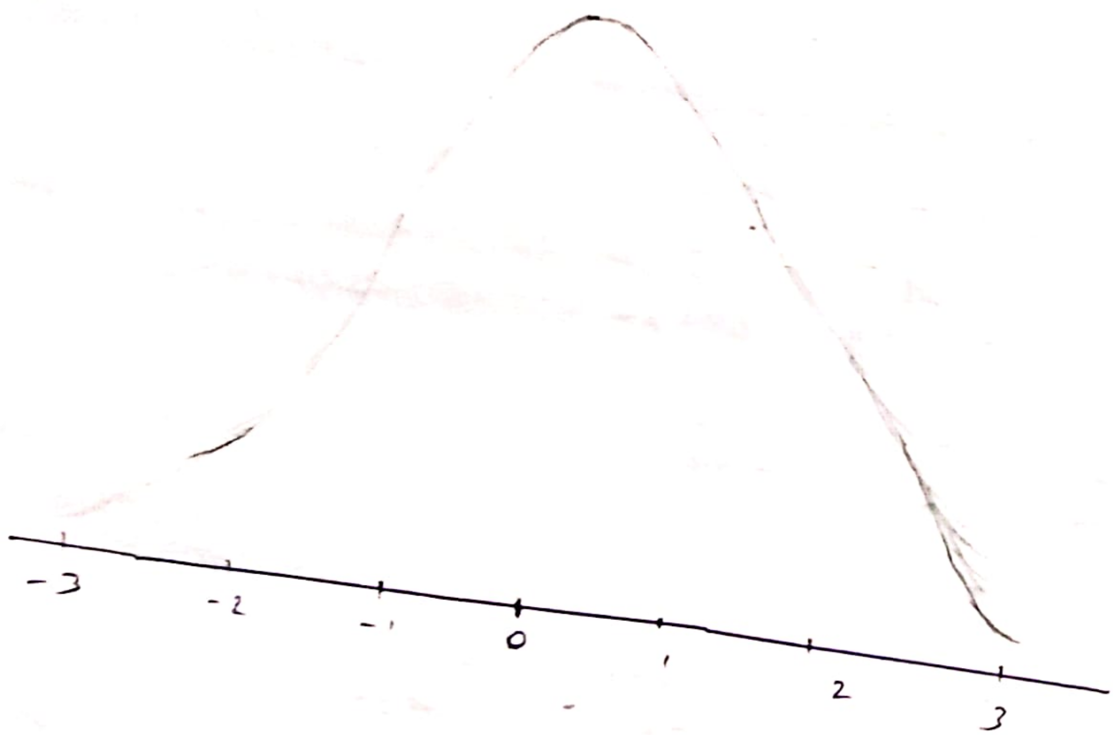
sd (a)
[1] 5.080309

σ² (sd(a)^2)
[1] 25.80954

n = 10
variance = (n-1) * (x)/n
[1] 23.22858.

SD = sqrt(varianc)
[1] ~~5.080309~~ 4.819604

4). x = seq(-3, 3, by = 0.1)
y = dnorm(x)
plot(x, y, xlab = "x value", ylab = "probability", main = "standard normal distribution", "s").



Q.
Sol":
Q.-7
Sol":

Z-Distribution.

Test for Hypothesis.

$H_0 : \mu = 10$ against $H_1 ; \mu = 10$

A sample of size 400 is collected which gives a mean 10.2 and the standard deviation 2.25

Test the hypothesis at 5% level of significant

Ans:- m_0 = (mean of population) = 10
m_x = (mean of sample) = 10.2
sd = 2.25
n. (sample size) = 400
zcal = (mx-m0) /(sd / sqrt (n)))
cat= ("z cal is: ", zcal)
>> zcal is 1.7778
pvalue = 2*(1-pnorm (abs (zcal)))
pvalue
>> 0.075

Note: If result of tested Hypothesis is > 0.05 then assumed of $H_0 : \mu_0 = 10$ is accept.

⇒ Test the hypothesis $H_0 : \mu = 75$ against $H : \mu \neq 75$

A sample of size 100 is selected & the sample mean is 80 with the S.D = 3.

test of hypothesis at 5% level of significant

Ans:- mo = (mean of population) = 75
mx = (mean of sample) = 80
sd = 3
n = (sample size = 100
zcal = (mx-mo) /(sd / (sqrt(n)))

IA1

cat: ("zcal is : ", zcal)

>>> z cal is = 16.6667

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

pvalue

>>0.

Q) Test of Hypothesis $H_0: \mu = 25$ against $H_1: \mu \neq 25$ at 5% level of significance.

The following sample of 30 selected.

```
x=c(20,24,27,35,30,40,26,27,10,20,30,37,35,21,22,23,
24,25,26,27,28,29,30,39,27,15,19,22,20,18)
>mx= mean(x)
 mx
 [1] 26.0667
>n=length(x)
>n
[1]30
variance = (n-1)* var(x)/n
variance
 52.9956
>sd = sqrt(variance)
 sd
[1]   7.279805
   mo=25
   mx= 26.0667
   n=30
   sd=7.279805
> zcal = (mx-mo)/(sd/sqrt(n)))
 zcal
   0.8025454
```

pvalue = 2* (1-pnorm (abs (zcal)))
pvalue
0.422375

The assumed Hypothesis is verified

Q2) Experience has show that 20% student of a college are smoke a sample of four hundred student reveal that out of 400 only 50 smoke test the hypothesis that the experience give the correct proportion or not

> P = 0.2
Q = 1-P
p = 50/400
p
0.125
n = 400
zcal = (p-P)/(sqrt (P*Q/n))
zcal
-3.75

pvalue = 2* (1-pnorm(abs(zcal)))
pvalue
[1] 0.0001768346

Q→ Test the hypothesis $H_0: p = 0.5$ against $H_1: p \neq 0.5$. A sample of 200 is selected and the sample portion is cal p = 0.56 test the hyposesis 1% level of significant

```
n = 200
p = 0.56
P = 0.5
Q = 1-P
zcal = (p-P)/(sqrt(P*Q/n))
cat("zcal is", zcal)
 zcal is 1.697056
pvalue = 2*(1-pnorm(abs(zcal)))
 pvalue
[1] 0.08968602
```

Pratical = 7

large small test.

1) A study of Noise level on 2 hospital is calculate below test of hypotesis that of Noise ... on 2 hospital are same or not.

	Hos A	Hos B
Nos of Sample	84	34
mean	61	54
SD	7	8

Sol^n:- The Noise level are same

$n_1 = 84$

$n_2 = 34$

$mx = 61$

$my = 54$

$sdx = 7$

$sdy = 8$

z = (mx-my) / sqrt((sdx^2/n1) + (sdy^2/n2))

z = 1.273662

cat("z calculated: ", z)

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

? pvalue = 0.2076.

2) Random sample of size 1000 & 2000 are drawn from 2 population with a means 67.5 and 68 respectively with a same S.D of 2.5 Test the hypothesis that the mean of the population are equal.

Sol^n:- H_0: Two population mean are equal.

$n_1 = 1000$

$n_2 = 2000$

$mx = 67.5$

$my = 68$

$sdx = 2.5$

$sdy = 2.5$

```
>z=(mx-my)/sqrt((sdx^2/n1)-(sdy^2/n2))
>z
[1] -5.163478
>cat("Z calculated is=",z)
Z calculated is= 5.1639787
>pvalue = 2*(1-pnorm(abs(z)))
pvalue
[1] 2.417564e-07
```

Q=) In a first year class, 20% of a random sample of 400 students had defective eye sight. In 2nd class 15.5% of 500 sample student had the same defect is the difference of population is same?

Soln

Ho: The proportion of 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
[1] 0.825
>z = (p1-p2)/sqrt(p*q*(1/n1+1/n2))
> z
[1] 1.76547
```

```
>cat("z calculated is = 1.765467)
>pvalue = 2*(1-pnorm(abs(z)))
>pvalue
[1] 0.07748489
```

Q=> From each of the box of the apples a sample size of 200 is collected. It is found that there are 44 bad apples in the first sample & 30 bad apples in the second sample. Test the hypothesis that the two boxes are equivalent in terms of no of bad apples.

Q=> In a M.A Class out of the sample of 60, mean height is 63.5 inch with a S.D of 2.5. In a M.com class out of 50 student mean height is 69.5 inch with a S.D of 2.5. Test of hypothesis that the mean of M.A and M.com class are same.

Solⁿ:-

H_0: The two boxes are equivalent in terms of bad apples

```
>n1=200
>n2=200
>p1=44/200
>p2=30/200
>p= (n1*p1+n2*p2)/(n1+n2)
>p
[1] 0.185
>q=1-p
>q
[1] 0.815
>z=(p1-p2)/sqrt(p*q*(1/n1+1/n2))
>z
[1] 1.802741
>cat("z calculated is =1.802741")
>pvalue = 2*(1-pnorm(abs(z)))
>pvalue
[1] 0.07142888
```

pvalue > 0.05 we accept the H_0 at 5% level of significance

Sol^n 5 =)

H_0 : The mean of M.A & M.com are same

$n_1 = 60$

$n_2 = 50$

$m_x = 63.5$

$m_y = 69.5$

$sd_x = 2.5$

$sd_y = 2.5$

```
> z = (mx-my)/sqrt((sdx^2/n1)+(sdy^2/n2))
> z
[1] 12.5359
> cat("z calculated is =", z)
z calculated is = -12.53359
> pvalue = 2*(1-pnorm(abs(z)))
> pvalue
[1] 0
```

As pvalue < 0.05, we reject the H_0 at 5%

Scanned with CamScanner

Practical - 8

Samall sample text

1) The flowers 10 selected and the haght founded 63, 63, 68, 69, 71, 71, 72 cm. Test the Hypothesis that the mean height 66 cm or not at 1%

A:- Mean = 66 cms

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 = 5.822 e-09.

alternative hypothesis: True mean is not equal to 0

95 percent confidence interval:

64.66479 71.62092

Sample estimate:

mean of x.

68.14286

pvalue less than 0.01 we reject Ho of 1% level of signifiet.

2 Random sample drawn from 2 different population Sample:- 8, 10, 12, 11, 16, 15, 18, 7

Sample 2:- 20, 15, 18, 9, 8, 10, 11, 12. Test the hypothesis that there is no difference mean b/w the polpola mean at 5%

72

Ho: There is no difference in population mean

x=c(8,10,12,11,16,15,18,7)
y=c(20,15,18,9,10,11,12,8)
t-test(x,y).

Welch Two Sample t-test.

data: x and y

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

alternative hypothesis: True difference in means not equal to 0

95 percent confidence interval:

-5.192719 3.692719

Sample estimates.

mean of x and y

12.125 12.875

p-value greater than 0.05 we accept Ho at 5% level of significant.

Q. Following are the weight 10 people before and after diet program. Test hypothesis that it effe or not.

Before 100, 125, 95, 96, 98, 112, 115, 104, 109, 110

After 95, 80, 95, 98, 90, 100, 110, 85, 100, 101

Soln: Ho The diet program is not effective

x=c(100,125,95,96,98,112,115,104,109,110)

y=c(95,80,95,98,90,100,110,85,100,101)

t-test (x,y, paried=T, alternative = "Less")

Welch Two sample t-test.

data : x and y

t = 2.6991 df = 17.801, p-value = 0.985

alternative hypothesis: True difference in means is not ~~equal~~ less than to 0.

95 percent confidence interval:

-Inf 18.72908

Sample estimates :

mean of the difference

''

p-value greater than 0.05 we accepting H_0 at 5% level of significant.

Mark before and after a training program as given below.

before: 20, 25, 32, 28, 27, 36, 35, 25

After: 30, 35, 32, 37, 37, 40, 40, 23

Test the Hypothesis is effective or not.

Sol:- H_0: The trainning program not effective

x=c(20,25,32,28,27,36,35,25)

y=c(30,35,32,37,37,40,40,23)

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

Paired t-test

data : x and y

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

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

-8.967399 Inf

Sample estimates:

mean of the difference

-5.75

p-value greater than 0.05. We accept H_0 at 5% level of significant.

Q.) 2 Random Sample were drawn from 2 random population and the value 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 have same varience at 5% of significant.

Sol

H_0: The varience of 2 population are equal.

```
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

p-value greater than 0.05 We accept H_0 at 5% level of significant

The arithemite mean of sample of 100 observation is 52 if S.D=7 test the hypothesis that the population mean =55 or not at 5% of LOS

Ho: population mean = 55

n = 100

mx = 52

mo = 55

SD = 7

z = (mx-mo) /(sd / sqrt(n))

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

pvalue = 4.

1.82153 e-05

p-value is < 0.05 we reject the 5% level of significant

M/12/2

Practical-9

Chi Square Distribution and Anova

Q-> Use the following data to test whether the cleanliness of new home is dependent upon other cleanliness condition or not

		O.H	home
cleanliness of child	clean	20	50
	fairly clean	80	20
	Dirty	35	45

H0: condition of home and the child are independent

```
x = c(70,80,35,20,45)
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)
data = y
x-squared = 25.646  df=2 , p-value = 2.6989-06
```

Hence pvalue is less than 0.5, we reject H0 at 5% LOS.

Q. Table below shows the relation performance b/w mathematics and computer

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

H_0: Performance in math and computer

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

> pvalue = chisq.test(y)
> pvalue
        pearson's chisquared test

data: y
  X-squared = 145.78,   df = 4,   pvalue < 2.2e-16
```

∴ Since pvalue less than 0.05 we reject H_0 at 5% los

Q=>

Q=>

∴ clear
of cl

Perform Anova for the following data.

Varities	observations.
A	50, 52
B	53, 55, 53
C	60, 58, 57, 56
D	52, 54, 54, 55)

H_0 = The means of variety of 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, val.equal = T)
        one-way analysis of means.
data: values and ind
F = 12.779 , num df = 3 , denom df = 10 , pvalue = 0.007

> anova = aov(value ~ ind, data = d)
> summary(anova)
```

	Df	sum sq	mean sq	F value	Pr(>F)
ind	3	[illegible]	30.988	12.78	0.00[illegible]
Residuals	10	[illegible].25	24.42[illegible]		

Hence pvalue is less than 0.05, we reject at 5% LoS.

Perform Anova

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

Solⁿ:

```
> x1 = c(6,7,5)
> x2 = c(4,6,5)
> x3 = c(5,6,10)
> x4 = c(6,7,9)
> d = stack(list(b1=x1, b2=x2, b3=x3, b4=x4))
names(d)
> oneway.test(values ~ ind, data=d, var.equal=T)
```

oneway analysis of means

data: values and ind
F = [illegible], num df = 3, denom df = 8, pvalue = [illegible]

Summary (anova).

	Df	sumsq	meansq	Fvalue	Pr(>F)
d		18	6.00	2.667	0.119
Residuals	8	18	2.25		

Type the data in Excel and save as in desktop filename.csv (MS.DOS)
Then open r-software and type x=read.csv(" ")
minimise r-software then right click on the file saved an desktop, click on properties, copy the location and paste it on.

x = read.csv(' (location)')

Replace \ this with / then after location add filename.csv)

Enter x
The data within excel sheet will be seen on R-software.

Practical - 10 => Permatrix test.

Following are amount of supharoxide emitted

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

apply sing test to test hypothesis that population median 29.5 against the alternatey it is sess than 21.5

Ho: Population median 21.5
H1: It is less than 21.5

x = c(17, 15, 20, 29, 19, 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, 17, 0.5)
pv.
[1] 0.4119091

We aceept the Ho.

If the alternatie is greater than median
pv = pbinom (sn, n, 0.5)

For the observation • 12, 19, 31, 28, 43, 40, 55, 49, 70 63. apply sing test to test population median is 25 against the alternative it is more than 25.

Ho: population median 25
H1: It is more than 25
x = c(12, 19, 31, 28, 43, 40, 55, 49, 70, 63)
m = 25

sp = length (x[x>m])
sn = length (x[x<m])
n = sp+sn
pv = pbinom (sp, n, 0.5)
pv
[1]
0.5981903
We reject the Ho.
If the alternative less than median
pv = pbinom (sn, sn, 0.5).

Q7 For following data
60, 65, 63, 89, 61, 71, 58, 51, 48, 66. test the hypothesis using welcoxon signed rank test. for testing hypothesis median is 60. against the alternative it is greater than 60.

Sol^n -

Ho : median is 60.
H1 : it is greater than 60

x = c(60,65,63,89,61,71,58,51,48,66)
wilcox.test (x, alter = "greater", mu = 60)

data: x
V = 29 , pvalue = 0.2346
alternative hypothesis : True location is greater than

Note

In alternative is less than

wilcox.test (x, alter= "less", mu=)

If alternative is not equal.

wilcox.test (x, alter="2.sided", mu=).

54

Q) Using wilcox.test test the hypothesis median is 12 and against the alternative of less than 12

12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20.

Sol^n H_0: median is 12

H_1: less than 12

x = c(12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20).

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

Wilcoxon signed rank test with continuity correction

data: x

V = 20.5, p-value = 0.1426

alternative hypothesis: True location is less than 12.

M
4.3.20