Sem II

Practical No. 1

P41

BASICS OF R-SOFTWARE

1. R is a software for statistical analysis & data computing.
2. It is an effective data handling software & outcomes storage is possible.
3. It is capable of graphical display.
4. It is a free software.

Q.1] Solve the following:

1) >4 + 6 + 8÷ 2-5
→ [1] 7.333333

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

3) > 5^3 + 7×5 × 8 + 46/5.
→ [1] ~~3~~414.2.

4) $\sqrt{4^2 + 5 \times 3 + 7/6}$
→ [1] 5.671567

5) round off (46÷7 + 9×8).
→ [1] 79.

Q.2] Solve the following:

1) a = c(2,3,5,7)*c(2)
[1] 4 6 10 14

2) b = c(2,3,5,7)*c(2,3,~~6~~)
b
[1] 4 9 10 21

3) c = c(2,3,5,7)* c(2,3,6,4)
c
[1] ~~Warning message.~~ 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.

Q.3] $x = 20$ $y = 30$ $z = 2$.

P42

find i) $x^2 + y^3 + z$.

ii) $\sqrt{x^2 + y}$

iii) $x^2 + y^2$.

→ ~~So~~ i) > $x^2 + y^3 + 2$.

[1] 27402.

ii) $\sqrt{x^2 + y^2}$

[1] 36.05551

iii) c = x^2 + y^2.

c

[1] 1300

4] find the matrix

> x = matrix (nrow = 4, ncol = 2, data = c(1, 2, 3, 4, 5

> x

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

5) find x+y and 2x+3y when ~~x~~

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

→ x = matrix (nrow = 3, ncol = 3, data = c(4, 7, 9, -2, 0, 6, 7, 3)).

x

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

Y = matric (nrow = 3, ncol = 3, data = c(10, 12, 15, -5, -6, 7, 9, 5

\> a = x + Y

\> a

	[,1]	[,2]	[,3]
[1,]	14	-7	13
[2,]	19	-4	16
[3,]	24	-11	8

\> b = 2 * x + 3 * Y

b.

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

Scanned with CamScanner

P18

6) Marks of statistic of comp. sci student
59, 20, 25, 24, 46, 56, 55, 45, 27, 22, 47, 58, 54, 40, 50,
32, 36, 29, 35, 39

```
> x = c(59, 20, 25, 24, 46, 56, 55, 45, 27, 22, 47,
        58, 54, 40, 50, 32, 36, 29, 35, 39)
> 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)	3
3	[30, 35)	1
4	[35, 40)	3
5	[40, 45)	1
6	[45, 50)	3
7	[50, 55)	2
8	[55, 60)	4

A.Ku
28.11

Scanned with CamScanner

Practical NO - 2.

PROBABILITY DISTRIBUTION.

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

i)

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

Since, $p(2) = -0.5$, can't be a probability mass function. In p.m.f $P(x)$ is greater than 0 for all X.

ii)

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~~ solution $P(x) = 1$.

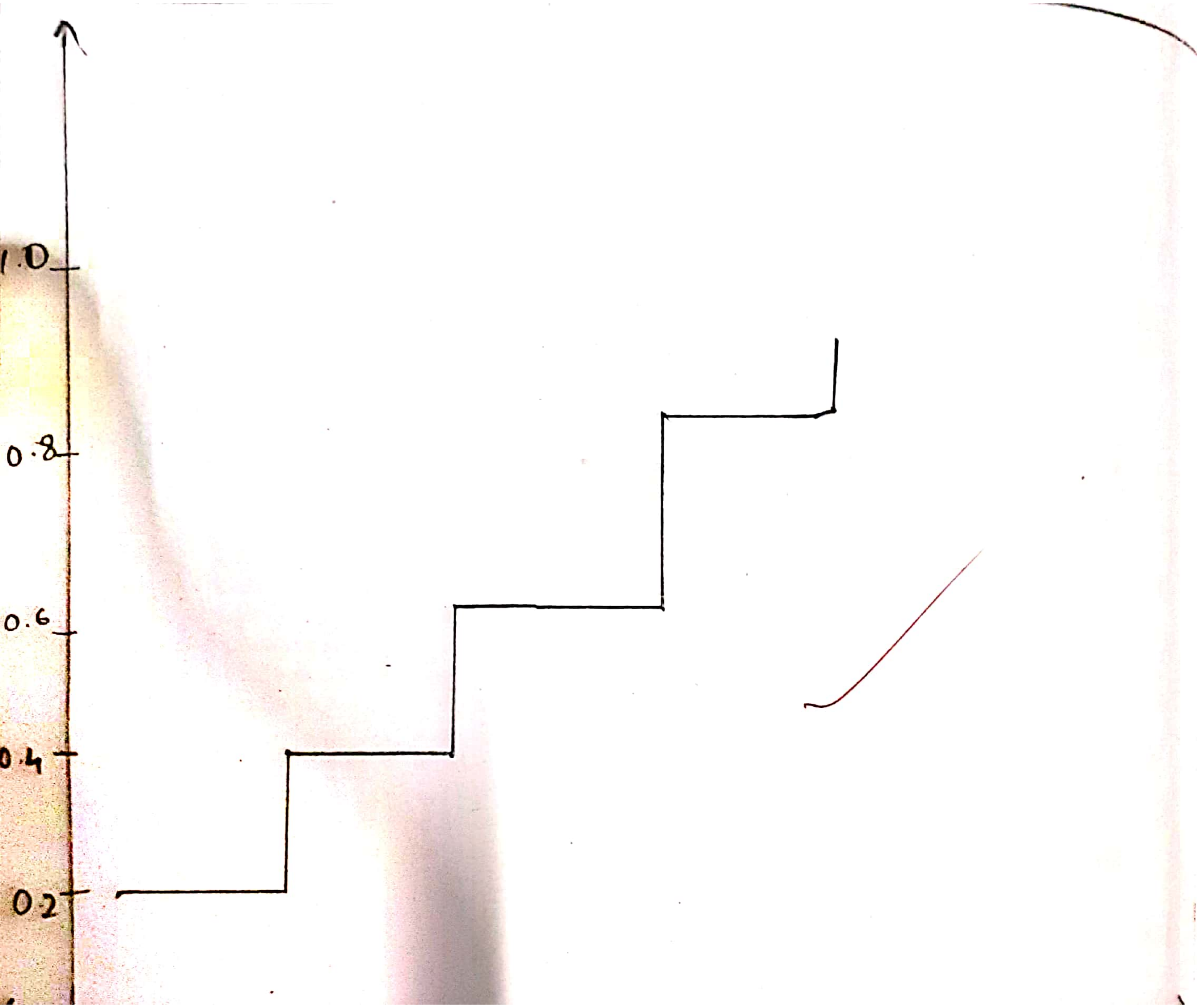
iii)

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

Since it's ~~to~~ sum is 1.

∴ It is p.m.f.

Scanned with CamScanner


1.0
0.8
0.6
0.4
0.2

PQD

i) Find c.d.f for the following p.m.f
  e) Sketch the graph.

x	1	2	3	4	5
f(x)	0.2	0.2	0.35	0.15	0.1.

F(x) = 0	x < 10
= 0.2	10 ≤ x < 20
= 0.40	20 ≤ x < 30
= 0.75	30 ≤ x < 40
= 0.90	40 ≤ x < 50
= 1.00	x ≥ 50.

ii)

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).
> cumsum (Prob).
x = c(1, 2, 3, 4, 5, 6).
Plot (x, xlab="probality", ylab= "x", main=
    "probability distribution", cumsum(prob).
```

Q.A?

3) Check whether the following is p.d.f or not.

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

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

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

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

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

$= 2$.

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

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

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

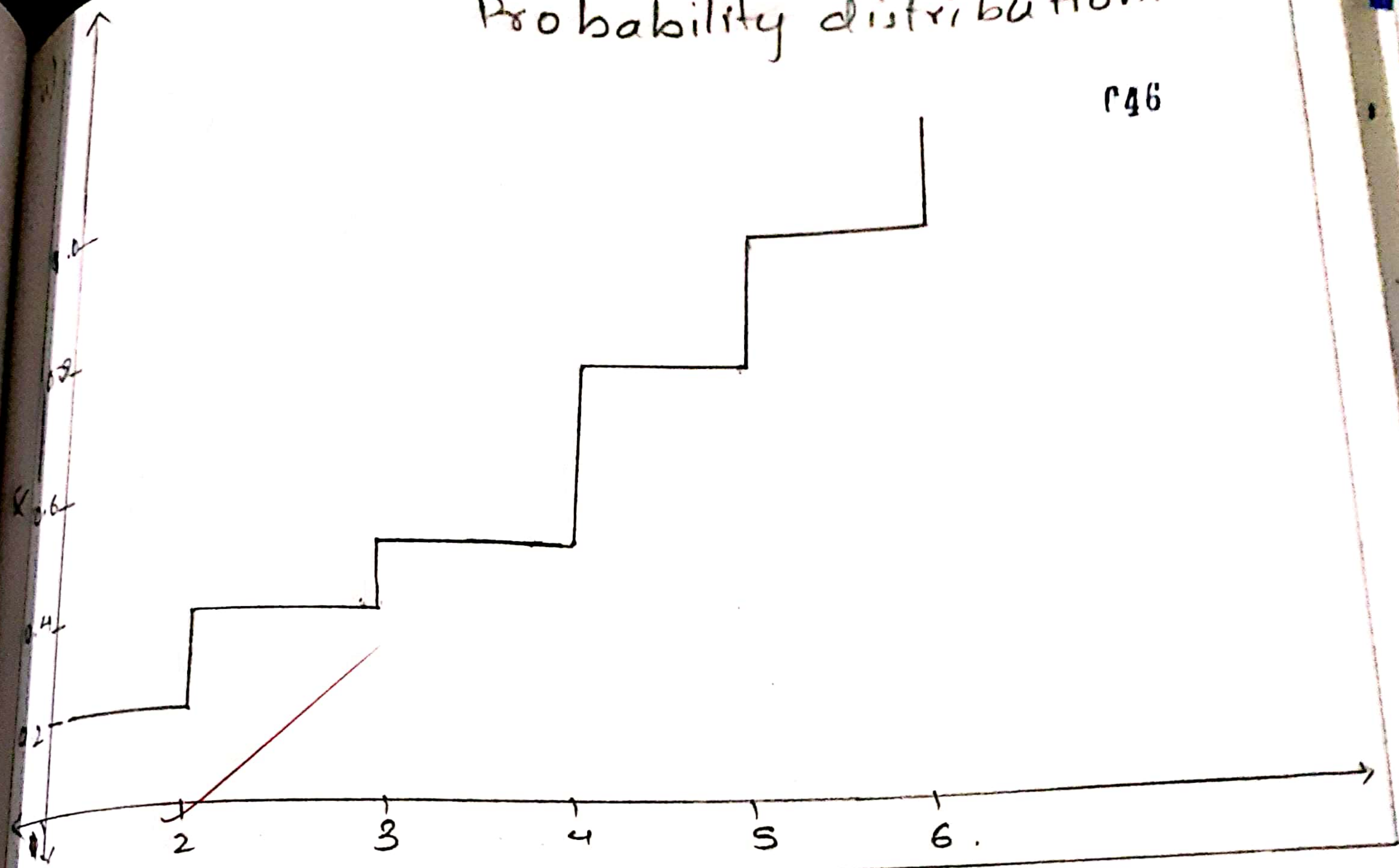
$= \left[\frac{\cancel{3}x^3}{\cancel{3}}\right]_0^1$

$= [1 - 0]$

$= 1$.

Arr 5/7/19

Probability distribution.



Probability.

PRACTICAL - 3.

1] Find the probability of exactly 10 in 100 trial with $p = 0.01$.

2] Suppose there are 12 MCQ. Suppose each Question has 5 opt out of which 1 is correct: find the probability of having
i) exactly 4 correct answer.
ii) atmost 4 correct answer.
iii) More than 5 correct answer.

3] Find the complete distribution when $n = 5$ & $p = 0.1$.

4] $n = 12$, $p = 0.25$, find the following probability
i) $P(X = 5)$
ii) $P(X < 5)$.
iii) $P(X > 7)$.
iv) $P(5 < X < 7)$.

The probability of a Salesman making a sell to a customer 0.15. find a probability of
i) no sells out of 10 customer.
ii) more than 3 sells out of 20.

Answer :

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

2) >dbinom (4, 12, 0.2)
[1] 0.1328756

ii) pbinom (4, 12, 0.2)
[1] 0.9274445.

iii) >pbinom (5, 12, 0.2).
[1] 0.01940528

3) >dbinom (0:5, 5, 0.1).

[1] 0.590499	0.32805	0.07290
0.00810	0.00045	0.00001

4) >dbinom (5, 12, 0.25).
[1] 0.1032414.

>pbinom (5, 12, 0.25).
[1] 0.9455978

>1-pbinom (7, 12, 0.25).
[1] 0.0027815

(0.25)).

[1] 0.04014945

i) >dbinom (0, 10, 0.15)

[1] 0.1968744

ii) >1-Pbinom (3, 20, 0.15)

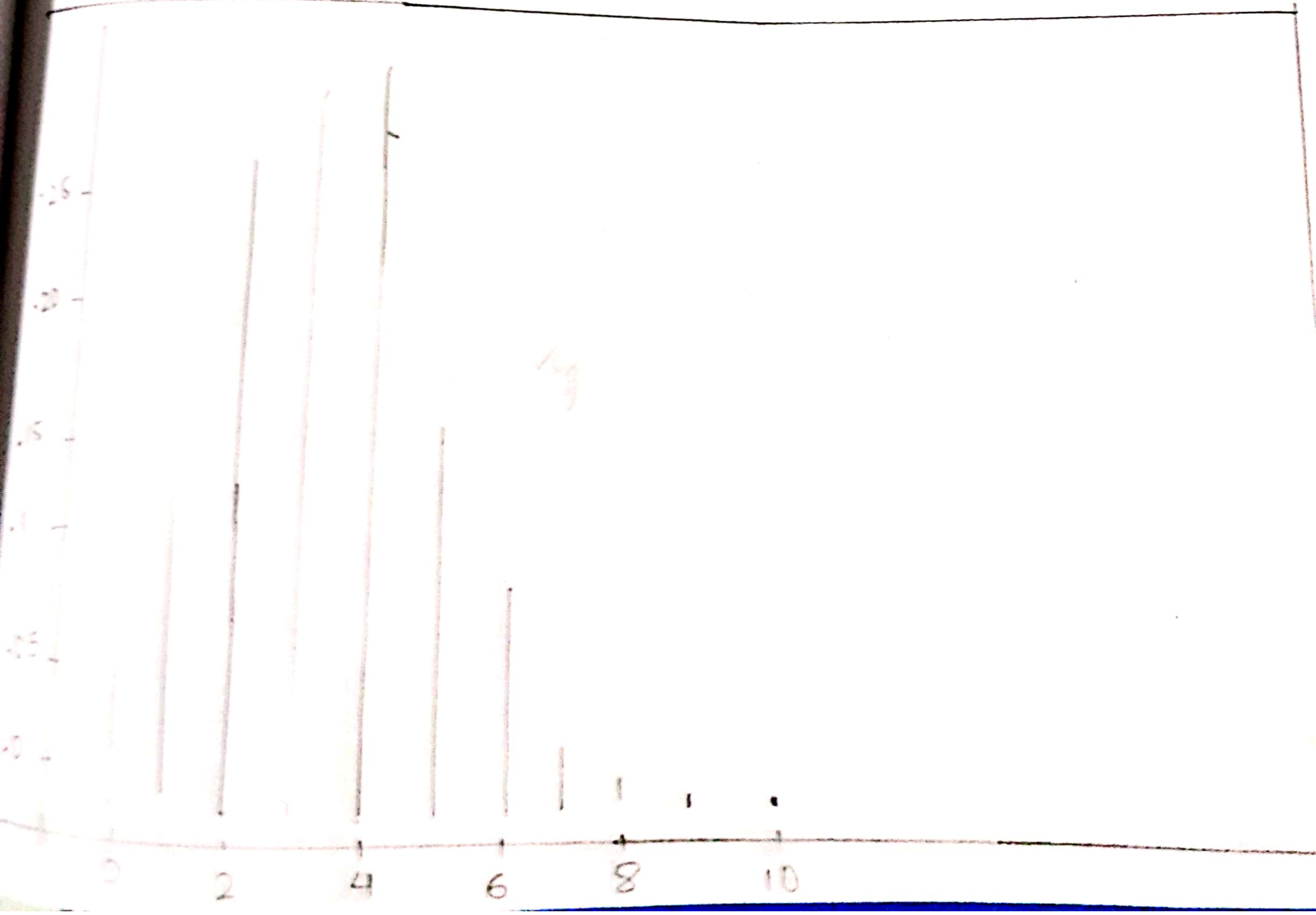
[1] 0.3522748

A sellsmen has a 20% probability of making a selling to a customer. Out of 30 customers what is a minimum no of sells hen can make with 88% probability.

>qbinom (0.88, 30, 0.2)

[1] 9

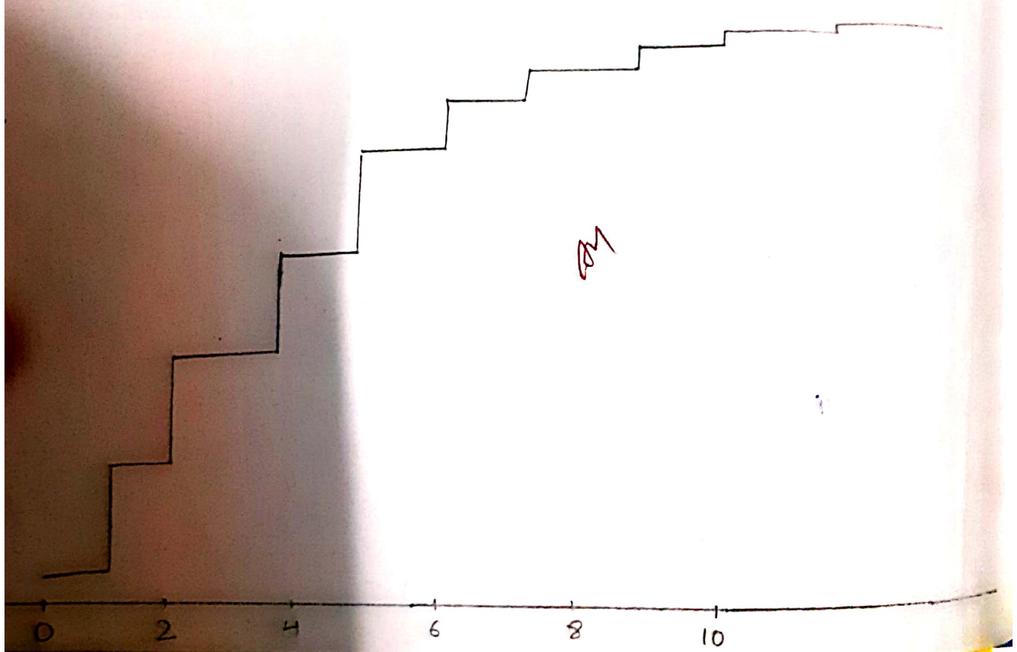
x follows binomial distributions with n=10 p=03 plot a graph of pmf & c.d.f.



> n = 10
> p = 0.3
> x = 0 : n
> prob = dbinom (x, n, p)
> cumprob = pbinom (x, n, p)
> d = data frame ("x values" = x, "probability" = prob)
> print(d)
[1]

	x values	probability
1	0	0.0282475249
2	1	0.1210608216
3	2	0.2334744405
4	3	0.2668279320

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



PRACTICAL - 4.

NORMAL DISTRIBUTION.

$p(x = x) = $ dnorm (x, μ, σ)
$p(x < x) = $ pnorm (x, μ, σ)
$p(x > x) = 1 - $ pnorm (x, μ, σ).
To generate random no from a normal distribution in (stadu random no). The rnorm (n, μ, σ).

1) A random variable X follows normal distribution with mean = $\mu = 12$ & standard deviation = $\sigma = 3$. find
i) $P(x \le 15)$
ii) $P(10 \le x \le 13)$
iii) $P(x > 14)$.
iv) Generate 5 observation. (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] 0.2524925
> rnorm (5, 12, 3).
[1] 11.253625  7.028228  10.848236  5.045339  13.
```

2) X follows normal distribution with $\mu = 10$ & $\sigma = 2$ find

i) probability $(X < 7)$

ii) $P(5 \le x \le 12)$.

iii) $P(X > 12)$.

iv) generate 10 observations.

v) find k. such that $P(X < k) = 0.4$.

```
→ > P1 = pnorm(7,10,2).
> P1.
[1] 0.0668072.
> P2 = pnorm(12,10,2) - pnorm(5,10,2).
> P2.
[1] 0.8351351.
> P3 = 1 - pnorm(12,10,2).
> P3.
[1] 0.1586553.
> rnorm(10,10,2).
[1] 9.021622   11.981599   9.079404   8.882875
    9.017746   13.423563   11.354241  10.159907
    12.327565  11.588792.
> qnorm(0.4,10,2)
[1] 9.493306.
```

Generate 5 Random no. for normal distribution with μ = 15, σ = 4, find Sample mean, median, S.D & print it.

```
> x = rnorm (5, 15, 4).
> x
[1] 14.29548  16.45377  14.20567  19.73212  11.87906
> am = mean (x).
> am.
[1] 15.30922.
> me = median(x).
> me.
[1] 14.29548.
> n=5
> variance = (n-1)*var(x)/n.
> variance.
[1] 6.967641.
> sd = sqrt (variance)
> sd.
[1] 2.639629.
> cat (" sample mean is = ", am).
Sample mean is = 15.30922 .
> cat ("sample median is = ", me).
Sample median = 14.29548.
> cat ("sample sd = ", Sd).
sample sd = 2.639629.
```

x follows normal x ~ N (30

i) find P(x ≤ 40)

ii) P(x > 35)

iii) P(25 < x < 35).

iv) 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).

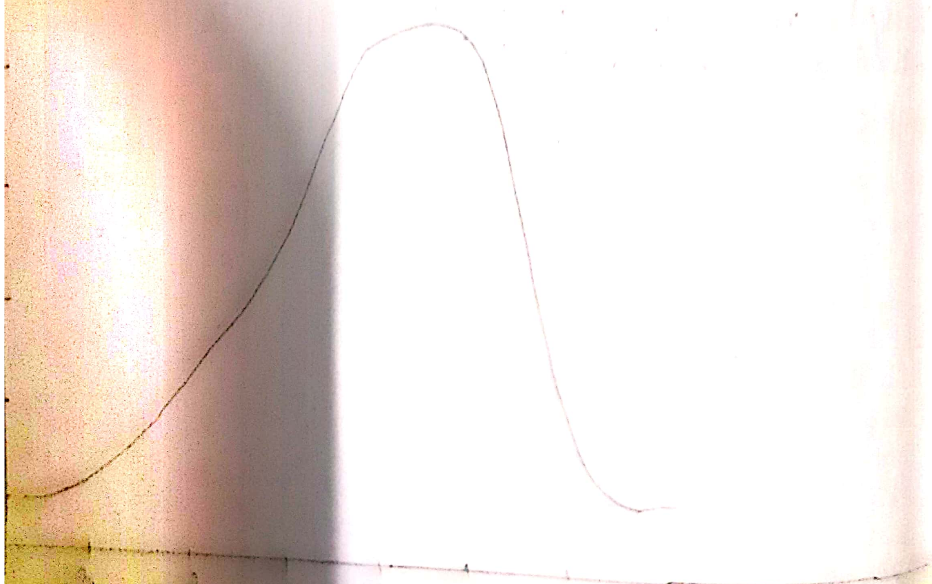
p3

[1] 0.3829249.

qnorm (0.6, 30, 10).

[1] 32.53347.

Standard normal Graph.



plot Standard normal graph.

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

Ap
2.1.20

PRACTICAL - 5

NORMAL & t - test.

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

Formula: $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $\quad$ $mo = \mu$ $\quad$ $mx = \bar{x}$

$sd = 6$ $\quad$ $n = n$.

→ z. mo = 15, mx = 14, sd = 3, n = 400.

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

z

[1] -6.666667.

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

pvalue.

[1] 2.616796e-11.

Since, P value is ~~more~~ less than 0.05, we ~~accept~~ reject ~~it~~. $H_0: \mu = 15$.

2) Test the Hypothesis. $H_0: \mu = 10$, $H_1: \mu_0 \neq 10$ Random Sample of Size 400 is drawn with sample mean 10.2 & S.D = 2.25. Test the Hypothesis at 5% level of significance.

→ mo = 10, mu = 10.2, sd = 2.25, n = 400.

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

z

[1] 1.777778

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

pvalue.

[1] 0.07544036.

Since, the pval is ~~less~~ more than 0.05

∴ we accept $H_0: \mu = 10$.

3) Test the Hypothesis H_0: proportion of smokers in college is 0.2.

A sample is collected & the sample porpor. is ~~calle~~ calculated = 0.125. test the hypo at level of significance. (sampe size = 400).

→ P = 0.2, p = 0.125, n = 400, Q = 1 - P.

z = (p - P)/(sqrt(P*q/n)).

z.

[1] -3.75

Pvalue = 2*(1-pnorm(abs(z))).

Pvalue

[1] 0.0001768346.

So we reject H_0: proportion = 0.2.

4) last year farmer lost ~~h~~ 25-20% of their crop Random sample of 60 fields are collected & is found that 9 fields crops are insect poll. Test the hypo at 1% level of significance.

→ ~~Ho = 0.2~~. Ho : ~~d~~P = 0.2.

H1 : p = 9/60 =

P = 0.2 , p = -0.9 n = 60 , Q = 1 - P.

→ P = 0.2 , p = 0.15 .

z = (p - P)/(sqrt(P * q/n)).

z

[1] -0.9682458

Pvalue = 2*(1 - pnorm(abs(z))).

Pvalue.

[1] 0.3329216.

5) Test the hypo $H_0 : \mu = 12.5$ from the following sample at 5% level of sig:

→ 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 = ~~length~~ mean (x).

mx

[1] 12.07

Variance = (n-1) * var(x)/n.

Variance

[1] 0.019521

```
sd = sqrt (variance)
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.
```

since the value is less than 0.5, value rejected

AM
16.1.20.

PRACTICAL-6.

LARGE SAMPLE TEST.

Q.1] 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 as the standard Deviation is 30. Test the hypothesis of the population is 250 or not at 5% level of significance.

→ mx = 275, mo = 250 SD = 30 n = 100.

```
z = (mx-mo)/(sd/(sqrt(n))).
[1] 8.333333.
pv = 2*(1-pnorm(abs(z))).
pv
[1] 0.
```

Since the p value is less than 0.05
∴ we reject Ho : μ

In a random Sample of 1k Student it is found that 750 use blue pen Test the hypothesis that the population proportion is 0.8 at 1 level of significance

n = 1000 p = 750/1000 P = 0.8 Q = 1 - P

p = 0.75.

z = (p - P)/(sqrt(P* q/n))

z

[1] -3.952847.

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

pv

[1] 7.72268e-05.

Since the value of p is less than 0.05.

∴ we reject Ho: P = 0.8.

3) Two random Sample of size 1000 & 2000 are drawn from 2 population with the same S.D 2.5. The Sample means are 67.5 & 68. respectively. Test the hypothesis Ho: $\mu_1 = \mu_2$ against H1: $\mu_1 \neq \mu_2$ at 5% l.O.S.

$n_1 = 1000$ $mx_2 = 68$

$n_2 = 2000$ $sd_1 = 2.5$

$mx_1 = 67.5$. $sd_2 = 2.5$.

z = (mx1-mx2)/sqrt((sd1^2/n1)+(sd2^2/n2)).

z

[1] -5.163978.

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

pv

[1] 2.417564e-07.

~~Accept~~. reject.

4) A study of noice level in two hospital that given below test the claim that the 2 hospitals have same level of noice at 1% L.O.S.

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

n1 = 84

n2 = 34

mx1 = 61.2.

mx2 = 59.4.

sd1 = 7.9

sd2 = 7.5

z = (mx1 - mx2)/sqrt((sd1^2/n1) + (sd2^2/n2))

[1] 1.162528.

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

pv

[1] 0.2450211.

Accept.

On a sample of 600 students in a college 400 use blue ink. In another college from a sample of 900 student 450 use blue ink. Test the hypo that the proportion of students using blue ink in 2 colleges are equal or not at 1% l.o.s.

→ $H_0 : P_1 = P_2$. $H_1 : P_1 \neq P_2$.

$n_1 = 600$

$n_2 = 900$.

p1 = 400/600 = 0.6

p2 = 450/900 = 0.5.

p = (n1 * p1 + n2 * p2)/(n1 + n2)

q = 1-p.

∴ z = (p1 - p2)/sqrt(p * q/(1/n1 + 1/n2))

[1] 6.381534.

pv = 2*(1 - pnorm(abs(z))).

pv

[1] 1.753222e-10.

∴ we reject $H_0 : P_1 = P_2$.

$H_0: P_1 = P_2$ $H_1: P_1 \neq P_2$

6) n1 = 200   n2 = 200
p1 = 44/200   p2 = 30/200.
q = 1-p.
p = q
[1] 0.815.

z = (p1-p2)/sqrt(p*q*(1/n1 + 1/n2)).
z
[1] 1.802741

pv = 2*(1-pnorm(abs(z))).
pv
[1] 0.07142888.

∴ Accept.

AM 23.01.20

PRACTICAL - 7

SMALL SAMPLE TEST

P56

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

$H_0 : \mu = 66$

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 = 1.588e-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

Since, the p-value is smaller ∴ we reject
the significance of at 5%.

2)

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

3) Two groups of students scored the following marks. Test the hypothesis that no significant difference between the two groups.

Grp-I = 18, 22, 21, 17, 20, 17, 23, 20, 22, 21

Grp-II = 16, 20, 14, 21, 20, 18, 13, 15, 17, 21

→ Ho: there is no diff betwn 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).

~~t0~~ two Sample t-test.

data: x & Y.

t = 2.2573 ; df = 16.376, p-value = 0.037

alternative hypothesis : true diff ~~betw~~ 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,

rejected.

The Sales data of 6 shops before & after are Special campaign given below

before data : 53, 28, 31, 48, 50, 42
After : 58, 29, 30, 55, 56, 45
Test the hypo that the campaign is effective.

H0 : There is no sig diff of sales before & 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 , p-value = 0.9806.
alternative hypo : true diff in means is greater than 0 95 percent confiden interval :
-6.035547 Inf
Sample estimates:
mean of the diff
-3.5.

Accepted.

Two medicines are applied to two groups of patient respectively

g1 : 10, 12, 13, 11, 14
g2 : 8, 9, 12, 14, 15, 10, 9

Is there any sig diff between two medicines?

H0: there is no signi diff.

x = c(10, 12, 13, 11, 14)
y = c(8, 9, 12, 14, 15, 10, 9).
t.test(x, y).

Sample t. test

t = 0.80384 df = 9.7994, p-value = 0.44

95 percent confidence interval:
-1.781171 3.781171.
mean of x and y.
12 11.

Since, the ~~val~~ p-value is greater than 0.05 therefore accept. 5%
of significance.

The following are the weights before & after the diet program. Is the diet program effective?

before : 120 , 125 , 115 , 130 , 123 , 119

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

H0: There is no sig diff.

x = c(120 , 125 , 115 , 130 , 128 , 119)

Y = c(100 , 114 , 95 , 90 , 115 , 99)

data : x & y.

t = 4.9293 , df = 5 , p.value = 0.997

- Inf 29.11493.

Mean of the differences

20.66667.

Since, the p-value is greater than 0.05

∴ accept 5% of significance.

6.1.20

Practical - 8.

Large & Small Sample test

① H0 : μ = 55 H1 : $\mu \neq$ 55

n = 100

mx = 52

m0 = 55

sd = 7

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

zcal

[1] -4.285714

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

Pvalue

[1] 1.82153e-05

we reject.

2) p = 350/700 Ho : P = $\frac{1}{2}$ = ?

n = 700 P = 0.5.

q = 1 - P.

z = (p - P)/(sqrt (P*q/n)).

z

[1] 0

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

pv

[1] 1.

Accept the pvalue at H0 : p = $\frac{1}{2}$.

Scanned with CamScanner

n1 = 1000
n2 = 1500
p1 = 0.02
p2 = 0.01.
p = (n1 * p1 + n2 * p2)/(n1+n2)
p
[1] 0.014
q = 1 - P
z = (p1 - p2)/sqrt (p* q*(1/n1 + 1/n2)
z
[1] 0.003474737.
pv = 2 * (1 - pnorm (abs(z)))
pv
[1] 0.03708364.

we reject

Ho : μ = 100.
mx = 99   sd = 8
mo = 100   n = 400.
z = (mx - mo) /(sd / (sqrt(n)))
z
[1] -2.5.
pv = 2 * (1 - pnorm (abs (z)))
pv
[1] 0.01241933

we reject.

Scanned with CamScanner

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

One Sample t-test

t = 47.94 df = 6 pvalue = 5.50

alternative

64.66479 71.62092.

mean of x

: 68.14286 :

we reject. Ho

7). Ho

n = 100

m0 = 1200

mx = 1150

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

z

[1] -4

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

pv

[1] 6.334248e-05.

we reject.

HO: $\sigma_1 = \sigma_2$.

x = c(66, 67, 75, 76, 82, 84, 88, 90, 92)

y = c(64, 66, 74, 78, 82, 87, 92, 93, 95, 97).

var.test (x, y).

F test to compare two variances.

= 0.70686, numdf = 8, denomdf = 10, p-value = 0.6359

95 percent confidence interval:

0.1833662 3.036093.

sample estimates:

ratio of variances

0.7068567.

we accept. Ho

Ho: $P_1 = P_2$

8). n1 = 200

n2 = 300

p1 = 44/200

p2 = 56/300

p = (n1 * p1 + n2 * p2)/(n1 + n2)

q = 1 - p.

z.

[1] 0.9128709.

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

pv

[1] 0.3613104

p

0.2.

We ~~reject~~ accept. Ho

A.M
13.2.20

→ 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-06
```

∴ H0: is rejected, ~~as~~ since pv is less than 0.05.

Test the hypothesis the vaccination and disease are independent or not.

Vaccine.

Disease.	Affected	Not Affected.
Affected	70	46
Not Affected	35	37

H0: Condition of disease & vaccination are independent.

x = c(70, 46, 35, 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.

X-squared = 2.0275, df = 1, pv = 0.1545

H0: is Accepted, Since, pv is more than 0.05.

3) Perform a ANOVA for the following data

TYPE	Observations
~~ABCD~~	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 & ind
F = 11.735, num df = 3, denom df = 9, p-value

> 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		

062

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

The following data gives the life of the tire 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

> H0: The 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 (~~values ~ind~~ (b1 = x1, b2 = x2, b3 = x3, b4 = x4))
names (d).
> oneway.test (values ~ ind, data = d, var.equal = T).
one - way analysis of means.
data :
F = 6.8445, num df = 3, denom df = 20,
p - value = 0.002349.
> anova = aov (values ~ ind, data = d)
summary (anova).

stats	maths
40	60
45	48
42	47
45	20
37	25
36	27
49	57
59	58
20	25
27	27

NON - PARAMETRIC TEST.

following are the amounts of sulphur oxide ~~ammitted~~ by an industry in 20 days. Sign test to test the hypothesis that the population medium 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 medium is 21.5

```
x = c(17, 15, 20 ... 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.
```

Accepted.

2) give the alternative is following is a data of 10 observation apply Sign test to test the hypothesis that the population medium is 625 against the alternative it is more than 625.

612,

→ H0 : population medium is 625.

```
x=c(612, 619, 631, 628, 643,640, 655, 649, 670,
    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.
```

Accepted.

The following are the values of a sample test the hypothesis that the population medium is 160 against the alternative it is more than 160 at 5% LOS using
63, 65, 60, 89, 61, 71, 58, 51, 69, 62, 63, 39, 72, 69, 48, 66, 72, 63, 87, 69.

H0: population medium is equal to 60
H1: population median is greater than 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.

Since, pvalue is less than 0.05, ∴ we rejected.

is 12 or less than 12.

→ Ho: population median is less than

x = c(15, 17, 24, 25, 20, 21, 32, 28, 12, 12, 24, 26).

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

wilcoxon S R T with continuity c

data : x.

V= 66 , p.value = 0.9986..

alternative hypothesis : true location is less than 12.

Since pvalue is more than 0.05
∴ it is accepted.

5). The weights of students before or after they stopped smoking are given below. using WSRT test that there is no significant change.