PRACTICAL-01

TITLE : Random Variable.

Q1. Find the mean and variable 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	0.16	0.64
TOTAL	$\Sigma = 1$	$\Sigma = 1$	$\Sigma E(X)^2 = 0.20$	$\Sigma [E(X)]^2 = 0.74$

$\therefore$ 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·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 $V(X) = 83/64$.

Scanned with CamScanner

(c)

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

Solution:

X	$P(X)$	$X \cdot P(X)$	$E(X)^2$	$[E(X)]^2$
-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 $= E(X) = \Sigma X \cdot P(X) = 6.05$

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

$= 94.85 - 27.7525$

$= 67.095$

Mean $E(X) = 6.05$ & variance $V(X) = 67.097$.

Q2. If $P(X)$ is pmf of a random ~~raw~~ variable X

If $P(X)$ represents pmf for r.v. X, find value of K. Then evaluate mean & variance.

Solution: As $P(X_i)$ is a pmf it should satisfy the properties of pmf which are

(a) $P(X_i) \geq 0$ for all sample space.

(b) $\Sigma P(X_i) = 1$

Scanned with CamScanner

X	-1	0	1	2
P(X)	K+1/13	K/13	1/13	K-4/13

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

$\therefore 1 = \frac{K+1+K+1+K-4}{13}$

$\therefore 13 = 3K-2$

$\therefore 15 = 3K$

$\therefore \boxed{K=5}$

X	P(X)	X·P(X)	E(X)²	[E(X)]²
-1	6/13	-6/13	6/13	36/189
0	5/13	0	0	0
1	1/13	1/13	1/13	1/189
2	1/13	2/13	4/13	4/189
TOTAL	Σ=1	Σ=-3/13	Σ=11/13	Σ=41/189

$\therefore$ Mean $= E(X) = \Sigma X \cdot P(X) = -3/13$

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

$= \frac{11}{13} - \frac{41}{189}$

$= \frac{143-41}{169}$

$= \frac{102}{169}$.

$\therefore$ Mean $= -3/31$ & variance $= 102/169$.

Q3. The pmf of random variable 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 ① $P(-1 \le X < 2)$, ② $P(1 \le x \le 5)$, ③ $P(X \le 2)$, ④ $P(X \ge 0)$

Solution:

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	1.0

① $P(-1 \le x \le 2) = P(X \le 2) - P(x \le -1) + P(x = -1)$

$= F(x_b) - F(x_a) + p(a)$

$= F(2) - F(-1) + P(-1)$

$= 0.75 - 0.3 + 0.2$

$= 0.25$

② $P(1 \le x \le 5) = F(x_b) - F(x_a + P(a)$

$= F(5) - F(-1) + P(1)$

$= 0.95 - 0.65 + 0.2$

$= 0.15$

③ $P(x \le 2) = P(X=1) + P(X=-1) + P(X=3) + P(X=0) + P(X=2)$

$= 0.1 + 0.2 + 0.15 + 0.2 + 0.1$

$= 0.75$

43

(7) $P(x \geq 0) = 1 - F(0) + P(0)$
$= 1 - 0.45 + 0.15$
$= 0.40$

(4) Let f be continuous random variable with pdf : $f(x) = \frac{x+1}{2}; -1 < x < 1$

Obtain cdf of x. Find mean :-

Solution : By definition of cdf, we have :

$F(X) = \int_{-1}^{x+1} t\,dt$
$= \int_{-1}^{1} \frac{x+1}{2} \cdot dx$
$= \frac{1}{2}\left(\frac{1}{2}x^2 + x\right)$ for $-1 \leq x \leq 1$

Hence the cdf is
$F(X) = 0$ for $x \leq 0$ -1.
$= \frac{1}{4}x^2 + \frac{1}{2}x$; $-1 \leq x \leq 1$
$= 0$; $x \geq 1$

5. Let f be c.r.v. with pdf

$$f(x) = \frac{x+2}{18} \quad ; -2 \le x \le 4.$$
$$= 0. \quad ; \text{otherwise}$$

calculate cdf.

Solution: By definition of cdf, we have,

$$F(X) = \int_{-2}^{4} t \cdot dt.$$

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

$$= \frac{1}{18} \left(\frac{1}{2} x^2 + 2x\right) \text{ for } -2 \le x \le 4$$

Hence cdf is

$$F(X) = 0 \quad \text{for } x < -2$$
$$= \frac{1}{18} \left(\frac{1}{2} x^2 + 2x\right), \text{ for } -2 < x < 4$$
$$= 0 \quad , \text{ for } x \ge 4.$$

Neeh

PRACTICAL-02

TITLE: Binomial Distribution :

Q1. An unbiased coin is tossed 4 times calculate probability of obtaining no head, atleast one head & more than one tail.

NO HEAD :

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

AT LEAST ONE HEAD :

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

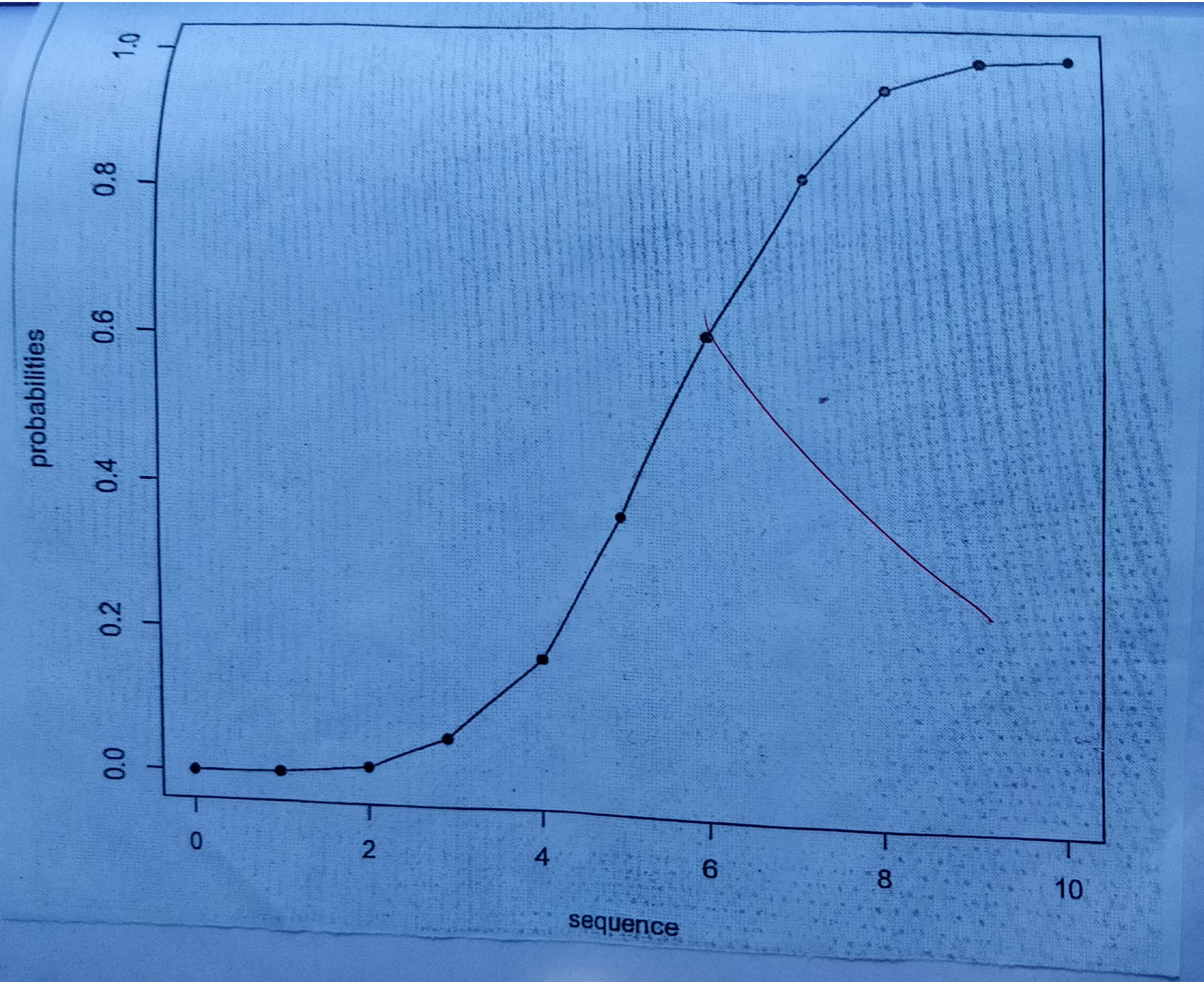
MORE THAN ONE TAIL :

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

Q2. If probability of student being accepted is 0.3, 5 students supply, what's the probability of almost 2 are accepted :

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

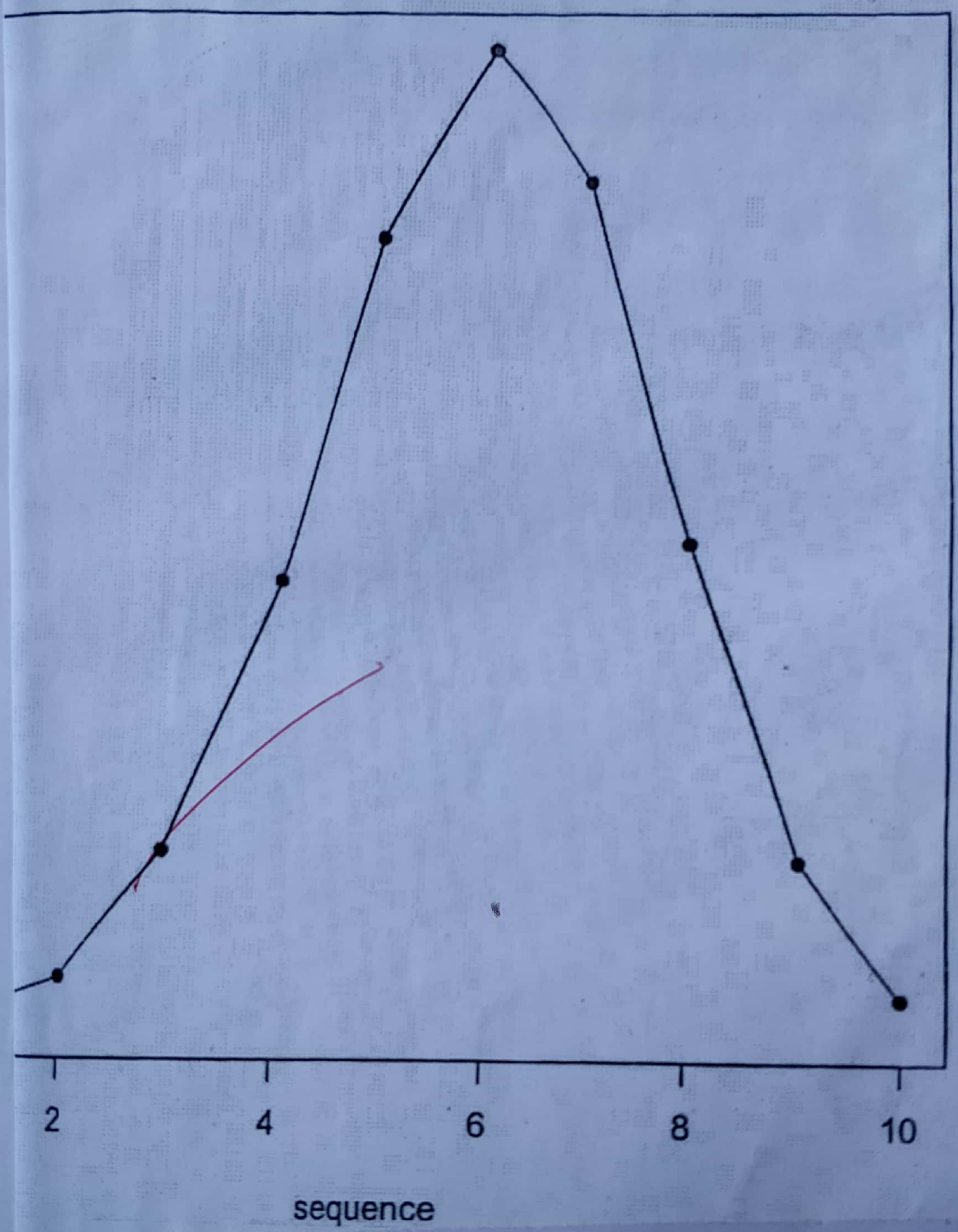
Q3. An unbiased coin is tossed 6 times, probability of head at any toss = 0.3. Let 'x' be no. of head that comes up. Calculate $P(X=2)$, $P(X=3)$, $P(1<x<5)$


probabilities
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
sequence

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

For n=10, p=0.6, evaluate binomial probabilities and plot graphs of pmf & 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
0.2403107840 0.0060466176 0.1209323520
>plot (x,y,xlab="sequence", ylab="probabilities", "o", pch=16)


2
4
6
8
10
sequence

Q7. Bits are sent for communication channel in packet of 12. If probability of bit being corrupted is 0.1. What is probability of no more than 2 bits are corrupted in a packet?

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

PRACTICAL-03

TITLE : Normal Distribution

A normal distribution of 100 students with mean = 40, SD = 15. Find no. of students whose marks are ① $P(X<30)$, ② $P(40<x<70)$, ③ $P(25<x<35)$, ④ $P(x>60)$

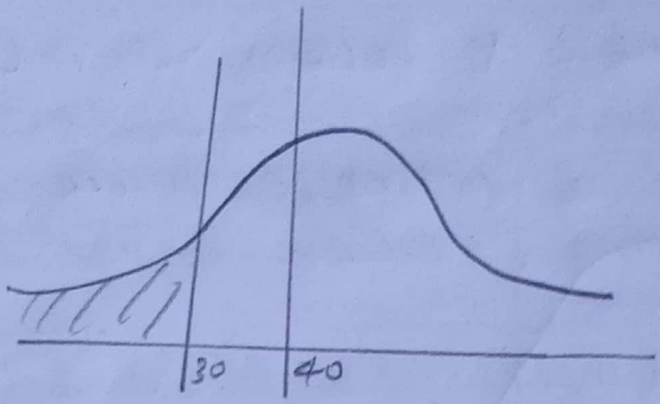
```
> pnorm (30,40,15)
[1] 0.2524925

> pnorm (70,40,15)-pnorm(40,40,15)
[1]0.4772499

>pnorm (35,40,15)-pnorm (25,40,15)
[1] 0.2107861

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

Q2. If random variable x follows normal dist^n with mean = 50, σ = 100. Find ① $P(x \leq 70)$, ② $P(x>65)$, ③ $P(x<32)$, ④ $P(35<x<60)$ ⑤ $P(20<x<30)$

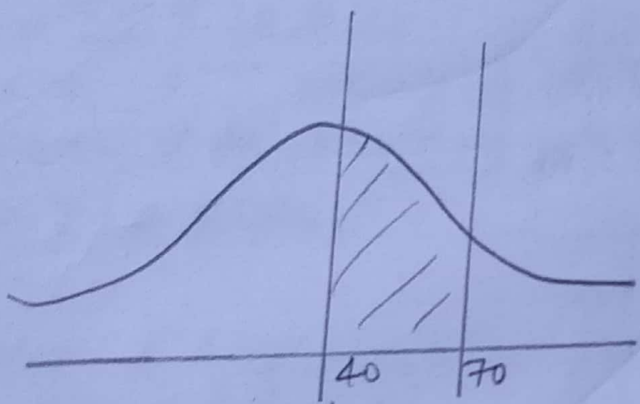
```
> pnorm (70,50,10)
[1]0.9772499
```

Q1.

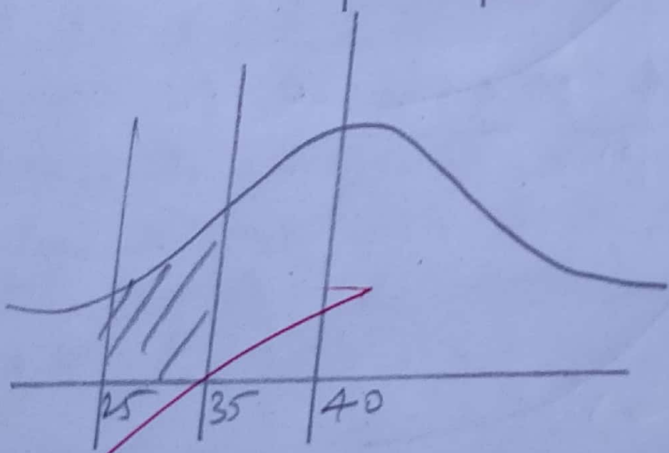
(a)



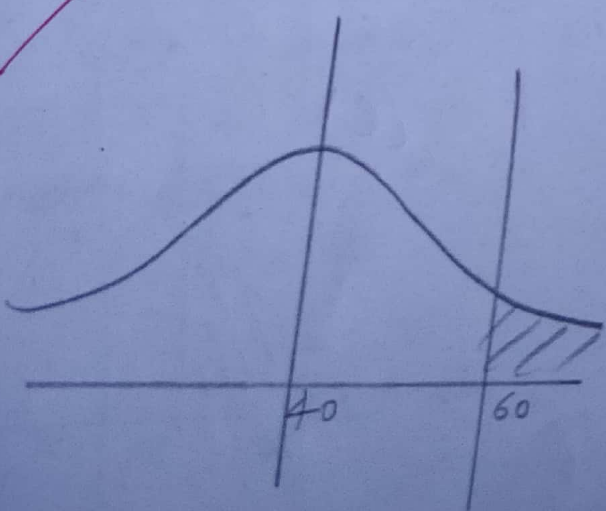
(b)



(c)



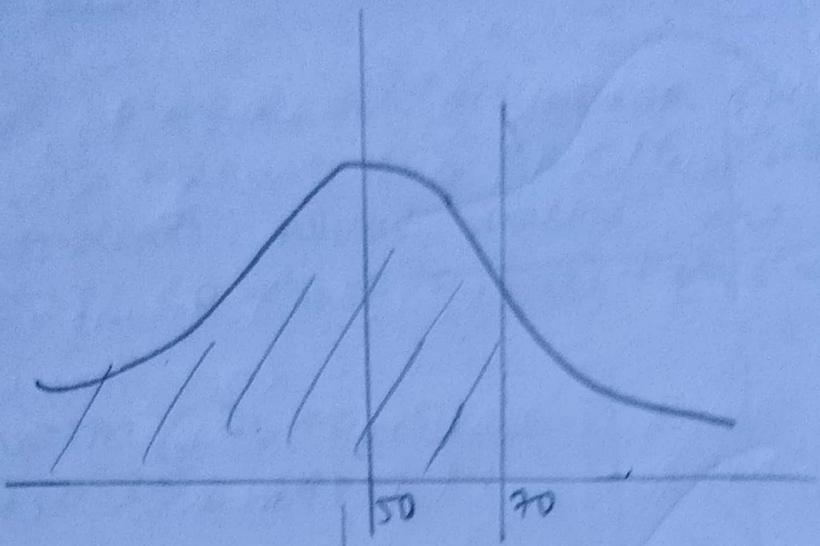
(d)



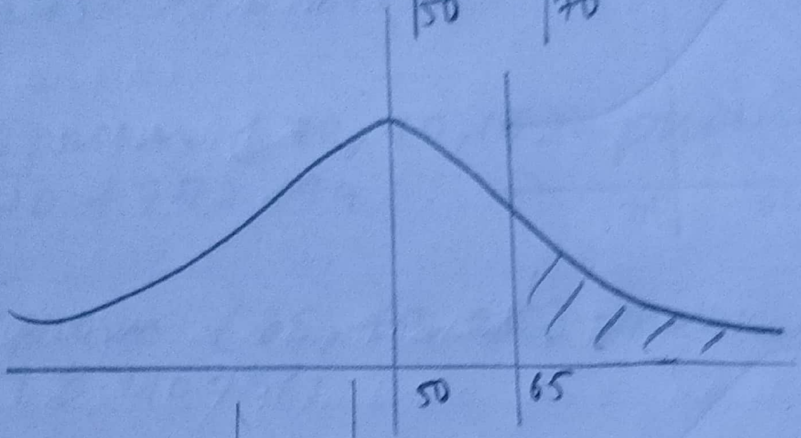
1/2

Q2.

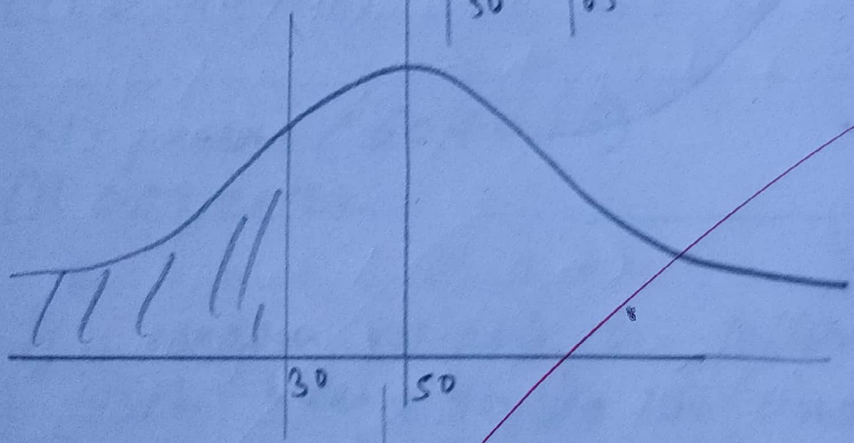
(a)



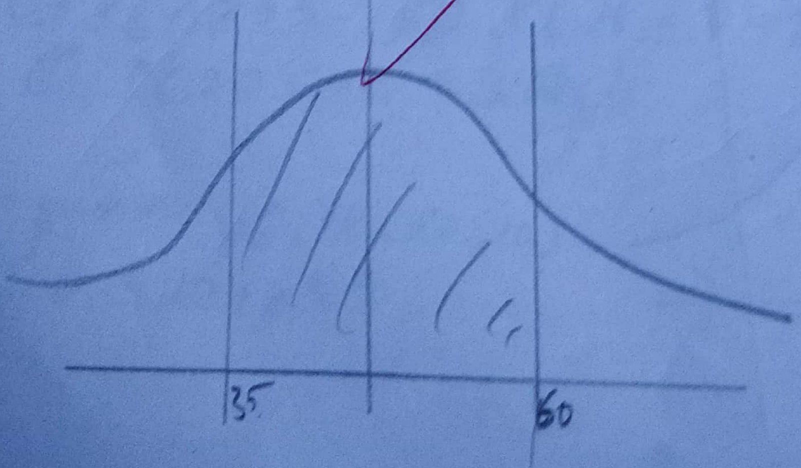
(b)



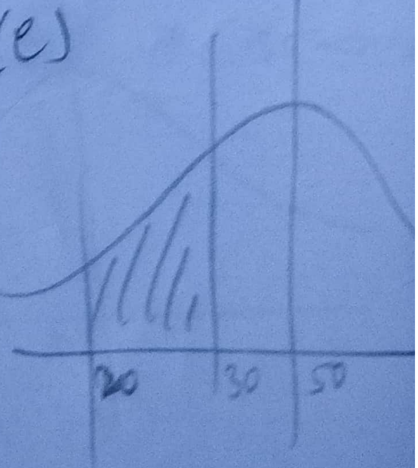
(c)



(d)



(e)



```
> 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 $P(X<K_1) = 0.6$ & $P(X>K_2) = 0.4$

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

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

Q4. A random variable X follows normal distribution with $\mu = 10$, $\sigma = 2$. Generate 100 observations & ~~vari~~ evaluate its mean & variance, median.

```
> x = rnorm (100, 10, 2)
> summary (x)
```

[1] Min	1st Q	Median	Mean	3rd Q	Max
5.713	8.444	9.723	9.914	11.325	14.238

```
> Var (x)
[1] 3.648924
```

Q5. Write a command to generate 10 random variable/ numbers for normally distribution 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

PRACTICAL-04

TOPIC : TESTING OF HYPOTHESIS.

case I : Sample mean and standard derivation given single population.

Q1. Suppose the food label on cookies bag states that it has atmost 2g of saturated fats in a single cookie. In sample of 55 cookies, it was found that mean and of saturated fat per cookies is 2.1g. Assume that sample std. significance can at 15%. level of significance can be rejected the claim of food.

→ $\sigma = 0.3$, $\bar{x} = 2.1$

$n = 35$, $\mu = 2$

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

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

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{2.1 - 2}{0.3/\sqrt{3}}$$

$$= \boxed{1.972027}$$

pvalue = 1 − pnorm (z)
= $\boxed{0.0243}$

∴ Reject the null hypothesis .

∴ pvalue < 0.05
Accept alternate hypothesis .

Q2. A sample of 100 customers was randomly it was found that avg. spending was 275 /– . The S.D. = 30 using 0.05 level of significance could you conclude that the amt. 250 /. whereas restaurant claims that it is not 250 /–

→ $\bar{x} = 275$; $\mu = 250$
$H_0 = \mu < 250$
$H_1 = \mu > 250$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{275 - 250}{30/\sqrt{100}}$$

$= \boxed{8.333}$

pt. (2.99 , lowertail = F)
∴ $\boxed{\text{pvalue} = 2.305736e\text{-}15}$

Reject the null hypothesis.

$\because$ pvalue < 0.05.

$\therefore$ Accept alternative hypothesis.

$\{\because \mu > 250\}$

33- A quality control engineer finds that sample of 100 bulbs have avg. life of 470 hours. Assuming population test whether the population mean is 480 hours v/s population mean < 480. at 1%

$\rightarrow$ $\mu = 100$, $\mu < 480$, $\sigma = 25$, $\mu = 480$, $\bar{x} = 470$.

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

pt. (z, 99, lower.tail = T).

$= \boxed{6.112576e.05}$

$\therefore$ Reject the null hypothesis

$\therefore p < 0.05$

$\therefore$ Accept the alternative hypothesis ($\mu_1 < 480$)

Q4. A principal at school claims that IQ is 100 of students. A random sample of 30 students whose IQ was found to be 112. The S.D. of population is 15. Test the claim of principal.

→ Method - 1 :

$H_0 = \mu = 100$
$H_1 = \mu < 100.$

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

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{112-100}{15/\sqrt{30}}$$

$$= \boxed{4.38178}$$

$\therefore$ pvalue $= \boxed{5.88567e\cdot06}$

$\therefore$ Reject the null hypothesis = claims of principal $[\mu = 100]$.

Scanned with CamScanner

Method - 2 :

$H_0 = \mu = 100$
$H_1 = \mu = 100$

$\therefore$ P value = 2 (1- pnorm (abs(z)))
$= \boxed{1.177134 5e.05}$

$\therefore$ Reject the null hypothesis.

pvalue < 0.05.

* Single Population Hypothesis Proportio

It is believed that coin is fair. The coin is tossed 40 times, 25 times head. Indicate whether the coin is fair or not at 95% LOC.

$\rightarrow$ $Z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$.

$p_0 = 0.5$, $q_0 = 1 - 0.5 = 0.5$

$p = \frac{28}{40} = 0.7$, $n = 90$.

$\therefore z = \dfrac{0.7-0.5}{\sqrt{\dfrac{0.5\times0.5}{40}}}$

$H_0 = \mu = 0.5$

$H_1 = \mu \neq 0.5$

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. To confirm that male & female are equal in —

$Z = \dfrac{p-p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$; $p = \dfrac{520}{1000} = 0.52$, $p_0 = 0.5$, $q_0 = 0.$

$n = 1000$, $H_0 = [p = p_0]$, $H_1 = [p \neq p_0]$

z = (p-p0) (sqrt(abs(z)))

z = 1.2645.

pvalue = 2×[1-pnorm(abs(z))]

pvalue = 0.2060506

∴ Reject null hypothesis.

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.

$\rightarrow$ $$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$; $p \rightarrow 325/600 = 0.541607$

$p_0 = 0.5$, $q_0 = 0.5$, $n = 600$, $H_0 = [p_0 = p]$
$H_1 = [p \neq p_0]$

Z = (0.5416 - 0.5) / (sqrt (0.5 * (0.5/600))
Z = 2.037975

pvalue = 0.04155239.

Reject the null hypothesis $\because$ pvalue < 0.5.

9. Experience shows that 20% of manufactured products are of top quality. In 1 day production of 400 articles only 50 are of top quality. Test hypothesis that experience product is wrong?

$$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$, $p = 0.125$ (50/400) , $p_0 = 0.2$, $q_0 = 0.8$,

$n = 400$

$H_0 = [p = 0.2]$
$H_1 = [p \neq 0.2]$

z = (0.125 - 0.2) / (sqrt(0.2 * 0.8 / 400))
z = -3.75
pvalue = 2 x (1 - pnorm (abs(z)))
pvalue = 0.001768346.

∴ Reject the null hypothesis ∵ pvalue < 0.2.
Accept null hypothesis i.e. $p \neq 0.2$.

Q1. In an election campaign, a telephone of 800 registered voters shows favour 460. second note opinion 520 of 100 registered voters favoured the condition of 0.5%. LOC is there sufficient evidence that popularity has decreased.

→ n = 800, p_1 = 460/800 = 0.575, m = 100, p_2 = 520/100 = 0.52
p = (0.575 * 800 + 0.52 x 1000)/(180)
p = 0.54444.

$z = sqrt(0.54 * 0.46) + 1/8$

$z = 0.00112139+$

$H_0 = p = 0.544$

$H_1 = p < 0.544$

$pvalue = 2 \times (1 - pnorm(abs(z)))$

$\therefore pvalue = 0.9991053$

$\therefore$ Accept the null hypothesis $\because$ pvalue > 0.5.

$\therefore$ Accept $p = 0.5444$.

P2. From a consignment A, 100 articles are drawn & 44 were found defectives from consignment B, 200, samples are drawn out of which 30 are defective. Test whether proportion of defective items in 2 consignments are significantly different.

$\rightarrow H_0 = p_1 = p_2$

$H_1 = p_1 \neq p_2$

$p2 = 30/200 = 0.15$

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

19

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

z = (sqrt (0.185 * 0.815) * (2/200))
z = 0.00388 2976.

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

∴ pvalue = 0.996 9018.
pvalue > p.

∴ Accept the null hypothesis i.e p1 = p2.

Neeta

PRACTICAL-05

CHI SQUARE

Q1. Use the foll^g. 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
     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.

∴ Reject null hypothesis.
Both are dependent.

Q2. A disc is tossed 120 times & follg. results are obtained:

No. of terms	Freq.
1	30
2	25
3	18
4	10
5	22
6	15

Test ~~the~~ 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.

Q3 IQ test was conducted of 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 test after training.

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

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

[1] 0.1135950.

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

Q4. Online face   Graduate
to face

20

40

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

∴ H_0 = Independent H_1 =Dependent
> z = matrix (x, nrow=2)
> chisq.test (z)

Pearson's chi-squared test with Yates continuity correction.

Scanned with CamScanner

data = z

x-squared = 18.05, df = 1, p-values = 2.15 e^{-05}

∴ Reject null hypothesis.
∴ Both are dependent.

Q5. A dice is tossed 180 itemes

Use of terms	frequency.
1	20
2	30
3	35
4	40
5	12
6	43

Test the hypothesis that dice is unbiased

Ho = dice is biased
H_1 = dice is unbiased.

```
> x = c(20,30,35,40,12,43)
> chisq.test(x)
```

Chi-squared test for given probability.

data = z

x-squared = 23.933, df = 5, pvalue ~~000~~ = 0.000223.

∴ Reject null hypothesis

Dice is unbiased.

PRACTICAL-06

Topic - T-test

Q. Let x = 3366, 3337, 3361, 3410, 3316, 3357, 3348, 3356, 3376, 3382, 3377, 3355, 3408, 3401, 3398, 3424, 3388, 3374, 3384, 3374.

Write test hypothesis.

① $H_0 : \mu = 3400$, $H_1 : \mu \neq 3400$
② $H_0 : \mu = 3400$, $H_1 : \mu > 3400$
③ $H_0 : \mu = 3400$, $H_1 : \mu < 3400$.
at 95% level of confidence. Also, check at 97% level of confidence.

→ ① $H_0 : \mu = 3400$
$H_1 : \mu \neq 340$

```
> x = c(3366, 3337, 3361, 3410, 3316, 3357, 3348,
  3356, 3376, 3382, 3377, 3355, 3408, 3401,
  3398, 3424, 3383, 3374, 3384, 3374).
> t-test(x, mu = 3400, alter = "two.sided",
  conf.level = 0.95).
```

data : x
t = -4.4865, df = 19, pvalue = 0.0002528
alternative hypothesis : True mean is not equal to 3400 ..

95% confidence level:

3361.797   3386.103

Sample estimates:
mean of x:
3373.95

∴ Reject H_0.
∴ Accept H_1.

> t.test (x, mu = 3400, alter = "two-sided",
conf.level = 0.97)

One sample test:
data: x
t = -4.4865, df = 19, P-value = 0.000258
Sample estimates:
mean of x:
3373.95
∴ Reject H_0.
∴ Accept H_1.

→ ⓒ $H_0 = \mu = 3400$
$H_1 = \mu > 3400$.
> t.test (x, mu = 3400, alter = "greater",
conf.level = 0.95)
One sample t-test:

data: x

t = -4.4865, df = 19, P-value = 0.9999

alternative hypothesis: true mean is greater than 3400, 3363.91 Inf.

sample estimates:

mean of x:

3373.95

∴ Accept H_0.

(3) $H_0: \mu = 3400$

$H_1: \mu < 3400$

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

one sided t-test.

data: x

t = -4.4865, df = 19, pvalue = 0.0001264

alternative hypothesis: true mean is less than 3400.

-Inf = 3383.49.

Sample estimates:

mean of x:

3373.95:

∴ Reject H_0

∴ Accept H_1

15

>t.test (x, mu = 3400, alter = "less", conf.level = 0.99)

one sample t-test.

data : x
t = -4.4865 df = 19, p-value = 0.000126
alternative hypothesis : true-mean is less than 3400.

93 % level of confidence
-Inf 3385.563.
Sample estimates:
mean of x.
3378.95
∴ Reject H_0
∴ Accept H_1

Q2. Below all the data of gains in weighs. diets A and B.

Diet A: 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 32, 25
Diet B: 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21
→ ∴ $H_0 = a - b = 0$
∴ $H_1 = a - b \neq 0$
>a = c(25, 32, 30, 43, 24, 24, 14, 32, 24, 31, 31, 18, 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 = -0.62782, df = 11, p-value = 0.5429.
alternative hypothesis : true difference in means is equal to 0.

confidence interval
-14.267330 7.933997.

sample estimates :
mean of differences
-3.166667

∴ Accept Ho.
There is no difference in ~~congress~~. weighs.

Q3. Eleven student gave the test after 1 month they again gave the last after tuitions, do the marks gives evidence that students have benefitted by coaching.
E1 : 23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19.
E2 : 24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 17.
test at 99 level of confidence.

→ ∴ $H_0 : e_1 = e_2$
∴ $H_1 : e_1 < e_2$.
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: true difference in mean is less than 0.

99 % confidence interval :
-Inf 0.863333
sample estimates ;
mean of differences
-1

∴ Accept H_0.

Q4. Two drugs for BP was given and data was collected
D1: 0.7 , -1.6, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0.2
D2: 1.4, 0.8, 1.1, 0.1, 0.1, 4.4, 5.5, 1.6, 4.6, 3.4.

The 2 drugs have same effect, check whether 2 drugs have same ~~diff~~ effect on patient or not.

→ t.test (d_1 , d_2 , alter = "two sided", paired = T, conf.level = 0.95).

Paired t-test.
data : d1 & d2
t = -4.0621, df = 4, P-values = 0.002833
∴ Reject H_0
∴ Accept H_1

Q5. If there is difference in salaries for the same job in 2 different countries.

CA : 53000, 49958, 41974, 44366, 40470, 36963.

CB : 62990, 58850, 49475, 52263, 47674, 43552.

→ $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

> CA = c(53000, 49958, 41974, 44366, 40470, 36963)

> CB = c(62490, 58850, ~~[illegible]~~ 47495, 52263, 47674, 93552)

> t.test (CA, CB, paired = T, alter = "two sided", conf.level = 0.95)

paired t-test.

data : CA and CB.

t = -4.4569, df = 5, p-values = 0.00666,

alternative hypothesis = true difference in means is not equal to 0.

95 % confidence level :

-10404.821 -2792.846

Sample estimates :

mean of differences :

-6598.833

∴ Reject H_0.

∴ Accept H_1.

PRACTICAL-07

Topic: F-test.

Q1. Life expectancy in 10 regions of India in 1990 and 2000 are given below. Test whether variance at the two times are the same.

1990: 37, 39, 36, 42, 45, 44, 46, 49, 50, 51
2000: 44, 45, 47, 43, 42, 49, 50, 41, 48, 52, 58, 4~~2~~.

```
> x = c(37,39,36,42,45,44,46,49,50,51)
> y = c(44,45,47,43,42,49,50,41,48,52,42,5
> var.test(x,y)
    F test to compare two variances:
data : x and y
F = 1.1449, num df = 9, denom df = 11, pvalue = 0
alternative hypothesis : true ratio of variances is
not equal to 1
95 % confidence interval :
0.3191005      4.4789350
sample ~~differences~~ estimates:
ratio of variances :
        1.1449
```

Q2. for following data, test hypothesis for –

① Equatity of two population mean (→ T-test)

② Equality of two proportion variance (→ F test)

Sample 1: 175, 168, 145, 190, 181, 185, 175, 200
Sample 2: 180, 170, 153, 180, 179, 183, 187, 205.

→ x = c(175, 168, 145, 190, 181, 185, 175, 200)
> y = c(180, 170, 153, 180, 179, 183, 187, 205).
> var.test(x, y)

F test to compare two variances.

data: ~~x~~ and y.
F = 1.25, num df = 7, denom df = 7, p-value 0.7759
alternative hypothesis: true ratio of variances is not equal to 1.
0. 2502589   0.2437398
sample estimates:
ratio of variances
1.250021.

t.test ~~(...)~~ (x)

One sample t-test

data: 0
Sample estimates:
mean of x:
177.375.

Q3. The following are prices of commodities in sample of shop selected at random from different cities.

C-A: 74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.80, 76.80, 77.10, 76.40.
C-B: 70.80, 74.90, 76.20, 72.80, 78.10, 74.70, 69.80, 81.20

```
> a = c(74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.8
   76.80, 77.10, 76.40)
> b = c(70.80, 74.90, 76.20, 72.80, 78.10, 74.4
   69.80, 81.20)

> var.test(a, b)

        F-test to compare two variances.
data: a and b.
F = 0.22579, num df = 9, denom df = 7,
pvalue = 0.04249.
alternative hypothesis: true ratio of variance
  is not equal to 1
95% confidence level:
sample estimates:
ratio of variances:
              0.225791.
```

Q4. Prepare csv file in excel, import file in R & apply the test to check equality of variance of 2 datas.

Obs.1 → 10, 15, 17, 11, 16, 20
Obs 2 → 15, 14, 16, 11, 12, 19.

```
> Data = read.csv(file.choose(), header = TRUE)

> Data
   Obs.1  Obs 2
1   10     15
2   15     14
3   17     16
4   11     11
5   16     12
6   20     19

> attach(Data).
> var.test(obs1, obs2).
    F test to compare two variances
data: obs.1 and obs 2
F = 1.7068, num df = 5, denom df = 5, pvalue =
  0.5717.
alternative hypothesis: true ratio of variances
is not equal to 1
95 % confidence interval :
      0.238838      12.197640
sample estimates:
```

ratio of variances:
1.706827.

PRACTICAL-08

SIGN TEST.

Q1. The time of failures (in hours) of 10 randomly selected 9-Volt battery of a certain company is as follows —

28.9, 15.2, 28.7, 72.5, 48.6, 52.4, 37.6, 49.5, 62.1, 54.5.

Test hypothesis that population median is 63 against alternative that is less than 63 at 5% level of significance.

→ x = c(28.9, 15.2, 28.7, 72.5, 48.6, 52.4, 37.6, 49.5, 62.1, 54.5)

H_0 : median = 63

H_1 : median < 63.

y = which (x>63)

length(y)

[1] 4

z = which (x<63)

length(z)

[1] ~~4~~ 9

n→ Total negative and positive signs.

n = y+z

>qbinom(0.5, 10, 0.5)

[1] 2

∴ qbinom < z, Accept null hypothesis.

Q2. Following data gives weight of 40 students in random sample —

46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48, 65, 61, 66, 54, 50, 48, 49, 62, 47, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 54

Use sign test to test whether mean weight of population is 50 kg against alternative that its greater than 50 kg.

→ x = c(46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48, 65, 61, 66, 54, 50, 48, 49, 62, 47, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 54)

H_0 : mean weight = 50
H_1 : mean weight > 50

```
y = which (x > 50)
length (y)
[1] 25
z = which (x < 50)
length(z)
[1] 12
qbinom(0.05, 40, 0.5)
[1] 15.
```

∴ Reject null hypothesis.

Q5: median age of tourists is visiting a certain place is claimed to be 41 years. A random sample of 17 tourists have age —

48, 52, 25, 29, 57, 39, 45, 36, 30, 49, 28, 39, 44, 63, 32, 65, 42.

Use sign test to check claim.

→ x=c(48, 52, 25, 29, 57, 39, 45, 36, 30, 49, 28, 39, 44, 63, 32, 65, 42)

H_0: age = 41 years
H_1: age ≠ 41 years.

z = which (x > 41)
length(z)
[1] 9
j = which(x < 41)
length(j)
[1] 8
qbinom (0.05, 17, 0.5)
[1] 5.

∴ Accept null hypothesis.

Q4. The time in minutes that a patient has to wait for consultation is recorded as follows-

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

Use Wilcox sign test to check whether the median in time is greater than 20 at 5% level of significance.

→ H_0 : minutes > 20

H_1 : minutes < 20.

```
x = c( 15,17, 24,25, 20,21, 32, 28,12,25, 24,26)
  z = which(x>20)
  length(z)
 [1] 8
  j = which (x<20)
   length(j)
 [1] 3
wilcox.test (x, alternative = "greater")

[1]        Wilcoxon signed rank test with
   continuity correction.
data : x
V = 78 , p-value = 0.001253
alternative hypothesis : true location is greater
                              than 0
```

∴ Accept null hypothesis.

35. The weight in kgs of the person before and after stop smoking is —

Before	After
65, 75, 75, 72, 62	72, 72, 72, 66, 82.

Use wilcoxon test to check weight of a person is less after smoking at 5% level of significance.

→ x = c(65, 75, 75, 72, 62)
y = c(72, 72, 72, 62, 82)
z = x - y
z
[1] -7 0 3 -7 -4

wilcox.test (z, mu = 0, alternative = "less")
[1]
Wilcoxon signed rank test with continuity correction
data : z
V = 1 , p-value = 0.09873
alternative hypothesis : true location is less than 0.

∴ Reject null hypothesis

Neat

PRACTICAL - 09

* Topic :

Q1. The following data gives effect of 3 treatments

T_1	T_2	T_3
2	10	10
3	8	13
7	7	14
2	5	13
6	10	15

Test the hypothesis of all the treatements are equally effective.

→ t1 = c(2, 3, 7, 2, 6)
t2 = c(10, 8, 7, 5, 10)
t3 = c(10, 13, 14, 13, 15)
data = data.frame(~~t1~~, t2, t3)
data.

	t1	t2	t3
1	2	10	10
2	3	8	13
3	7	7	14
4	2	5	13
5	6	10	15

c = stack (data)
c

	values	ind
1	2	t1
2	3	t1
3	7	t1
4	2	t1
5	6	t1
6	10	t2
7	8	t2
8	7	t2
9	5	t2
10	10	t2
11	10	t3
12	13	t3
13	14	t3
14	13	t3
15	15	t3

aov (values ~ ind, e)

Call:

aov (formula = values ~ ind, e)

Terms:

	ind	Residuals
Sum of squares	203.3333	54.0000
Deg. of Freedom	2	12

Residual standard error : 2.12132

Estimated effects may be unbalanced

oneway.test (values ~ ind, e)

One-way analysis of means (not assuming equal variances)

data : values and ind

F = 21.537 , num df = 2.0000 , denom df = 7.9314

p-value = 0.0006232

∴ Reject NULL HYPOTHESIS.

Q2. The following gives life of types of 4 brands A, B, C, D;

A		B		C		D	
20	22	19	16	21	19	15	14
23	24	17	20	22	17	16	18
18	17	17	16	20		14	16

Test that average life of all brands are same.

```
a = c(20,23,18,17,22,24)
b = c(19,15,17,20,16,17)
c = c(21,19,22,17,20)
data = data.frame(a,b,c)
data
```

	a	b	c	d
	20	19	21	15
1	22	16	19	14
2	23	17	17	16
3	24	17	22	18
4	18	20	20	16
	17	16		14

e = list(a1 = a, b2 = b, c1 = c, d1 = d)

~~e~~
a
~~c~~
e

aov (values ~ ind , f)

call :

aov (formula = values ~ ind, data = f)

Terms :

	ind	Residuals
Sum of squares	97.46812	82.96667
Deg. of Freedom.	3	19

Residual standard error : 2.059657

Estimated effects may be damaged.

One way . test (values ~ ind , data = f)

p-value = 0.00629 1.

∴ Reject NULL HYPOTHESIS.

Q3. 3 types of wax is applied for protection of cars and number of days of protection per noted. Test whethere these are equally effective.

A	B	C
46, 46, 44, 45, 48, 49	40, 42, 51, 52, 55	50, 53, 58, 59.

→ a = c(44, 45, 46, 47, 48, 49)
b = c(40, 42, 51, 52, 55)
c = c(50, 53, 58, 59)
e = list(a1 = a, b1 = b, c1 = c)
d = stack(e)
aov(values ~ ind, data = d)
call:
aov(formula = values ~ ind = d)

Terms:	ind	Residuals.
Sum of squares	182.4333	245.5000
Deg. of Freedom	2	12

Residual standard error: 4.523069
Estimated effect may be unbalanced.
one way.test(values ~ ind, date = d)

One-way analysis of mean

data: values and ind.
F = 6.2398, num df = 2.0000, denom df = 5.4

p-value = 0·03822

∴ Reject NULL HYPOTHESIS.

Q4. An experiment was performed on 8 people & observations were noted. Test hypothesis for these three groups.

G1	G2	G3
26, 23, 51, 48, 58, 37, 39, 44	22, 27, 29, 39, 46, 48, 49, 65	59, 66, 38, 49, 56, 60, 56, 62.

Test hypothesis for all groups have equal results on their health.

→
```
a = c( 23, 26, 51, 48, 58, 37, 39, 44)
b = c( 27, 22, 28, 39, 46, 48, 49, 65)
c = c( 59, 66, 38, 49, 56, 60, 56, 62)
d = data.fame (a, b, c)
e = stack (d)
e
~~one way~~ aov (values ~ ind, data = d)
```

call:

aov (formula = values ~ ind = d)

Terms:	ind	Residuals
Sum of squares	172·4009	267·0691
Deg. of Freedom	2	12

one way.test (values ~ ind, data = e)

One way analysis of means

data : values and ind .

F = 5.3711 , num df = 2.00 , denom df = 13.41 ,

pvalue = 0.01935

∴ Reject NULL HYPOTHESIS.

Nerb