Page No.
Date

<BCS-040> <Dec-2022>

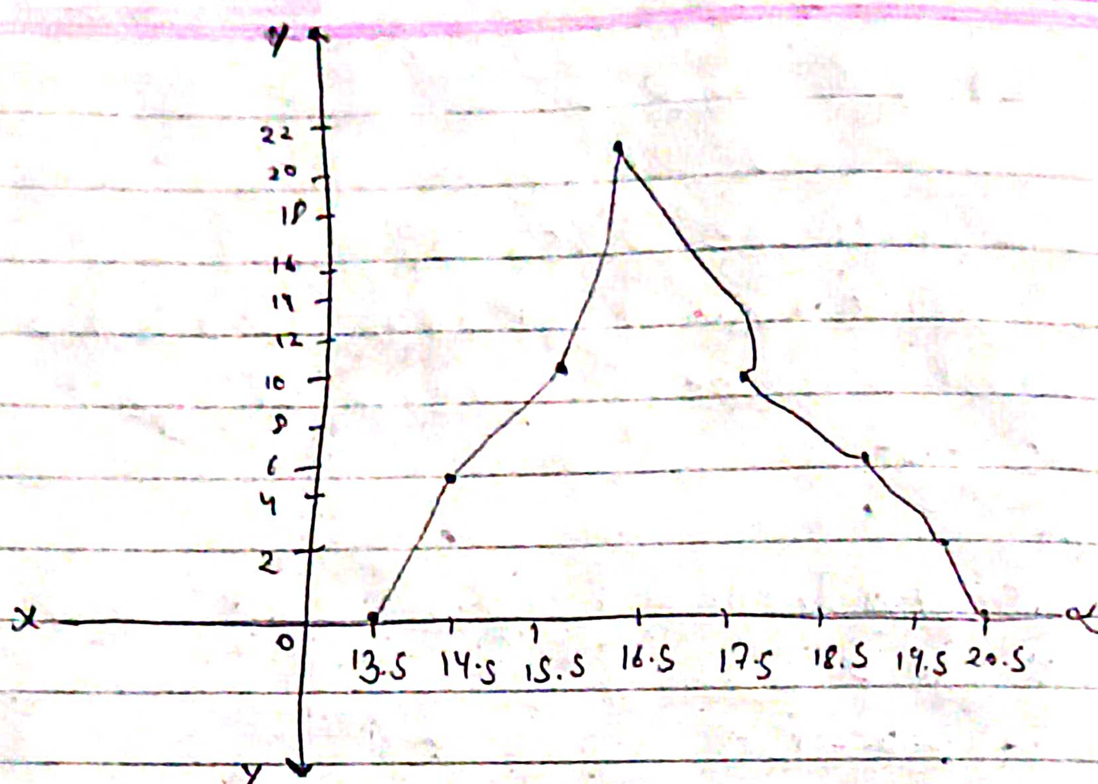
Q1. In a study on the per capita income for a particular year in a city, the following weekly observations were made:

Per capita Income (₹) (1k = 1000)	Number of weeks
14k - 15k	5
15k - 16k	10
16k - 17k	20
17k - 18k	9
18k - 19k	6
19k - 20k	2

Draw a histogram and frequency polygon on the same scale.

Ans

Per Cap. Inc.	Num. of weeks	Avg. cost
14k - 15k	5	14.5
15k - 16k	10	15.5
16k - 17k	20	16.5
17k - 18k	9	17.5
18k - 19k	6	18.5
19k - 20k	2	19.5



Scale

X, 1 cm = 10 ~~cost of~~ Per capital Income

Y, 1cm = 2 week

Q2. A problem of statistics techniques is given three student A, B, C whose changes of solving are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ ~~repea~~ respectively.

$\frac{1}{3} \alpha \frac{1}{2}$

$\frac{2 \alpha 3}{6}$

$\frac{6}{6}$

~~Q. select~~

Ay.

~~P(S)~~. $P(\bar{A} \cap \bar{B} \cap \bar{C}) = P(\bar{A})\, P(\bar{B})\, P(\bar{C})$

$\frac{1}{2} - \frac{1}{3}$

$\frac{3-2}{6}$

$\frac{1}{2} \times \frac{2}{3} \times \frac{3}{4}$

Page No.
Date:

i).
Ans

$P(S) = P(\bar{A} \cap \bar{B} \cap \bar{C})$

$= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4}$

$= \frac{1}{4}$ unsolve probeblity.

ii).

$P(S) = 1 - \frac{1}{4} = \frac{3}{4}$ Ans

solve probeblity

Q.3.
Ans

Class	Frequency	xi	fixi	C.F	
0-20	6	10	60	6	
20-40	8	30	240	14	
40-60	10 f_0	50	500	24	– C.F
60-80	f ← 12 high f_1	70	840	36	median class
80-100	6 f_2	90	540	42	
100-120	5	110	550	47	
120-140	3	130	390	50	

modal class

$\Sigma f_i = 50$ $\quad \Sigma x_i f_i = 3120$

i) mean $= \frac{\Sigma x_i f_i}{f_i} = \frac{3120}{50}$ 62.4 Ans

ii) mode $= l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$

Page No.
Date

(ii) Mode $= L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

$= 60 + \frac{12-10}{24-10-8} \times 20$

$= 60 + \frac{2}{8} \times 20 = 65$ Ans

$\boxed{\text{Mode} = 65}$

(iii) Median.

$L_1 + \frac{\frac{N}{2} - c.f}{f_1} \times i$

$= \frac{N}{2} = \frac{50}{2} = 25$

$= 60 + \frac{25 - 24}{12} \times 20$

$= 60 + \frac{1}{3} \times 5$

$= 60 + 1.6 = 61.5$ Ans

Q4. Box A contains 5 red and 4 blue balls, Box B contains 2 red 5 blue balls. A ball is drawn at random from each box. find the probability that one is red and the other is blue.

Ans. Probability formula :- ~~no. of~~

$$\frac{\text{No. of favourable outcomes}}{\text{Total. No. of outcomes}}$$

→ Probability of selecting a red ball from BOX A.

$P(R) = \frac{5}{9}$

~~Box~~ Blue ball Box A

$= P(b) = \frac{4}{9}$

→ Red ball from Box B

$P(R_2) = \frac{2}{7}$

blue

$P(b_2) = \frac{5}{7}$

→ Probability of selecting a red ball from Box A ~~and~~ and blue from Box B

$= \cancel{P(R_1 b_2) = \frac{5}{9} \times \frac{2}{7}}$

$= P(R_1 b_2) = \frac{5}{9} \times \frac{5}{7} = \frac{25}{63}$

→ Probability of selecting a blue ball from Box A and red ball from Box B.

$P(b R_2) = \frac{4}{9} \times \frac{2}{7} = \frac{8}{63}$

= Total probability

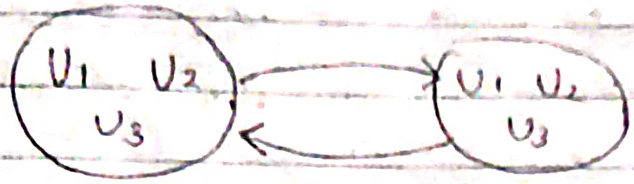
$= P(R_1 b_2) + P(b R_2)$

$= \frac{25}{63} + \frac{8}{63} = \frac{33}{63} = \boxed{\frac{11}{21}}$ Ans.

Q8. a). Compare simple random sampling with replacement and ~~simple~~ simple random sampling without replacement

Ans.

1). Simple Random sampling with Replacement.



Theorm:-

$P(U_1) = \frac{1}{N}$

$P(U_2) = \frac{1}{N}$

$P(U_N) = \frac{1}{N}$

2). Simple Random Sampling without Replacement.

Q5. Define time series and discuss various components of time series.

Ans. (The set of data collected on the basis of time is called time series (such as daily sales, weekly orders, monthly revenue, Yearly income).)

Q. This collection of Data are used to determine the ~~compny~~ company's growth.

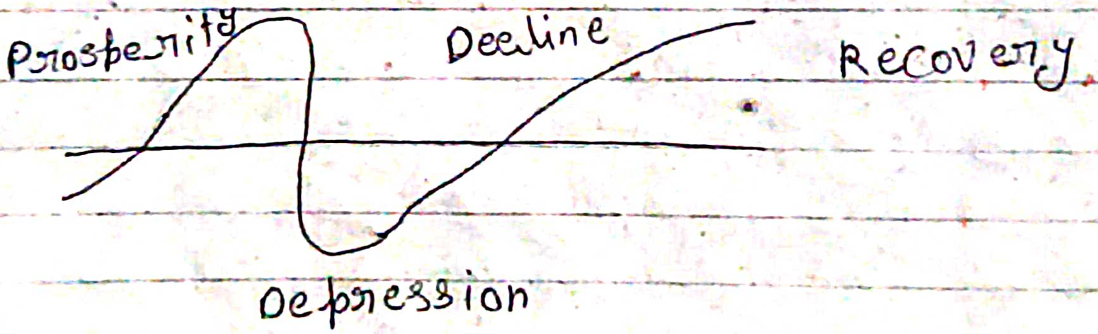
* Components *

•) Secular Trend :- long term Data (Grow) like in 1992 Sales in 200 and now in Present year sales is 200000.

•). Seasonal Variation :- Basically there are things in variation that is Depression Recovery.

Ex:- Ice cream sold in ~~se~~ summer not winter..

•). Cyclical variation :-



•). Irregular variation :- sometimes it happen through like natural calamity.

Page No.
Date:

(c). Write short notes on the folling.

i). **t-test** :- A t test is a statistical test that is used to compare the means of two groups.

- If you want to compare more than groups, or if you want to do multiple pairwise comparisons, use an anova test.

ii). **Properties of good estimator** :-

A good estimator is one which is as close to the true value of the parameter as possible. A good estimator possess the following properties or characteristics.

- **Estimators** :- Those sample statistics which are used to estimate the unknown population parameters are called estimators.

BCS-040 June 2022

Q1. Calculate the mean and standard deviation from the following Data:-

Marks	No. of students
0-10	10
10-20	9
20-30	25
30-40	30
40-50	16
50-60	10

Ans.

		m(x)	Σfx/Σfixi
0-10	10	5	50
10-20	9	15	135
20-30	25	25	625
30-40	30	35	1,050
40-50	16	45	720
50-60	10	55	550
Σfi =	100		3130

Σfixi

mean = $\frac{\sum f_i x_i}{\sum f_i}$ = $\frac{3130}{100}$ 31.

Q2. In order to find the correlation coefficient between two variables x and Y 20 pair of observations, the following calculation were made:

$$\Sigma x = 15, \quad \Sigma y = 46, \quad \Sigma xy = 50,$$
$$\Sigma x^2 = 61 \text{ and } \Sigma y^2 = 90$$

7.0

1024

725

Calculate the correlation coefficient and the slope of the regression line of Y on x.

A:

Correlation coefficient $(r) = \dfrac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2 \, \Sigma y_i^2}}$

$= \dfrac{\cancel{85}\,50}{\sqrt{61 \times 90}} = \dfrac{50}{5491}$

$$\begin{array}{r} 90 \\ \times 61 \\ \hline 91 \\ 540\times \\ \hline 5491 \end{array}$$

$= 109.82$ (appx.)

slope of regression line y on x

$b_{yx} = \dfrac{\Sigma x_i y_i}{\Sigma x_i^2}$ x on y Σy_i^2

$= \dfrac{50}{61} = 0.81$ (appx.)

~~Q3 Differentiate between Simple Random Sampling and Systematic Random Sampling.~~

Page No.
Date

Q3. chi-square test for independence of attributes.

Ans. • The chi-square test of independence is a statistical hypothesis test used to determine whether two categorical or normal variables.

• If you have only a table of values that show frequency counts, you can use the test.

Ex:-

• ~~Ques~~ We have a simple random sample 600 people who saw a movie were purchased.

• Our variables are the movie ~~ty~~ type and whether or not snacks were purchased. Both variable are categorical.

~~Sol~~

Q4. Compute the ~~th~~ three yearly moving average of the fallwing data.

Ans.

Day	Sales	3 yearly m/s	3 yearly Avg.
1	45	-	-
2	46	45+46+48=139	139/3 = 46.
3	48	46+48+47 = 141	141/3 = 47
4	47	153	51
5	58	163	54.3
6	58	161	53.6
7	57	156	52
8	52	150	50
9	53	149	49.6
10	45	149	49.6
11	51	157	52.3
12	61	-	-

Q5. The following contingency table presents the analysis of 300 persons according to hair colour and eye color.

Hair Colour	Eye Color		
	Blue	Grey	Brown
Fair	30	10	40
Brown	40	20	40
Black	50	30	40

Test the hypothesis that there is an association between hair colour and eye colour at 5% level of significance.

Page No.
Date:

step 1 :- Ho: Hair and eye colour are independent.

H1: Hair and eye colour are associated.

step 2 :- For testing the null hypothesis the formula is

$$x^2 = \sum\left(\frac{O_{ij} - E_{ij}}{E_{ij}}\right)^2$$

ij → column
↓
row

Now, under Ho, the expected frequencies can be obtained as:-

$$E_{i} = \frac{R_i \times C_i}{N} = \frac{\text{Sum of ith row} \times \text{Sum of jth column}}{\text{Total sample size}}$$

Therefore

$$E_{11} = \frac{R_1 \times C_1}{N} = \frac{80 \times 120}{300} = 32$$

$$E_{12} = \frac{R_1 \times C_2}{N} = \frac{80 \times 60}{300} = 16$$

$$E_{13} = \frac{R_1 \times C_3}{N} = \frac{80 \times 120}{300} = 32$$

$$E_{21} = \frac{R_2 \times C_1}{N} = \frac{100 \times 120}{300} = 40$$

$$E_{22} = \frac{R_2 \times C_2}{N} = \frac{100 \times 60}{300} = 20$$

$$E_{23} = \frac{R_2 \times C_3}{N} = \frac{100 \times 120}{300} = 40$$

$$E_{31} = \frac{R_3 \times C1}{N} = \frac{120 \times 120}{300} = 48$$

$$E_{32} = \frac{R_3 \times C_2}{N} = \frac{120 \times 60}{300} = 24$$

$$E_{33} = \frac{R_3 \times C_3}{N} = \frac{120 \times 120}{300} = 48$$

Step 3 :-

Observed frequency (O)	Expected frequency (E)	(O-E)	$(O-E)^2$	$\frac{(O-E)}{E}$
30	32	-2	4	0.12
10	16	-6	36	2.25
~~40~~ 40	32	8	64	2
20	40	0	0	0
40	20	0	0	0
50	40	0	0	0
30	48	0	~~4~~	0.08
40	24	2	4	1.50
	48	6 -8	36 64	1.33
Total = 300	300			7.29

$$x^2 = \Sigma \frac{(O-E)^2}{E} = 7.29$$

Page No. ____
Date

step 4:- The degree of freedom $= (r-1)(c-1)$
$= (3-1)(3-1)$
$= 2 \times 2$
$= 4$ Ans

steps:-
The critical value of chi-square with 4 d.f at 5% level of significance is 9.49 and the calculate value is 7.29.

step 6:-
since calculated value of test statistic chi-square = 7.29 is less than critical value = 9.49, so we do not reject the null hypothesis and reject the alternative hypothesis.

Thus, we can conclude that the sample provide us sufficient evidence againt the claim so hair colour is independent of eye colour.

Page No.
Date:

<BCS - 40 Dec-2021>

Q1 A Pharamacologist measured the amount of dopamine in the brainsg each of 10 rats. The observed data is given below:

6.8, 7.0, 5.3, 6.0, 5.9, 5.6, 6.8, 7.4, 6.2, 7.0

Calculate:

i). The mean and standard deviation.

Ans

x	$(x-\bar{x})$	$(x-\bar{x})^2$
6.8	6.8-6.4=0.4	0.16
7.0	7.0-6.4=0.60	0.36
5.3	-1.10	1.21
6.0	-0.40	0.16
5.9	-0.50	0.25
5.6	- 0.80	0.64
6.8	0.40	0.16
7.4	1.00	1
6.2	-0.20	0.04
7.0	0.60	0.36
64.00		4.34

$$\text{mean} = (\bar{x}) = \frac{\Sigma x_i}{n}$$

$$= \frac{64}{10} = 6.4$$

Page No. ______
Date: ______

* Standard Deviation $= \sqrt{\frac{\cancel{x}\,\Sigma(x-\text{mean})^2}{n - 1}}$ 10

$= \sqrt{\frac{4.34}{9}} = \sqrt{0.48229}$

$=$

0.6944

(ii). Coefficient of variation.

Ans.

$CV = \frac{SD}{\text{mean}} = \frac{0.6944}{6.4} = 0.1085$ Ans

Q3.

A

Before Drug	After Drug	Difference	$(Diff)^2$	D mean	D-D mean
68	71	3	9	2.66	0.33
71	70	-1	1	"	-3.66
84	82	-2	4	"	-4.66
93	97	4	16	"	1.33
67	73	6	36	"	3.33
74	80	6	36	"	3.33
		16	102		

$(D-D\text{ mean})^2$
0.11
13.44
21.77
1.77
11.11
11.11
59.33

473

973

Page No.
Date:

= mean = 2.66

~~(D - mean)²/n = 5~~

$\Sigma(D - mean)^2 = \cancel{\frac{59.33}{6}} = 9.88$

Step 2:-

$SD = \sqrt{\frac{\Sigma(D - mean)^2}{n}} = \sqrt{\frac{(59.33)^2}{6}}$

$= \frac{(16 - 2.66)^2}{6} = 3.84$

$= \sqrt{\frac{(59.3)^2}{6}} = 8.99.$

$= \sqrt{n} = \sqrt{6} = 2.44$

= Diff mean / SD o Diff / $\sqrt{n}$ = 2.077

Q8.

A.

	strate			
	I	II	III	IV
Ni = Sizes	5000	1000	2000	2000
Si	25	10	15	20

Page No.
Date:

A1. $N_1 = 5000$, $N_2 = 1000$, $N_3 = 2000$, $N_4 = 2000$

$N = \cancel{3000}\ 10000$ (total of $N_{1,2,3,4}$)

$n = 500$ given in question

$\frac{50}{250}$

1. formula = $n_1 = N_1 \times \frac{n}{N} = 5000 \times \frac{500}{10000}$

$= 250$ Ans

$n_2 = N_2 \times \frac{n}{N} = 1000 \times \frac{500}{10000} = 50$

$n_3 = N_3 \times \frac{n}{N} = 2000 \times \frac{500}{10000} = 100$

$n_4 = N_4 \times \frac{n}{N} = 2000 \times \frac{500}{10000} = 100$

2) S_i = formula = $n_1 = n \cdot \frac{N_1 S_1}{\sum N_i S_i}$

~~$n_1 = 500 \times 5000 \times 25$~~

$N_1 S_1 = 5000 \times 25 = 125,000$

$N_2 S_2 = 1000 \times 10 = 10,000$

$N_3 S_3 = 2000 \times 15 = 30,000$

$N_4 S_4 = 2000 \times 20 = 40,000$

$\qquad\qquad 205,000$

$n_1 = n \cdot \frac{N_1 S_1}{\sum N_i S_i} = 500 \times \frac{125000}{205000} = 250$

$n_2 = n \cdot \frac{N_2 S_2}{\sum N_i S_i} = 500 \times \frac{10000}{205000} = 20$

$n_3 = 500 \times \frac{30,000}{205,000} = \cancel{100}\ 60$

$n_4 = 500 \times \frac{40,000}{205000} = 80$

Q10.

A=, x y

Age	BP	$x-\bar{x}$	dx^2	$y-\bar{y}$	dy^2	$dx\,dy$
23	~~65~~ 65	29-29=-6	36	65-72=-7	49	-6×-7= 42
27	60	-5	25	-12	144	-5×-12 = 60
25	62	-4	16	-10	100	40
26	70	-3	9	-2	4	6
28	70	-1	1	-2	4	2
29	73	0	0	1	1	0
31	75	2	4	3	9	6
35	83	6	36	11	121	66
40	90	11	121	18	324	198
x=261	Y=648	Σdx=0	248	0	Σdy²=750	Σdxdy=420

$$= \Sigma x = \frac{\Sigma \bar{x}}{n} = \frac{261}{9} = 29 = \bar{x}$$

$$= \Sigma y = \frac{\Sigma \bar{y}}{n} = \frac{648}{9} = 72 = \bar{y}$$

$$= b_{yx} = \frac{\Sigma dx\,dy}{dx^2} = \frac{420}{248} = 1.69$$

$$= b_{xy} = \frac{\Sigma dx\,dy}{dy^2} = \frac{420}{756} = 0.56$$

= The regression equation Y on X is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$72 = 1.69(x = 29)$$

$$y - 72 = 1.69x - 49.01$$

Page No.
Date:

$y = 1.69x - 49.01 + 72$

$y = 1.69x + 22.99$

* ~~The~~ regression equation x on y

$x - \bar{x} = b_{xy}(y - \bar{y})$

$x - 29 = 0.56(y - \cancel{\ } 72)$

$x - 29 = 0.56y - 40.32$

$x = 0.56y = 40.32 + 29$

$x = 0.56y - 11.32$

Hence

- y on x is = $1.69x + 22.99$
- x on y is = $0.56y - 11.32$

Page No. ____
Date ____

5. ii. Assignable causes and chance causes.

A.

Chance Causes	Assignable Causes
• chance causes is a common causes.	• Assignable causes is a special causes
• It is naturally inherent we cannot remove it. It is random in nature.	• It is caused by human we can remove it. It non-random in nature
• Chance causes variation means the Process is in control.	• Assignable causes variation mean the Process is out of control.
• Result in small amount of variation.	• Result in large amount of variation.

iii. Producer's risk and Consumer's risk.

Ans.

Producer Risk	Consumer risk
• risk associated with rejection of good quality lot even though it consist of good Parts	• Risk associated with accepting Poor quality lot even though it consist of Poor Parts
• then also rejected due to sampling process	• then also accepted due to sampling Process.

Q3. Write short notes on any two of the following

~~Exponential smoothing method:~~

i). Cluster sampling :-

- Under this method, total population is divided into some recognizable sub-divisions, Known as cluster.

- Out of all the clusters, a given number of ~~obs~~ cluster are chosen at random.

- All the item covered by the selected ~~cls~~ clusters are included in the sample.

- In using cluster sampling, the following Points should be kept in mind.

- The number of sampling units in each cluster should be ~~ap~~ approximately same.

(ii). Correlation and Rank-correlation

~~[illegible]~~

Page No.
Date:

Ans. correlation :-

- correlation is a statistical measure that expresses the extent to which two variables are related.

- It is a common tool for describing simple relationship without making a statement about cause and effect.

- Ex:- Decrease in income, decreases expenditure.

*. Rank - correlation :-

- There are three types of ~~Prob~~ problem in Rank method:

1). when Ranks are given.
2). when Ranks are not given.
3). when Ranks are equal or repeated.

Q Producer's risk and Consumer's risk.

Ans Producers risk :- Risk associated with reject of good quality lot even though it consist of good Parts then also rejected due to sampling process.

- α = P (~~sega~~ rejecting a lot of P defective)

Page No. ____
Date: | | |

Consumer's risk :-

- Risk associated with accepting poor quality lot even through it consist of poor parts then also accepted due to sampling process

- β = P(accepting a lot having pt ~~of~~ defectives).

Q1. An Automobile Service centre performed the following number of car services each month.

15, 18, 25, 40, 25, 18, 25, 21, 30, 33, 25, 20, 10, 28, 36, 15, 26, 35, 20, 21, 32, 40, 32, 16, 12, 14, 22, 26, 37, 16.

Ans (i).

1) Range = Max - Min
$40 - 10 = 30$

2). class = 4

3) width of class interval = $\frac{\text{Range}}{\text{class}} = \frac{30}{4} = 7.5$

Round off = 8

i). continuous frequency distribution.

Class width	Frequency
10 - 17	7
18 - 25	11
26 - 33	7
34 - 41	5

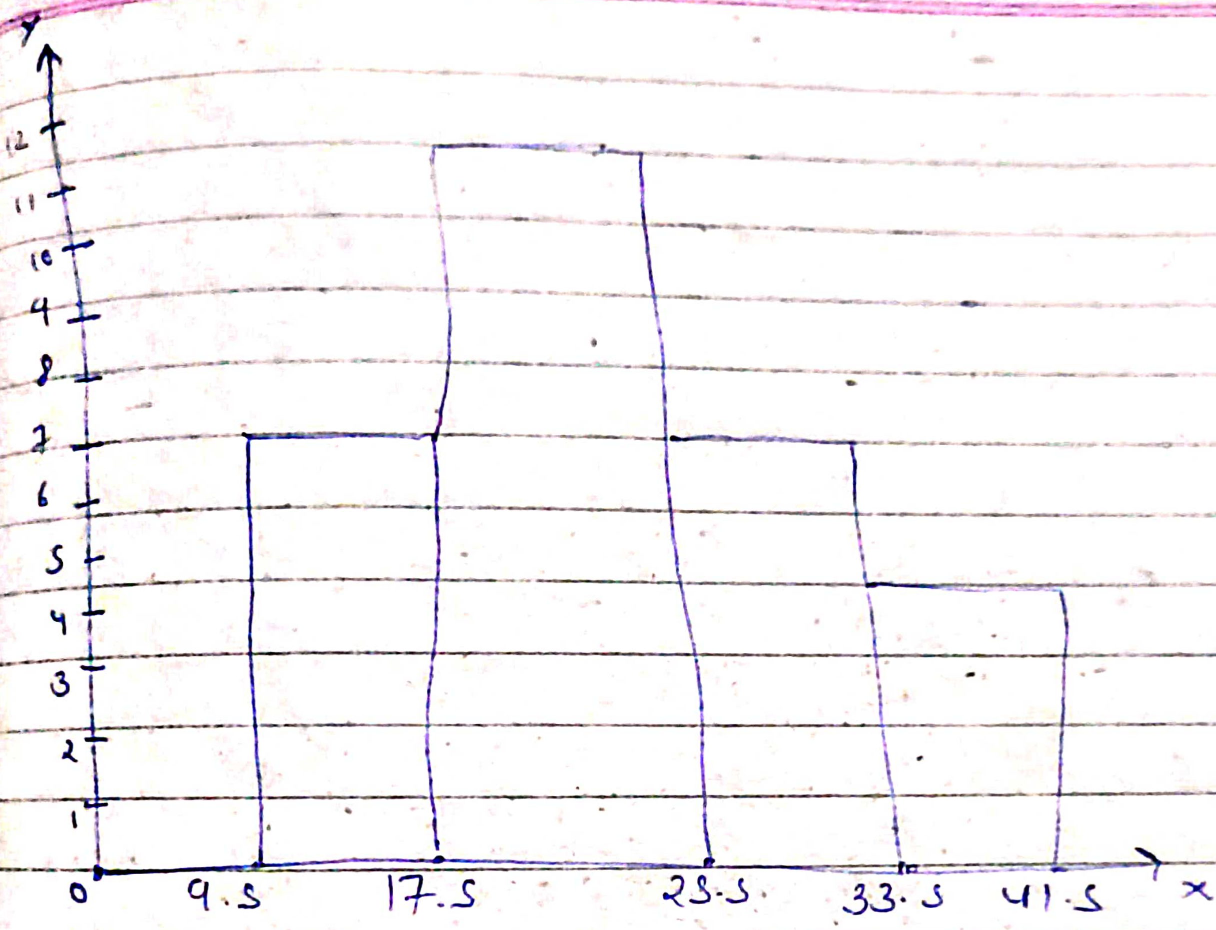
(ii).

An	Inclusive series	Exclusive Series
L ←	10 - 17 → H	9.5 - 17.5
	18 - 25	17.5 - 25.5
	26 - 33	25.5 - 33.5
	34 - 41	33.5 - 41.5

formula = $\frac{18-17}{2} = \frac{1}{2} = 0.5$

Histogram :-

Exclusive series	Frequency
9.5 - 17.5	7
17.5 - 25.5	11
25.5 - 33.5	7
33.5 - 41.5	5



y
12
11
10
9
8
7
6
5
4
3
2
1
0
9.5
17.5
25.5
33.5
41.5
x