PRACTICAL : 01

Title : Random Variable

] Find the mean and variance for the following:

a]

X	-1	0	1	2
P(x)	0.1	0.2	0.3	0.4

Solution:

X	P(x)	X.P(x)	$E(x)^2$	$[E(x)]^2$
-1	0.1	-0.1	0.1	0.01
0	0.2	0	0	0
1	0.3	0.3	0.3	0.09
2	0.4	0.8	1.6	0.64
Total	$\Sigma = 1$	$\Sigma = 1$	$\Sigma E(x)^2 = 2.0$	$\Sigma[E(x)]^2 = 0.74$

$\therefore$ Mean $= E(x) = \Sigma x_i \cdot p(x) = 1$

$\therefore$ Variance $= V(x) = \Sigma E(x)^2 - \Sigma[E(x)]^2$

$= 2 - 0.74$

$= 1.24$

$\therefore$ Mean $E(x) = 1$ & Variance $V(x) = 1.24$

b]

x	-1	0	1	2
$p(x)$	$1/8$	$1/8$	$1/4$	$1/2$

Solution:

x	$P(x)$	$x \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-1	$1/8$	$-1/8$	$1/8$	$1/64$
0	$1/8$	0	0	0
1	$1/4$	$1/4$	$1/4$	$1/16$
2	$1/2$	1	2	1
Total	$\Sigma = 1$	$\Sigma = 9/8$	$\Sigma = 19/8$	$\Sigma = 69/64$

$\therefore$ Mean $= E(x) = \Sigma x \cdot p(x) = 9/8$

$\therefore$ Variance $= V(x) = \Sigma E(x)^2 - \Sigma[E(x)]^2$

$$= \frac{19}{8} - \frac{69}{64}$$

$$= \frac{152 - 69}{64}$$

$$= \frac{83}{64}$$

$\therefore$ Mean $E(x) = 9/8$ & Variance $V(x) = 83/64$

X	-1	0	1	2
$P(x)$	$k+1/13$	$k/13$	$1/13$	$k-4/13$

$$\therefore \Sigma P(x_i) = 1 = \frac{k+1}{13} + \frac{k}{13} + \frac{1}{13} + \frac{k-4}{13}$$

$$1 = \frac{k+1+k+1+k-4}{13}$$

$$13 = 3k-2$$
$$15 = 3k$$
$$k = 5$$

x	$P(x)$	$X \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-1	$6/13$	$-6/13$	$6/13$	$36/169$
0	$5/13$	0	0	0
1	$1/13$	$1/13$	$1/13$	$1/169$
2	$1/13$	$2/13$	$4/13$	$4/169$
Total	$\Sigma = 1$	$\Sigma = -3/13$	$\Sigma = 11/13$	$\Sigma = 41/169$

$$\therefore \text{Mean} = E(x) = \Sigma . XP(x) = -3/13$$

$$\therefore \text{Variance} = V(x) = \Sigma E(x)^2 - \Sigma [E(x)]^2$$
$$= 11/13 - 41/169$$
$$= 143 - 41/169$$
$$= 102/169$$

$$\therefore \text{mean} = -3/13 \ \& \ \text{Variance} = 102/169$$

(iv) $P(X \geq 0) = 1 - F(0) + P(0)$
$= 1 - 0.45 + 0.15$
$= 0.40$

Q.4 Let f be continuous random variable with pdf
$$f(x) = \frac{x+1}{2} \qquad -1 < x < 1$$
$$= 0 \qquad \text{otherwise}$$
Obtain c.d.f of x

Soln. By definition of cdf,
we have
$$F(x) = \int_{-1}^{1} t\,dt$$
$$= \int_{-1}^{1} \frac{x+1}{2}\,dx$$
$$= \frac{1}{2}\left(\frac{1}{2}x^2 + x\right) \text{ for } -1 \leq x \leq 1$$
Hence the c.d.f is,
$$F(x) = 0$$
$$= \frac{1}{4}x^2 + \frac{1}{2}x \quad \text{for } -1 \leq x \leq 1$$
$$= 0 \quad \text{for } x \geq 1$$

The p.m.f of random ~~variable~~ variance x is given by

x	-3	-1	0	1	2	3	5	8
P(x)	0.1	0.2	0.15	0.2	0.1	0.15	0.05	0.05

Obtain cdf find (i) $P(-1 \le x \le 2)$ (ii) $P(1 \le x \le 5)$ (iii) $P(x \le 2)$ (iv) $P(x \ge 0)$

Soln

x	-3	-1	0	1	2	3	5	8
P(x)	0.1	0.2	0.15	0.2	0.1	0.15	0.05	0.05
F(x)	0.1	0.3	0.45	0.65	0.75	0.90	0.95	10

(i) $P(-1 \le x \le 2) = P(x \le 2) - P(x \le -1) + P(x = -1)$
$= F(x_b) - F(x_a) + P(a)$
$= F(2) - F(-1) + P(-1)$
$= 0.75 - 0.3 + 0.2$
$= 0.25$

(ii) $P(1 \le x \le 5) = F(x_b) - F(x_a) + P(a)$
$= F(5) - F(1) + P(1)$
$= 0.95 - 0.65 + 0.2$
$= 0.15$

(iii) $P(x \le 2) = P(x = -3) + P(x = -1) + P(x = 0) + P(x = 1) + P(x = 2)$
$= 0.1 + 0.2 + 0.15 + 0.2 + 0.1$
$= 0.75$

Let f be continuous random variable with p.d.f.

$\therefore f(x) = \frac{x+2}{18} \qquad -2 \le x \le 4$

$= 0 \qquad$ Otherwise

Calculate c.d.f

Soln: By definition of c.d.f we have,

$F(x) = \int_{-2}^{4} t \, dt$

$= \int_{2}^{4} \frac{x+2}{18} \, dx$

$= \frac{1}{18}\left(\frac{1}{2}x^2 + 2x\right)$

for $-2 \le x \le 4$

Hence c.d.f is

$F(x) = 0 \quad$ for $x < -2$

$= \frac{1}{18}\left(\frac{1}{2}x^2 + 2x\right)$

for $-2 < x < 4$

$= 0$ for $x \ge 4$

c]

X	-3	10	15
P(X)	0.4	0.35	0.25

Solution:

X	P(x)	X.P(x)	$E(x)^2$	$[E(x)]^2$
-3	0.4	-1.2	3.6	1.44
10	0.35	3.5	35	12.25
15	0.25	3.75	56.25	14.0625
Total	$\Sigma = 1$	$\Sigma = 6.05$	$\Sigma = 94.85$	$\Sigma = 27.7525$

$\therefore$ Mean $= E(x) = \Sigma X . P(x) = 6.05$

$\therefore$ Variance $= V(x) = \Sigma E(x)^2 - \Sigma [E(x)]^2$

$= 94.85 - 27.7525$

$= 67.0975$

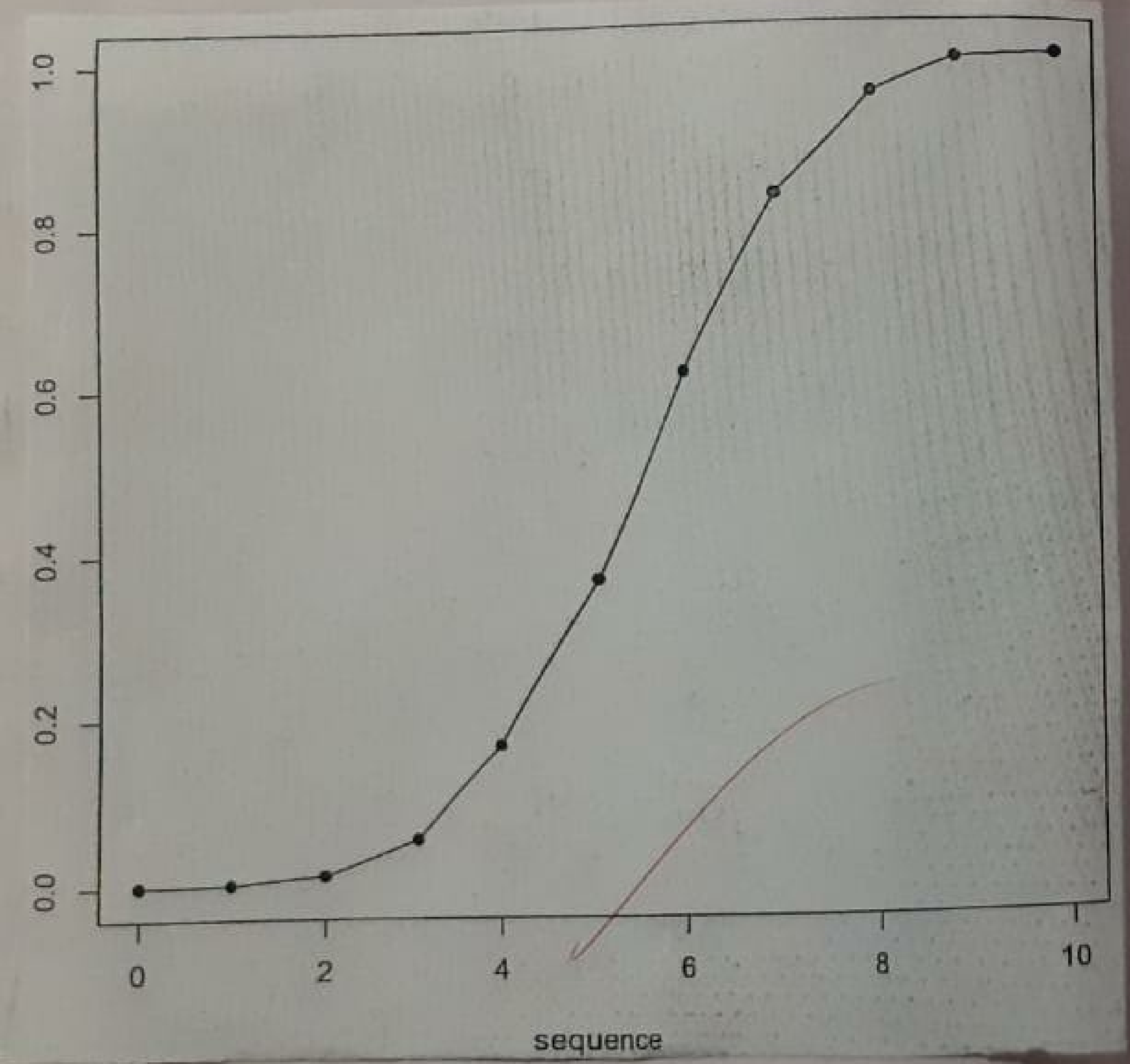
$\therefore$ Mean $E(x) = 6.05$ & Variance $V(x) = 67.0975$

Q2] If P(X) is p.m.f of a random variable X. If p(x) represents p.m.f for random variable X. Find value of k. Then evaluate mean & variance.

Solution: As P(xi) is a p.m.f it should satisfy the properties of p.m.f which are:

a] $p(x_i) > 0$ for all sample space

b] $\Sigma P(x_i) = 1$




```
> dbinom (3,6,0.3)
[1] 0.18522
> dbinom (2,6,0.3) + dbinom(3,6,0.3) + dbinom (4,6,0.3)
[1] 0.74373
```

4) For n= 10, p = 0.6, evaluate binomial probabilities and plot the graphs of p.m.f & cdf

```
> x = seq (0,10)
> y = dbinom (x, 10, 0.6)
> y
[1] 0.0001048576   0.0015728640   0.0106168320
    0.0424673280   0.1114767360   0.2006581248
    0.2508226560   0.2149908480   0.1209323520
    0.0403107840   0.0060466176

> plot (x, y, xlab = "sequence", ylab = "probabilities", "l", pch=16)
> x = seq (0,10)
> y = pbinom (x, 10, 0.6)
> plot (x, y, xlab = "sequence", ylab = "probabilities", "o", pch = 16)
```

5) Generate a random sample of size 10 for a B.D → B(8,0.3). Find the mean & the variance of the sample

```
> rbinom (8, 10, 0.3)
[1] 2 2 3 4 3 4 2 3
> mean (rbinom (8, 10, 0.3))
[1] 2.375
```

PRACTICAL 2
Title : Binomial Distribution

Q.1] An unbiased coin is tossed 4 times. Calculate probability of obtaining no head, atleast one h.. & more than one tail.

NO HEAD:

```
> dbinom (0,4,0.5)
[1] 0.0625
```

ATleast one Head

```
> 1-dbinom (0,4,0.5)
[1] 0.9375
```

More than one Tail:

```
> pbinom (1,4,0.5,lower tail=F)
[1] 0.9375
```

Q.2] The probability that student is accepted to a prestigious college is 0.3. If 5 students supply. what's the probability of almost 2 are accepted

```
> pbinom(2,5,0.3)
[1] 0.83692
```

Q.3] An unbiased coin is tossed 6 times the probabili.. of head at any toss = 0.3. Let x be no. of hea.. that comes up. Calculate P(x = 2), P(x= 3), P(1<x<5

```
> dbinom (2,6,0.3)
[1] 0.324135
```

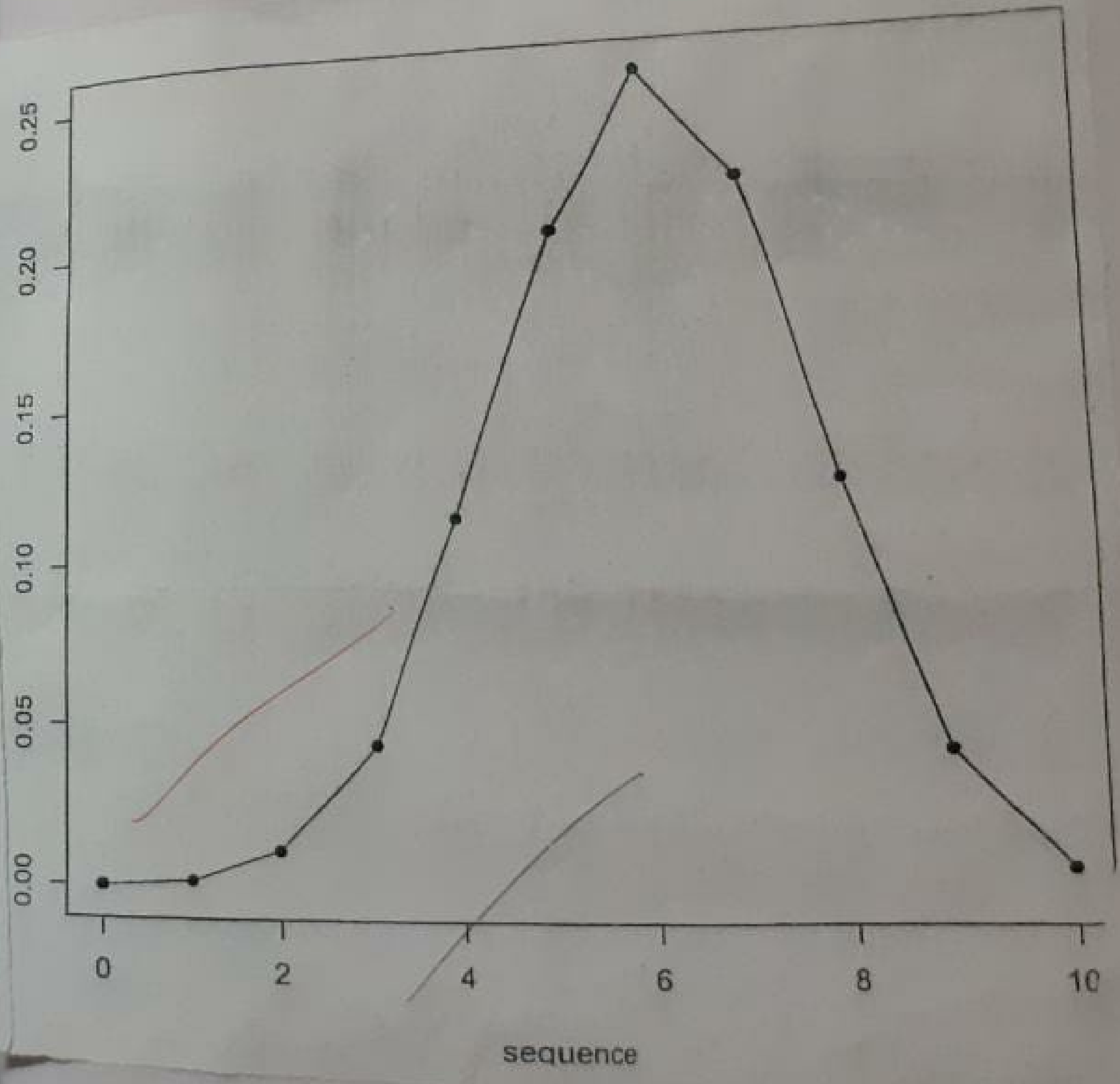
```
> var (rbinom(8,10,0.3))
[1] 1.696469
```

Q.6 The probability of men hitting the target is ... if he shoots 10 times what is the probabili... that he hits the target exactly 3 times, probabi... that he hits the target atleast one time.

```
> dbinom(3,10,0.25)
[1] 0.2502823
> 1- dbinom(1,10,0.25)
[1] 0.8122883
```

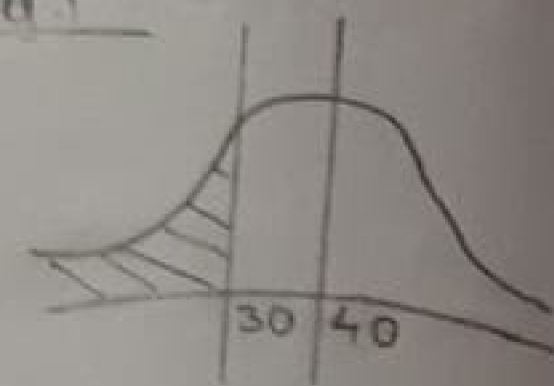
Q.7) Bits are sent for communication channel in packet ... 12. If the probability of bit being corrupted is 0... What is the probability of no more than 2 bits ... corrupts in a packet?

```
>pbinom (2,12,0.1,lower.tail=F)+dbinom(2,12,0.1)
[1] 0.3409977
```

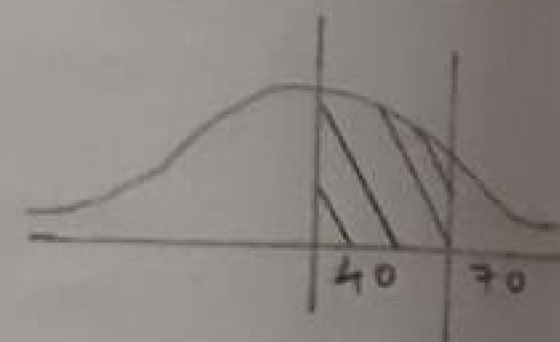


Q.1

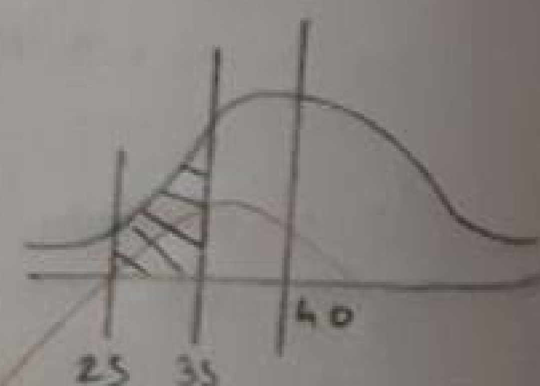
a]



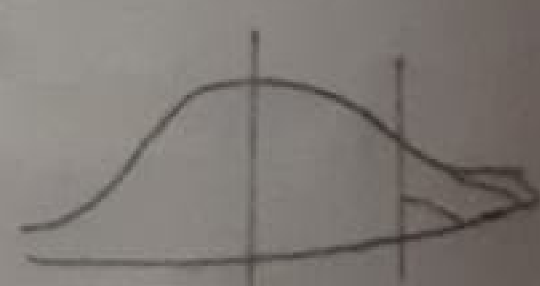
b]



c]



d]



PRACTICAL : 3

Title : Normal Distribution

1] A normal distribution of 100 students with mean = 40, 50:
Find no. of students whose marks are :
① $P(x < 30)$ ② $P(40 < x < 70)$ ③ $P(25 < x < 35)$ ④ $P(x > 60)$

```
> pnorm(30,40,15)
[1]   0.25 24925
```

```
> pnorm(70,40,15) - pnorm(40,40,15)
[1] 0.477 2499
```

```
> pnorm(35,40,25)-pnorm(25,40,15)
[1] 0.2107 861
```

```
> 1- pnorm(60,40,15)
[1] 0.09121122
```

4) A random variable x follows normal distribution with $\mu = 10, \sigma = 2$, Generate 100 observations and evaluate its mean, median & variance.

```
> x = rnorm(100,10,2)
> summary(x)
[1] Min    1st Q   Median   Mean    3rd Q    Max
    5.713  8.444   9.723    9.914   1.325    14.238
> Var(x)
[1] 3.648924
```

5) Write a command to generate 10 random numbers for normally distibution with $\mu = 50, \sigma = 4$. Find the sample mean & median.

```
> x = rnorm(10,50,4)
> summary(x)
Min     1st Q   Median   Mean    3rd Q    Max
44.73   50.46   52.01    52.35   54.39    58.85
```

Q1) If the random variable x follows the normal distribution with mean = 50., V = 100. Find (1) $P(x \le 70)$
(2) $P(x > 65)$ (3) $P(x \le 32)$ (4) $P(35 < x < 60)$
(5) $P(20 < x < 30)$

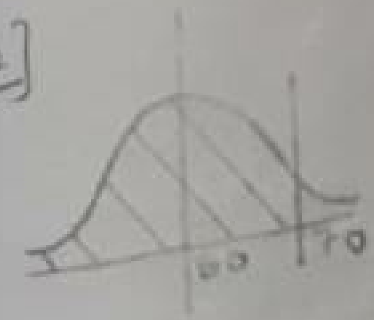
```
> pnorm (70,50,10)
[1] 0.9772499

> 1-pnorm (65,50,10)
[1] 0.0668072

> pnorm (32,50,10)
[1] 0.03593032

> pnorm (60,50,10) - pnorm (35,50,10)
[1] 0.7745375

> pnorm (30,50,10) - pnorm (20,50,10)
[1] 0.02140023
```

Q3) Let $X \sim N(160, 400)$, find k_1 & k_2 such that $P(x < k_1) = 0.8$ & $P(x > k_2) = 0.6$

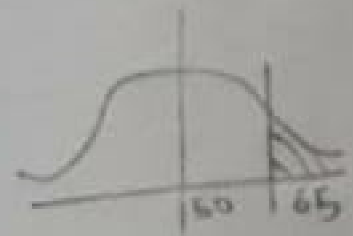
```
> qnorm (0.6,160,20)
[1] 165.0669
> qnorm (0.8,160,20)
[1] 176.8324
```

Q.2]

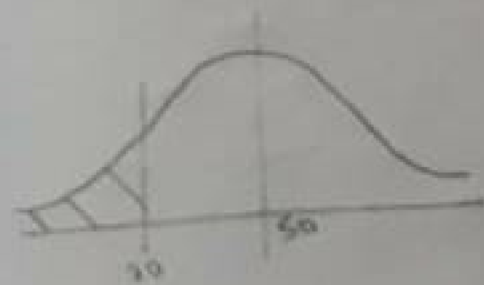
a]



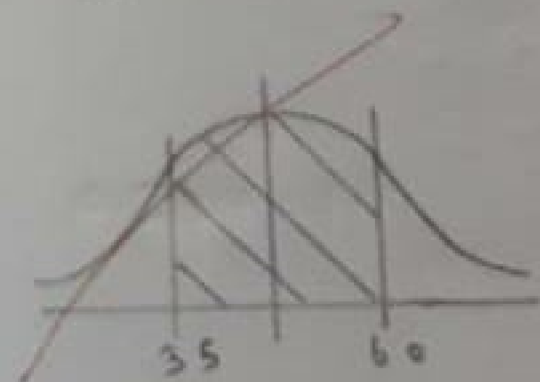
b]



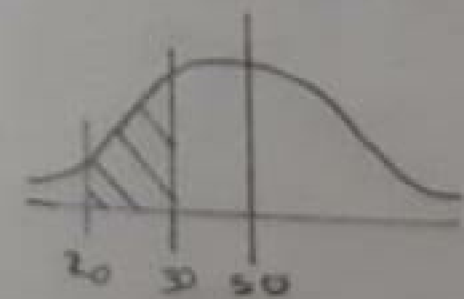
c]



d]



e]



E.P

Practical : 04

* Sample mean & d. deviation given single population

Q.1 Suppose the food level on the cookie bag's states that it has atmost 2 gms of saturated salt in a single cookies. In a sample of 55 cookies, it was found that mean and of saturated salt per cookie is 2.1 gm. Assume that the same std. deviation is 0.3 at 5% level of significance can be rejected the claim on food label.

To check whether reject or accept null hypothesis at 95% level of confidence or 5% level of significance.

$\Rightarrow \sigma = 0.3$

$n = 35$

$\bar{x} = 2.1$

$\mu = 2$

H_0 (null hypothesis) = $\mu < 2$

H_1 (alt. hypothesis) = $\mu > 2$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\therefore z = \frac{2.1 - 2}{\frac{0.3}{\sqrt{35}}} = 1.972027$$

p value = 1 - pnorm (2)

= 0.0243

$\therefore$ Reject all null hypothesis

$\therefore$ Accepted alternate hypothesis $\because$ p value < 0.05

A sample of 100 customers was randomly selected & it was found that average spending was 275/-. The SD = 30 using 0.05 level of significance, would you conclude that the amount spent by the customer is more than 250/- whereas the restaurant claim that it is not = 250/-

$\Rightarrow$ $\bar{x} = 275$, $\mu = 250$, $\sigma = 30$, $n = 100$.

$H_0 = \mu < 250$

$H_1 = \mu > 250$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{275 - 270}{\frac{30}{\sqrt{100}}} = 8.333$$

>> pt (z, 99, lower tail = F)

$\therefore$ p value = 2.30573 6e-13

$\Rightarrow$ Reject the null hypothesis $\because$ p value < 0.05

Accept the alternate hypothesis ($\mu > 250$)

pvalue = 5.88567e-06

Reject the null hypothesis = ~~1~~ claim of principle. $[\mu = 100]$

Method 2 tail test

$H_0 = \mu = 100$

$H_1 = \mu \neq 100$

→ Pvalue = 2x(1 - Pnorm(abs(z))) = 1.177134e-05

→ Reject the null hypothesis ∴ pvalue < 0.05

<u>Single population portion:</u>

Q) It is believed that coin is fair. The coin is tossed 40 times; 28 times - Head occurs. Indicate whether the coin is fair or not at 95% LOC.

→ $z = \dfrac{p - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$ → p: probability of sample; p_0 → probability of population

$p_0 = 0.5$

$q_0 = 1 - p_0 = 0.5$

$p = \dfrac{28}{40} = 0.7$

$n = 40$

$z = \dfrac{0.7 - 0.5}{\sqrt{\dfrac{0.5 \times 0.5}{40}}}$

$H_0 = \mu = 0.5$

$H_1 = \mu \neq 0.5$

Q.3] A quality control engineer finds that sample 100 have average life of 470 hours. Assuming pop test whether the population mean is 480 hours or population mean < 480 hours at 1.05 → 0.05

$n = 100, \bar{x} = 470, \mu_1 < 480, \sigma = 25, \mu = 480$

>> $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = -4$

>> pt (z, 99, lower.tail = F)
= 6.112576e.05

⇒ Reject the null hypothesis ∴ $p < 0.05$
Accept the alternate hypothesis ($\mu_1 < 480$)

Q.4) A principal at school claims that the IQ is 100 of the students. A random sample of 30 students whose IQ was found to be 112. The SD of population = 15. Test the claim of principal

Method 1 : 1 tail test

$H_0 = \mu - 100$

$H_1 = \mu = 100$

$\bar{x} = 112, SD = 15, \mu = 100, n = 30.$

$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

$= \frac{112 - 100}{\frac{15}{\sqrt{30}}} = 4.38178$

$H_0 = [p_0 = p]$
$H_1 = [p \neq p_0]$
>>> $z = [0.5416 - 0.5) / (sqrt(0.5 * 0.5(600))$
>>> $z = 2.037475$
>>> pvalue $= 2 * (1 - pnorm(abs(z)))$
∴ pvalue $= 0.04155234$

∴ Reject the null hypothesis = pvalue < 0.5
∴ Accept the alternate hypothesis i.e $p_0 \neq p_0$

Formula:

$$z = \sqrt{pq\left(\frac{1}{n} + \frac{1}{m}\right)} \quad \text{when} \quad p = \frac{p_1 n + p_2 m}{n + m}$$

In an electric campaign, a telephone phone of 800 registered voters shows favour 460. Second pole opinion s20 of 100 registered voters favored the candidate at 0.5 % LOC (level of confidence), is there sufficient evidence that popularity has decreased.

$n = 800$, $p_1 = 460/800 = 0.575$, $m = 1000$
$p_2 = 520/1000 = 0.52$
$(0.575 * 800 + 0.52 \times 1000) / (1800)$
$p = 0.54444.$
$z = sqrt(0.54 * 0.46) * 1/8$
$z = 0.00112139 4$
$H_0 = p = 0.544$
$H_1 = p < 0.544$

>> pvalue = 2 * (1 - pnorm(abs(z)))
∴ pvalue = 0.01141209

→ Reject the null hypothesis ∵ $p < 0.05$
Accept the alternate hypothesis

Q) In an hospital - 480 females & 520 males are born in a week. Do confirm that male & female born are equal in no.

$$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \qquad p \to \frac{520}{1000} = 0.52, \; p_0 = 0.5, q_0 = 0.5, n = 1000$$

$H_0 = [p = p_0]$
$H_1 = [p \neq p_0)$
>>> z = (p-p0) / (sqrt(p0q0/n))
>>> z = 1.2645
pvalue = 2 * (1-pnorm(abs(z)))
∴ pvalue = 0.2060506
∴ Reject the null hypothesis ∵ pvalue < 0.5
∴ Accept the alternate hypothesis ie $p \neq p_0$

Q.2) In a big city, 325 men out of 600 men are found to be self employed conclusion is that maximum men in city are self employed.

$$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \qquad p \to 325/600 = 0.541607, \; p_0 = 0.5, q_0 = 0.5, n$$

p value = (2 × (1 - pnorm(abs(z))))
∴ pvalue = 0.9991053

Accept the null hypothesis ∵ pvalue > 0.5
Accept p = 0.5444

Q.2 From a consignment A, 100 articles are dr.. & 44 were found defectives from consignment B, 200 samples are drawn out of which are defective. Test whether the proportion of defective items in 2 consignments are significant different.

$H_0 = p_1 = p_2$
$H_1 = p_1 \neq p_2$

$p_1 = 44/200 = 0.22$
$n = 200 = m$
$p_2 = 30/200 = 0.15$

>>> $p = \left(\frac{p_1 n + p_2 m)}{n+m} \right.$

>>> p = (0.22 * 200 + 0.15 * 200)/400
>>> p = 0.125

>>> z = (sqrt(0.185 * 0.815) * (2/200))
>>> z = 0.003882976

pvalue = 2 * (1-pnorm(abs(z)))
∴ pvalue = 0.9969018
p value > p
∴ Accept the null hypothesis i.e $p_1 = p_2$

Q.2 A dice is tossed 120 times & following results are obtained.

No. of times	Frequency
1	30
2	25
3	18
4	10
5	22
6	15

Test the hypothesis that dice is unbiased.

∴ H_0 = dice is unbiased , H_1 = dice is biased

```
> obs = c(30, 25, 18, 10, 22, 15)
> exp = sum(obs) / length(obs)
> exp
[1] 20

> z = sum((obs - exp)^2/exp)
> pchisq(z, df = length(obs)-1)
[1]   0.956659
```

∴ Accept the null hypothesis
∴ Dice is unbiased.

Q.4

	Graduate	Undergraduate
Online face	20	25
to face	40	5

Is there any association between student's prefrence for type of education & method.

$\therefore H_0 = $ Independent, $H_1 = $ Dependent

```
> x = c(20,40,25,5)
> z = matrix(x, nrow = 2)
> chisq.test(z)
      Pearson's chi-squared test with Yath's
      contenuity correction.
  data = z
  x - squared = 18.05 , df = 1, p-values = 2.157e-0.5
```

$\therefore$ Reject null hypothesis

$\therefore$ Both are dependent.

Practical 5

Title : Chi-square Test

1) Use the following data to test whether the attribute conditions of home & child are independent.

		Condition of Homes	
		Clean	Dirty
Condition of a child	clean	70	50
	f. clean	80	20
	Dirty	35	45

H_0 = Both are independent, H_1 = Both are dependent

```
> x = c(70,80,35)
> y = c(50,20,45)
> z = data.frame(x,y)
> z
[1]        x    y
  1       70   50
  2       80   20
  3       35   45
> chisq.test (z)
```

Pearson's chi squared test

data : z

x - squared = 25.646, df = 2, pvalue = 2.698e-06

∴ Reject the null hypothesis

∴ Both are dependent

A die is tossed 180 times

No. of terms	Frequency
1	20
2	30
3	35
4	40
5	12
6	43

Test the hypothesis that dice is unbiased

H_0 = dice is biased
H_1 = dice is unbiased

```
> x = c(20,30,35,40,12,43)
> chisq.test(x)
    Chi squared test for given probabilities

  data: x
  x-squared = 23.933, df = 5, p-value = 0.000223
```

∴ Reject null hypothesis
Die is unbiased.

Rew
Fals

An Iq test was conducted & the students were observed before & after training the result are following.

Before	After
110	120
120	118
123	125
132	136
125	121

Test whether there is change in the Iq after the training.

H_1 = no change in Iq
H_0 = Iq increased after training.

```
> a = c(120, 118, 125, 136, 121)
> b = c(110, 120, 123, 132, 125)
> z = sum((b-a)^2)/a)
> pchisq(z, df = length(b)-1)
[1]  0.1135959
```

Accept the ~~null~~ hypothesis
∴ ~~There is~~ change in Iq after training.

> t. test (e1, e2, paired = T, alter = "less",
conf. level = 0.99).

Paired t.test
data = e1 and e2

t = -1.4832 , df = 10, p-value = 0.08441
alternative hypothesis the difference in means
is less than 0.99 percent confidence interval
= Dni 0.863333
Sample estimates
Mean of the difference
∴ Accept H_0, Reject H_1

Q.4 Two drugs for BP was given 5 data was calculated

D_1 = 0.7, -1.6, -0.2, -1.2, 0-1, 3.4, 3.7, 0.8, 0.2
D_2 = -1.9, 0.8, 1.1, 0.1, -0.1, 4.4, 5.5, 1.6, 4.6, 2.3

The two drugs have some affect, check whether two drugs have same effect on patient or not

7 b = c (44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21)

t.test (a, b, Paired = T, alter = "two sided", conf. level 0.1)
paired test
data: a and b

a = c(25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 16, 21)
b = c(44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21)

t.test(a, b, Paired = T, alter = "two sided", conf level = 0.95)
Paired t.test
data = a and b
t = -0.62787, df = 11, p.value = 0.5412
alternative hypothesis : true difference in means is not equal to 0 to 95 percent confidence
interval = 14.267330 7.933947
Sample estimates
Mean of the differences -3.166667

∴ Accept ~~H~~ H_0, Reject H_1
∴ There is no difference in weight

ii) Students gave the test after 1 month they again gave the test after the tuitions do the marks gives evidence that students have benefits by watching

~~E_1 = 23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19~~
~~E_2 : 24, 19, 22, 8, 20, 22, 20, 20, 23, 20, 17.~~
test at 99 level of confidence

E_1 : 23, 20, 19, 21, 18, 20, 18, 17, 23, 18, 19
E_2 : 24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 17
∴ $H_0 = e_1 = e_2$
$H_1 = e_1 < e_2$

PRACTICAL : 6

Q1)
> x = c(3366, 3337, 3361, 3410, 3316, 3357, 3348, 3356, 3376, 3382, 3377, 3355, 3408, 3391, 3390, 3424, 3383, 3374, 3384, 3374)

Test the hypothesis

① $H_0 = \mu = 3400$, $H_1 = \mu \neq 3400$
② $H_0 = \mu = 3400$, $H_1 = \mu > 3400$ } at 95% LOC
③ $H_0 = \mu = 3400$, $H_1 = \mu < 3400$

Also check at 97% LOC

p. value = 0.987

t.test (x, mu = 3400 alter = "two sides", conf level = 0.95) p value = 0.000258

p.value > = 0.05

∴ Reject null hypothesis

t.test (x, mu = 3400, alter = "less", conf level = 0.95)

p.value = 0.0001264

t.test (x, mu = 3400, alter = "greater", conf level 0.95)

p.value = 0.0001264

t.test (x, mu = 3400 alter = "less", conf. level = 0.97)

Q2 Below are the data of rain in weights onto diff with diets A & B.

list A : 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25

list B : 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21

∴ $H_0 = a - b = 0$

$H_1 = a - b \neq 0$

95 percent confidence interval
-1.n 207330 7.933997
Sample estimates
Mean of the differences
- 3.166667

∴ Accept H_0, Reject H_1
∴ There is no difference in weighten.

11 students gave the test alts 1 month they of gave the test after the tuitions do the marks gives evidence that student have benefits by coaching

E_1 : 23, 20, 19, 21, 18, 20, 16, 17, 23, 16, 19
E_2 : 24, 19, 22, 18, 22, 20, 23, 20, 17
test at 99 level of confidnce.

→ H_0: $d_1 = d_2$
H_1 = $d_1 \neq d_2$
> d1 = c(0.7, -1.01, -0.2, -1.2, -0.1, 3.4, 3.7, 0.6, 2.01)
> d2 = c(1.9, 0.8, 1.1, 0.1, - 0.1, 4.4, 5.5, 1.6, 4.6, 3.4)
> t.test (d1, d2, alter = "two sided", conf. level = paired=t.test
data = d1 and d2
t = -4.0621, df = 9, p-value = 0.002833
alternative hypothesis : true difference in means is not equal to 0
95 percent confidence interval.
Mean of difference -1.58

∴ Reject H_0
Accept H_1

Q.5 If there is difference in salaries for the same in 2 different countries

~~cf~~ < A = 53000, 44956, 41974, 4436, 40570, 31913

< B = 62490, 58850, 49445, 52263, 47674, 43552

⇒ $H_0 = S_1 = S_2$

∴ $H_1 = S_1 \neq S_2$

> cA = c(53000, 44958, 41914, 344366, 40570, 33963)

> cB = c(62440, 88850, 44493, 52263, 42674, 43552)

Paired T.test

data : ca and cb

t = -4.4564, df = 5, p.value = 0.00666

alternative hypothesis : true difference in mean is not equal to 0

95 percent confidence interval

-10304.821 -2792.846

sample estimates

Mean of the difference. -6548.833

∴ Reject ~~the~~ H_0

Accept H_1

> val . test (x, y)

f test to compare two variables

p-value = 0.02756

∴ Reject H_0

∴ equality of 2 population mean are not same

> t. test (x, y, val.equal = F, paired - F)

which Two sample t - test

p-value = 0.3243

∴ Accept H_0

∴ Mean of two population is same.

Q.5 Prepare a csv file in excel import the in R and apply the test to check the equality of variance of 2 data.

Observed 1 : 10, 12, 17, 12, 16, 20

Observed 2 : 15, 14, 16, 11, 12, 19

$\therefore H_0 = \sigma^2_1 = \sigma^2_2$

$H_1 - \sigma^2_1 \neq \sigma^2_2$

Save the above observation : extract file in csv (ms - os) format

> data = read csv (file . choose(), header = T)

> data

∴ Accept H_0
∴ equality of 2 population mean are same.

Q equality of proportion variance
> var.test (x, y)
F. test to compare two variances
P - value = 0.7751
∴ Accept H_0
∴ equality of proportion variance are same.

The following are the price of commodity in there sample of shops selected at random from different city.

City A = 74.10, 77.70, 75.35, 74, 73.60, 79.30, 75.80, 76.80, 77.10, 76.40

City B = 70.80, 74.90, 76.20, 72.80, 78.10, 74.70, 69.80, 81.20

∴ $H_0 = \sigma_1^2 = \sigma_2^2$
$H_1 = \sigma_1^2 \neq \sigma_2^2$

> x = c (74.80, 77.10, 75.36, 74, 73.80, 79.30, 75.80, 79.30, 6.80, 77.10, 76.40)
> y = c (70.80, 74.90, 76.20, 72.80, 78.10, 74.70, 69.34, 81.20)

PRACTICAL 7

Title : F Test

1. Life exceptancy in 10 region of India in 1990 and 2000 all given below test whather the variance at the 2 time are same

1990 37, 39, 36, 42, 45, 44, 46, 49, 50, 51
2000 44, 45, 47, 43, 42, 49, 50, 41, 48, 52, 42, 59

```
> x = c (37, 39, 36, 42, 45, 44, 46, 49, 50, 51)
> y = c (44, 43, 47, 43, 42, 49, 50, 41, 48, 51, 42, 59)
  var. test (x, y)
    F test to compare two variances data : x
    and y :
F = 1.0548 ; num df = 9 , dnom df = 11 , p. value = 0.9176
alternative hypothesis : true : ratio of variance is
not equal to 95 percent confidence interval
  0. 2939977 . 4.1265887
  sample estimates
  ratio of variance
       1.083834
```

∴ Accept H_0
∴ Variance at 2 times are same.

2. I 25 28 26 22 29 31 31 26 31
II 30 23 31 32 23 25 36 26 31 82 27 37
38, 21

13

at 95% of confidence - level check the ratio of two population of variance.

> > x = c ∴ $H_0 = \sigma_1^2 = \sigma_2^2$

$H_1 = \sigma_1^2 \neq \sigma_2^2$

```
> x = c(25, 28, 26, 22, 29, 31, 37, 26, 31)
> y = c(30, 28, 31, 32, 23, 25, 36, 25, 31, 32, 32, 29,
        31, 35, 24)
> var.test(x, y)
```

F test to compare two variance

P - value = 0.4535

∴ Accept the H_0

∴ Variance at I and II are same

Q. For the following data test the hypothesis to

(1) equality of 2 population mean → t.test

(2) equality of proportion variance → f.test

Sample 1 : 175, 168, 145, 190, 181, 185, 175, 200

Sample 2 : 180, 170, 153, 180, 179, 183, 187, 205

→ (1) $H_0 = \mu_1 = \mu_2$

$H_1 = \mu_1 \neq \mu_2$

```
> x = c(175, 163, 145, 190, 181, 185, 175, 200)
> y = c(180, 170, 153, 180, 179, 183, 187, 205)
> t.test(x, y, alter = "two sided", conf.level = 0.95)
```

welch two sample t.test

P. value = 0.977,

x	O.B 1	OB 2
1	10	15
2	15	14
3	17	16
4	11	11
5	16	12
6	20	14

```
> attach(data)
> var.test (OB.1 , OB2 )
```

F test to compare two variance
p-value = 0.5717

∴ Accept H_0
∴ the variance of 2 data are same.

> H_0 = Median > 20
> H_1 = Median < 20
n = c (value = ...)
wilcox . test (x, alternative , = "less")
p value = 0.999
Accept = H_0

the next in legs of the person before after they stop smoking are as follows 65, 70, 70, 65, 72. After 72, 82, 72, 66, 73. Use wilcox test to check whether this the net of person there as after smoking use.

> H_0 = increased after this stopping of smoking
> H_1 = does not increases after of smoking
x = c values ...)
> z = x - y
+ z
> wilcox . test (z, mu = 0)
∴ P value = 0.1756
Accept H_0.

The following data gives the weight of 40 students in random sample.
46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48, 63, 61, 66, 84, 50, 48, 49, 62, 47, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 54.
Use the sign test to test whether the median rate of population is 50 kg against alternative it is > 50 kg

$\therefore H_0 = \text{median} - 50$
$H_1 = \text{median} - > 50$

x = c(46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48.
63, 61, 66, 54, 50, 48, 49, 62, 47, 49, 47, 55, 59, 63,
53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 53,

```
> sp = length which (x > 50)
> sp
  4
> sn = length (which(x < 50)
> sn
  13
> n = sp + sn
> qbinom (0.05, n, 0.5)
  1n
```

$\therefore$ q binom > 1 n
$\therefore$ Reject H_0

3) The median age of tourists visiting a certain place is claim to be 41 yrs A random sample of 20 tourists have the age.
25, 29, 52, 48, 57, 34, 45, 36, 30, 49, 28, 39, 44, 63, 32, 65, 42. Use the sign test to check the claim.

$\therefore$ H_0 : median $= H_1$
H_1 : Median $= H_1$

```
> x = c(25, 29, 52, 48, 57, 39, 45, 36, 30, 49, 21, 39,
   44, 63, 32, 65, 42)
> sp = length(which(x>4))
> sp
  9
> sn = length(which(x>41))
> sn
  8
> n = sp + sn
> qbinom(0.05, n, 0.5)
  5
```

$\therefore$ qbinom < sn
$\therefore$ Accept H_0
$\therefore$ median $= H_1$

Q.4 The times in minimum that a patient has to wait to consultation in recommended as following
24, 25, 20, 21, 32, 28, 12, 25, 24, 26
We will use sign test to check whether median waiting test more than 20 at 5%. is =

·a

PRACTICAL : 8

① The times of failure in hrs at 10 randomly selected 9 volt battery of a certain company is as follows
(28.9, 18.4, 28.7, 72.5, 48.6, 52.4, 37.6, 47.5, 62.1, 54.5)

Test the hypothesis that the population on Median is 63 against alternative is than 63 at 5% of level of significance.

∴ H_0 = Median = 63

H_1 = Median < 63

x = c (28.9, 18.2, 28.7, 72.5, 38.6, 52.4, 37.6, 47, 62, 1, 54.5)

> s P = length (which (x > 63))

> s A = length (which (x < 63))

> s P

→ x

> s n

1

> x = sP + sn

→ q binom (0.08, n, 0.5)

2

∴ qbinom < sn

Accept H_0

Median = 63

15

m
e = stack(m)
e
oneway.test(values ~ ind, data = e)
p.value = 0.03822
∴ Reject H_0

Q.4] An experiment was conducted on 8 persons & the observations were noted. Test the hypothesis that all groups have equal results on their reasons.

H_0 = equal results on their health.
H_1 = Not equal results
a = c(23, 26, 51, 48, 58, 37, 29, 44)
b = c(22, 27, 29, 39, 46, 48, 49, 65)
c = c(59, 66, 38, 49, 56, 60, 56, 62)
d = data.frame(a, b, c)
d
e = as.matrix(d)
e = stack(d)
aov(values ~ ind, data = e)
oneway.test(values ~ ind, data = e)
p.value = 0.01633
∴ Reject H_0

PRACTICAL : 9

Topic : Anova

Q] The following data gives the effect of 3 treatments

T_1 = 2, 3, 7, 2, 6

T_2 = 10, 8, 7, 5, 10

T_3 = 10, 13, 14, 13, 15

Test the hypothesis that all treatments are equally effective.

$H_0 = T_1 = T_2 = T_3$

$H_1 = T_1 \neq T_2 \neq T_3$

```
a = c (2, 3, 7, 2, 6)
b = c (10, 8, 7, 5, 10)
c = c (10, 13, 14, 13, 15)
d = data. frame (a, b, c)
d
x = as.matrix (d)
e = stack (d)
aov (value ~ T x d, dt = e)
oneway.test (values ~ ind, data = e)
```

p - value = 0.0006232

∴ Reject H_0.

The life cycle of different brands of types is given.
Test whether the driving life of all the types is same.

H_0 : life of all brands of tyres is same.
H_1 = life of all brands of tyres is not same

```
a = c (20,23,18,17,22,24)
b = c (17,15,17,20,16,17)
c = c (21,29,22,17,20)
d = c (15,14,16,18,14,16)
n
e = stack(m)
m = list (p = a, q = b, r = c, s = d)
n
e = stack(m)
e
oneway.list (value ~ ind, data = e, var.equal = TRUE)
```

p-value = 0.004056

Reject H_0

Three types of wax is applied for the protection of cars and no. of days of protections were noted. Test whether these are equally effective.

H_0 = Equally effective

H_1 = Not equally effective

```
a = c (44,45,46,47,48,49)
b = c (40,42,51,52,55)
c = c (50,53,58,59)
n = list (p = a, q = b, r = c)
```