BEST NOTES

A NOTES BY MUHIBB ALI

BIO-STATISTICS

For Universities Students

www.24hpdf.com

Date:

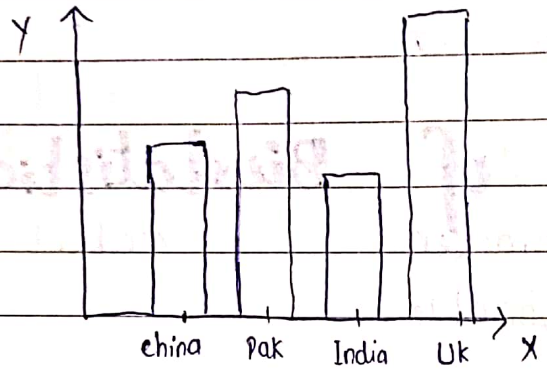
Bio-Statistics

Bio → life

Statistics → Analysis / Data collection

Statistics is the science that deals with the collection of data, analysis and interpretation of data.

- Sample
- Population



It suggests tables, graphics, figures.
It is commonly used in newspapers, books, TV, social media, lectures, speeches.

Bio-statistics :-

Biostatistics is the branch of biological science which deals with the study of method, data collection, analysis, presentation and interpretation of data of biological research. Biostatistics is also called as biometrices, since it involves many measurements and calculations.

Types :

(i) Descriptive statistics – Analysis and summerization of data.

(ii) Inferential statistics – conclusion.

Function of Bio-statistics :-

* Extent of illness
* Causation
* Outcome of health
* Research.

Some definitions of Biostatistics :-

* Statistic
* parameter
* constant
* characteristic
* variable

(i) Statistic:

Any numerical value which is calculated from sample is called statistic.

(ii) Parameter:

Any numerical value which is calculated from population is called parameter.

(iii) Characteristics:

A statistical measure that is used to summarize the value for a specific quantitative variable for all statistical units in a specific group.

Date: 18-11-2022

Scope of Statistics:-

It encompases the design of biological experiements, especially in medicine, pharmacy and agriculture the collection, summenization and analysis of data from those experiments and the interpretation and inference from the obtained results.

Population :-

A statistical population is a collection or set of all possible observations whether finite or infinite, relevant to some characteristics of interest.

e.g., A statistical population may be real such as the heights of all college students.

Sample:-

A sample is a representative part of population.

e.g., The set of 500 students being a subset of all college students.

Data Collection:

The most important part of statistical work is perhaps the collection of data. Statistical data are collected either by a complete enumeration of whole field is called census or data collection.

Methods:-

There are two methods of Data collection.

* Primary data
* Secondary data.

Date:

$K = 1 + 3.3 \log N$ → N = Number of total values

$K = 1 + 3.3 \log(60)$

$K = 1 + 3.3 \log(1.778)$

$K = 1 + 5.8678$

$\boxed{K = 6.867}$

Range = $R = X_{max} - X_{min}$

$R = 204 - 68$

$\boxed{R = 136}$

Now,

We have to find height (h)

$h = \frac{R}{k}$

$h = \frac{136}{7} \rightarrow 19.4285$

$\boxed{h = 20}$

Midpoint = $\frac{\text{Lower value} + \text{upper value}}{2}$

Class Interval	Enteries.	Tally	f	C.B	Mid point
65 – 84	76, 82, 70, 68, 84, 78, 75, 80, 82	~~||||~~ ||||	9	64.5 – 84.5	74.5
85 – 104	93, 95, 98, 86, 99, 99, 100, 98, 90, 104	~~||||~~ ~~||||~~	10	84.5 – 104.5	94.5
105 - 124					

Date:

Prepare a frequency table for the price data given below taking 5 units as the width of class-interval.

					84	82	80	78	91
100	96	92	88	86	73	71	69	58	56
87	83	79	77	75	73	71	69	58	56
73	50	57	55	53	51	48	46	63	59
55	51	49	47	45	43	41	58	54	50
56	44	42	40	38	36	46	53	50	43

Height = h = 5

Class interval	Enteries	Tally	Frequency	Class Boundries	Mid point.
35 - 39	38, 36	II	2	34.5 - 39.5	37
40 - 44	43, 41, 44, 42, 40, 43	~~IIII~~ I	6	39.5 - 44.5	42
45 - 49	48, 46, 49, 47, 45, 46	~~IIII~~ I	6	44.5 - 49.5	47
50 - 54	50, 53, 51, 51, 54, 50, 53, 50	~~IIII~~ III	8	49.5 - 54.5	52
55 - 59	58, 56, 57, 55, 59, 55, 58, 56	~~IIII~~ III	8	54.5 - 59.5	57
60 - 64	63	I	1	59.5 - 64.5	62
65 - 69	69	I	1	64.5 - 69.5	67
70 - 74	73, 71, 73	III	3	69.5 - 74.5	72
75 - 79	79, 77, 75, 78	IIII	4	74.5 - 79.5	77
80 - 84	84, 82, 80, 83,	IIII	4	79.5 - 84.5	82

Date:

85-89	88, 86, 87	|||	3	84.5-89.5	87
90-94	92, 91	||	2	89.5-94.5	92
95-99	96	|	1	94.5-99.5	97
100-104	100	|	1	99.5-104.5	102

b) Tabulate the following marks in a grouped frequency distribution:-

74	49	103	95	90	118	52	88	101	96	72	56	64	110	97
59	62	96	82	65	85	105	116	91	83	99	52	76	84	89
77	104	96	84	62	58	66	100	80	54	75	55	99	101	78
66	96	83	57	60	51	114	120	121	92	88	64	63	95	78

N = Number of total values = 60

$K = 1 + 3.3 \log N$

$K = 1 + 3.3 \log(60) \rightarrow K = 1 + 3.3(1.7781)$

$K = 1 + 5.8677 \rightarrow K = 6.8677 \rightarrow \boxed{K = 7}$

Range = $R = X_{max} - X_{min}$

$R = 121 - 49 \Rightarrow 72$

Now, we have to find the height

$$\text{Height} = \frac{R}{K}$$

$$h = \frac{72}{7} \rightarrow 10.2857$$

$$\boxed{h = 10}$$

Date:

Class interval	Enteries	Tally	Frequency	Class Boundaries	Mid point
40 - 49	49	|	1	39.5 - 49.5	44.5
50 - 59	52, 56, 59, 52, 58, 54, 55, 57, 51	𝍸 ||||	9	49.5 - 59.5	54.5
60 - 69	64, 62, 65, 62, 66, 66, 60, 64, 63	𝍸 ||||	9	59.5 - 69.5	64.5
70 - 79	74, 72, 76, 77, 75, 78, 78	𝍸 ||	7	69.5 - 79.5	74.5
80 - 89	88, 82, 85, 83, 84, 89, 84, 80, 83, 88	𝍸 𝍸	10	79.5 - 89.5	84.5
90 - 99	95, 90, 96, 97, 96, 91, 99, 96, 99, 96, 92, 95	𝍸 𝍸 ||	12	89.5 - 99.5	94.5
100 - 109	103, 101, 105, 104, 100, 104.	𝍸 |	6	99.5 - 109.5	104.5
110 - 119	118, 110, 116 114	||||	4	109.5 - 119.5	114.5
120 - 129	120, 121	||	2	119.5 - 129.5	124.5

Date:

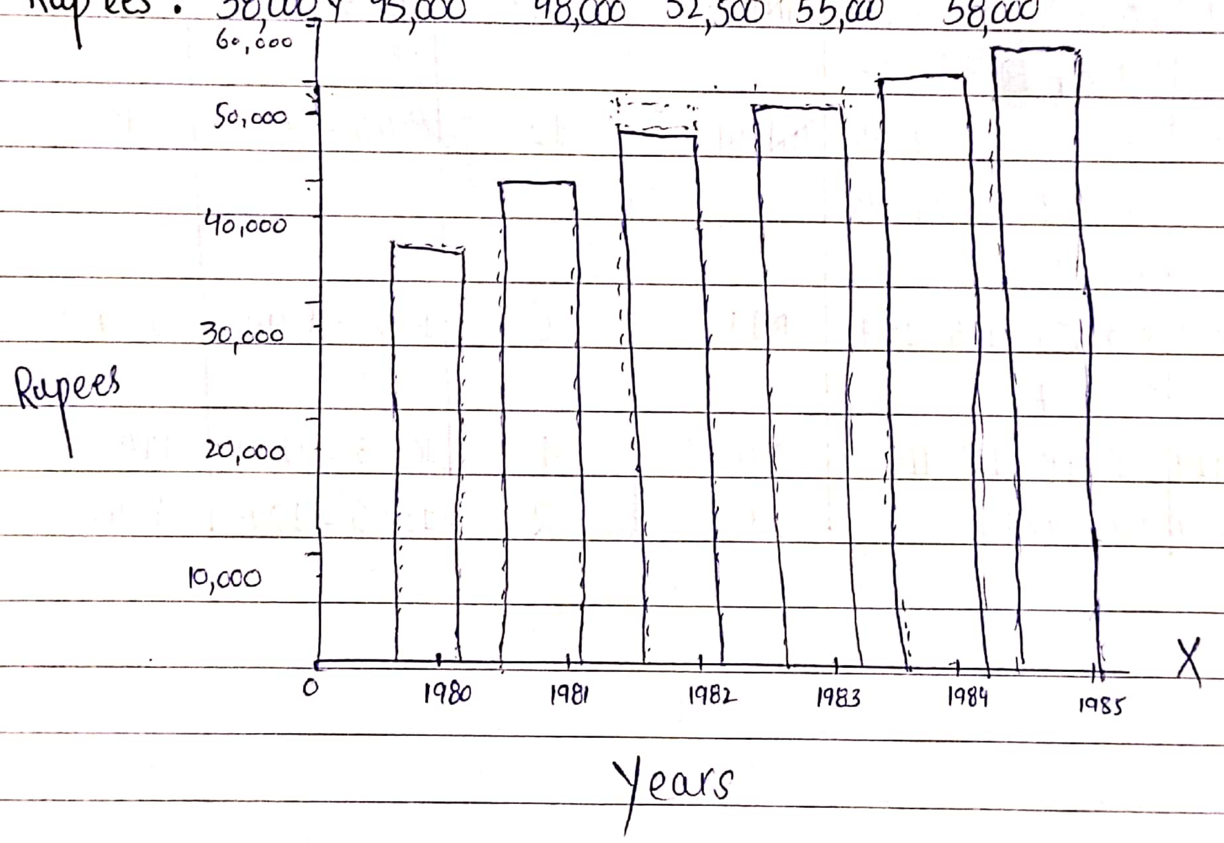
Formation of tables and charts :-

Simple bar chart :-

A simple bar chart consists of horizontal or vertical bars of equal widths and lengths proportional to the value of they represent.

E.g., (1)

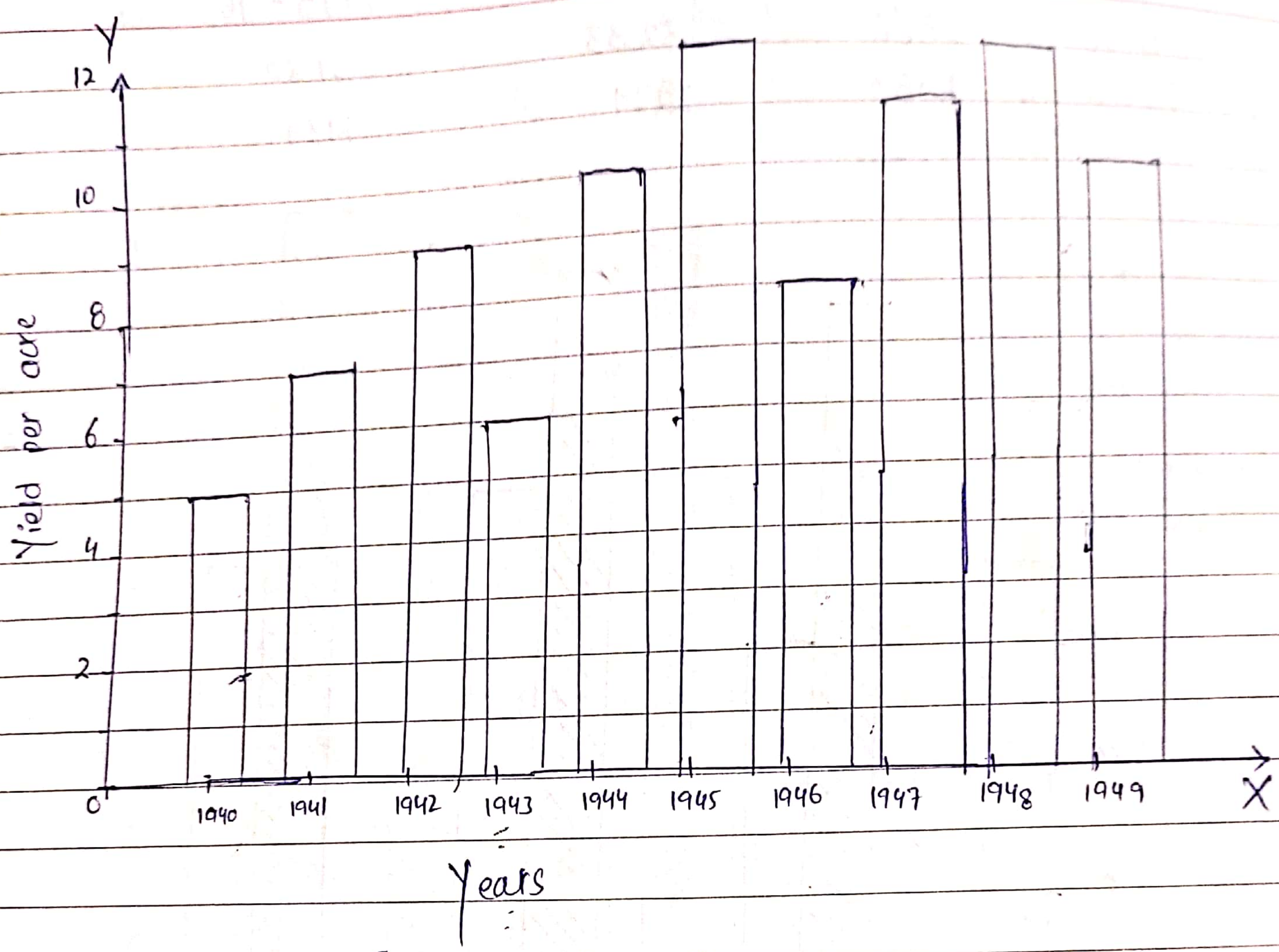
Years :	1980	1981	1982	1983	1984	1985
Rupees :	38,000	45,000	48,000	52,500	55,000	58,000



(2) Represent the following yield per acre data by a bar diagram.

Years :	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949
Yield per acre :	5	7	9	6	10	12	8	11	12	10

Date:



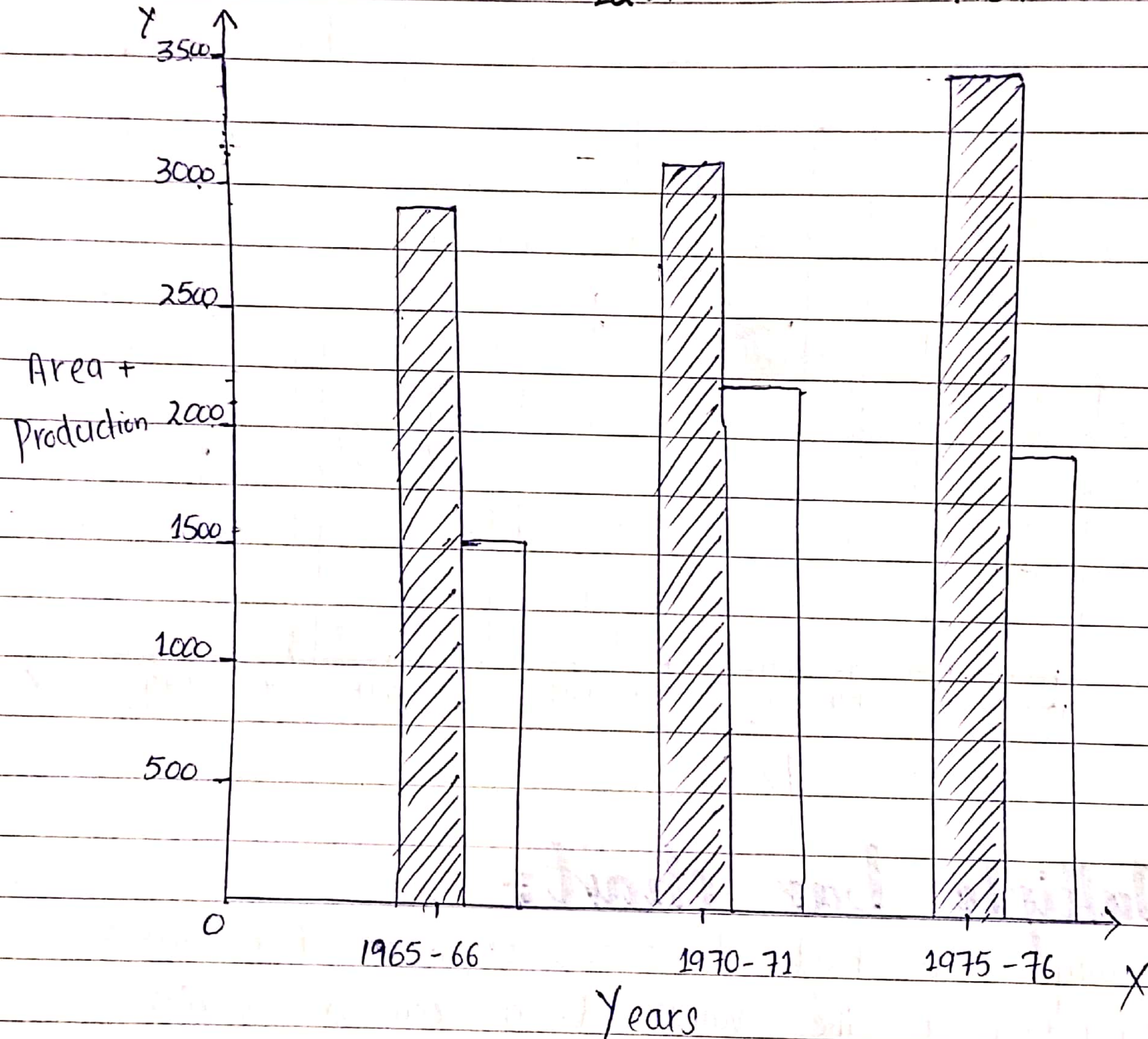
Multiple bar chart :-

A multiple bar chart shows two or more characteristics corresponding to the values of a common variable in the form of grouped bars, whose lengths are proportional to the values of the characteristics and each of which is shaded or coloured differently to aid identification.

Example :-

Date:

Year :	1965-66	1970 - 71	1975 - 76
Area :	2886	3233	3420
Production:	1588	2229	1937

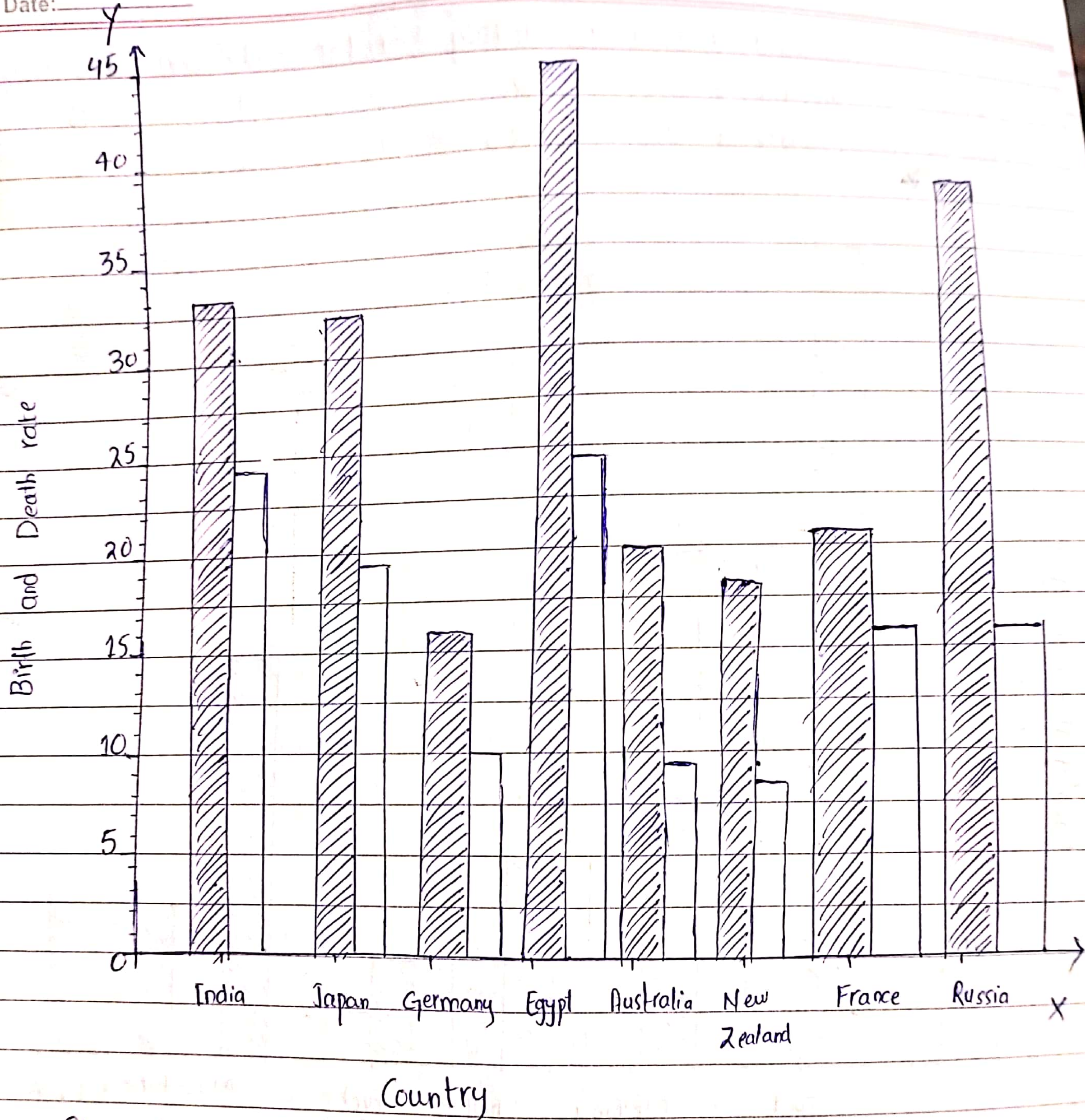


* Shaded portion shows the area.
* Blank portion shows the production.

Example 2:-

Country :	India	Japan	Germany	Egypt	Australia	New Zealand	France	Russia
Birth rate :	33	32	16	44	20	18	21	38
Death rate :	24	19	10	24	9	8	16	16

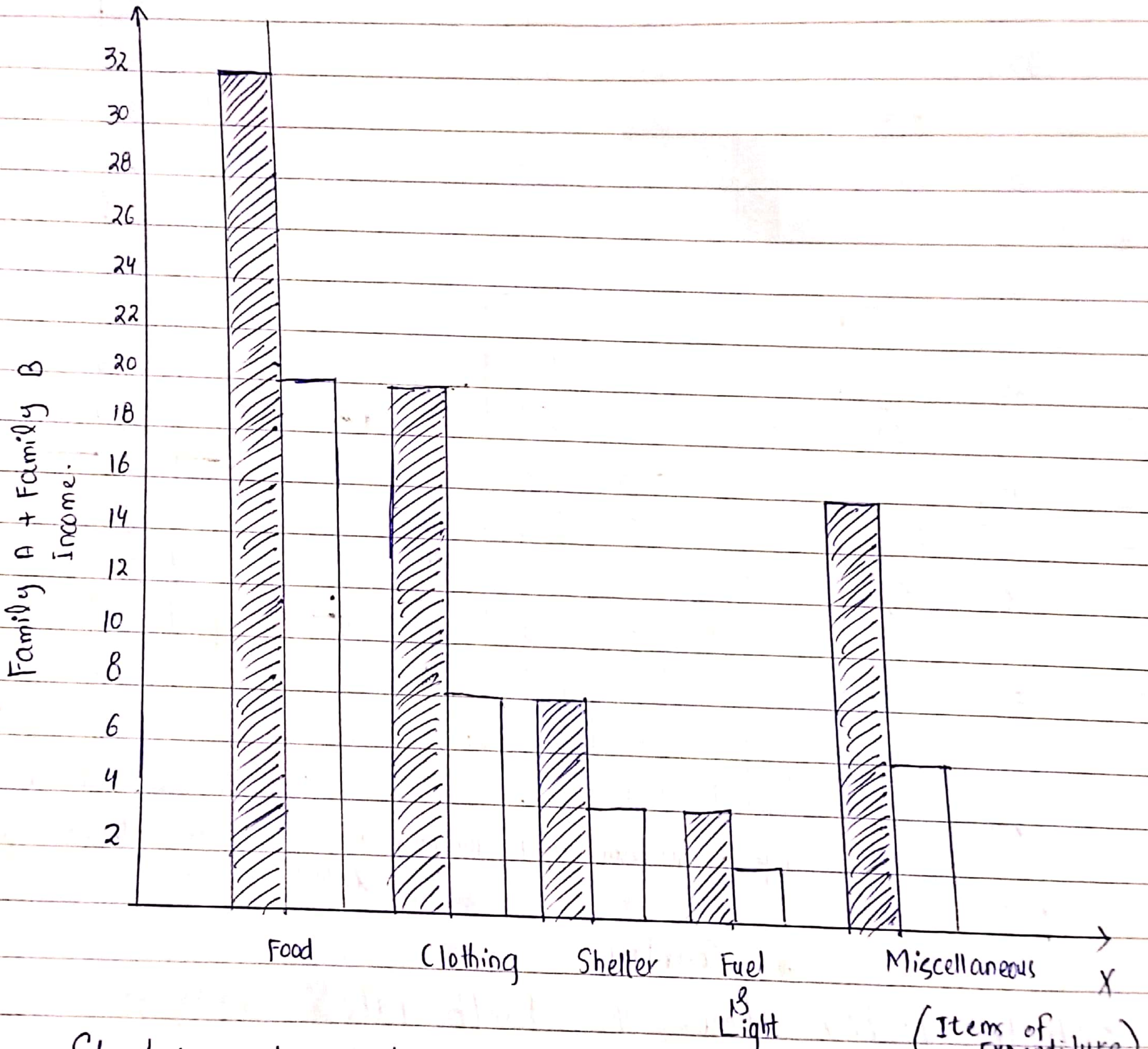
Date:



* Shaded portion shows the Birth rate.
* Blank portion shows the Death rate.

Date:

Items of Expenditure :	Food	Clothing	Shelter	Fuel & Light	Miscellaneous
Family A income :	32	20	8	4	16
Family B income :	20	8	4	2	6



* Shaded portion shows the family A income.
* Blank portion shows the family B income.

Scanned with CamScanner

Date:

Component Bar chart:-

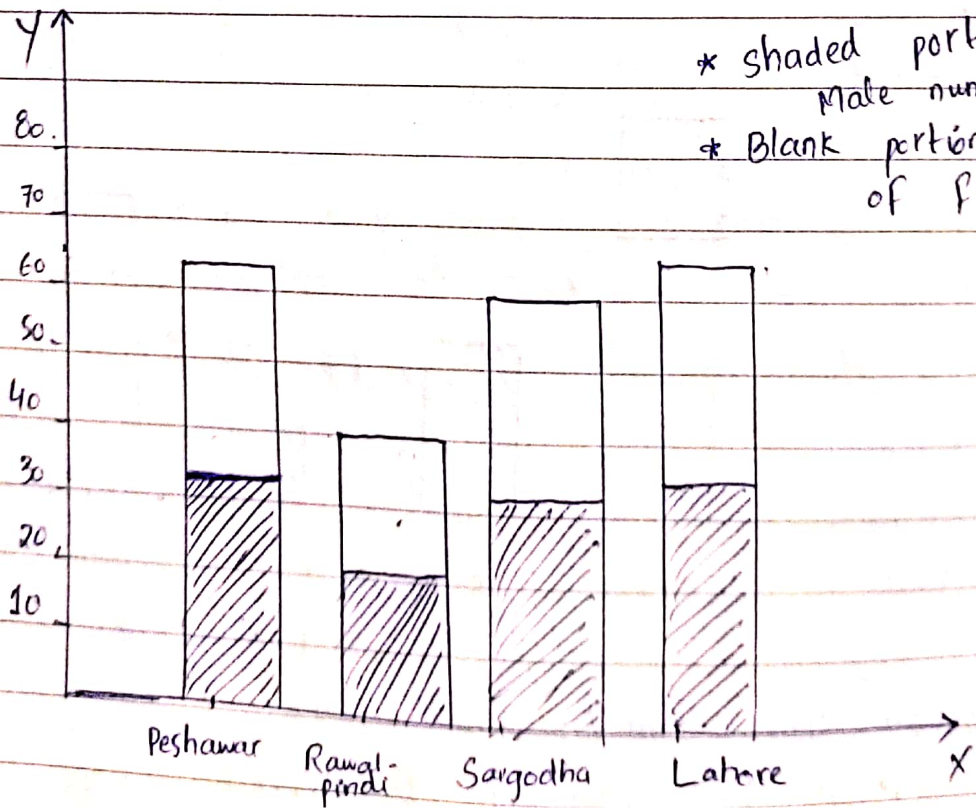
A component bar chart is an effective technique in which each bar is divided into two or more sections, proportional in size to the component parts of a total beig displayed by each bar.

Component bar charts are used to represent the cumulation of the various components of data and the percentages

They are also known as sub-divided bars.

Example:-

Division	Both genders	Male	Female
Peshawar	64	33	31
Rawalpindi	40	21	19
Sargodha	60	32	28
Lahore	65	35	30



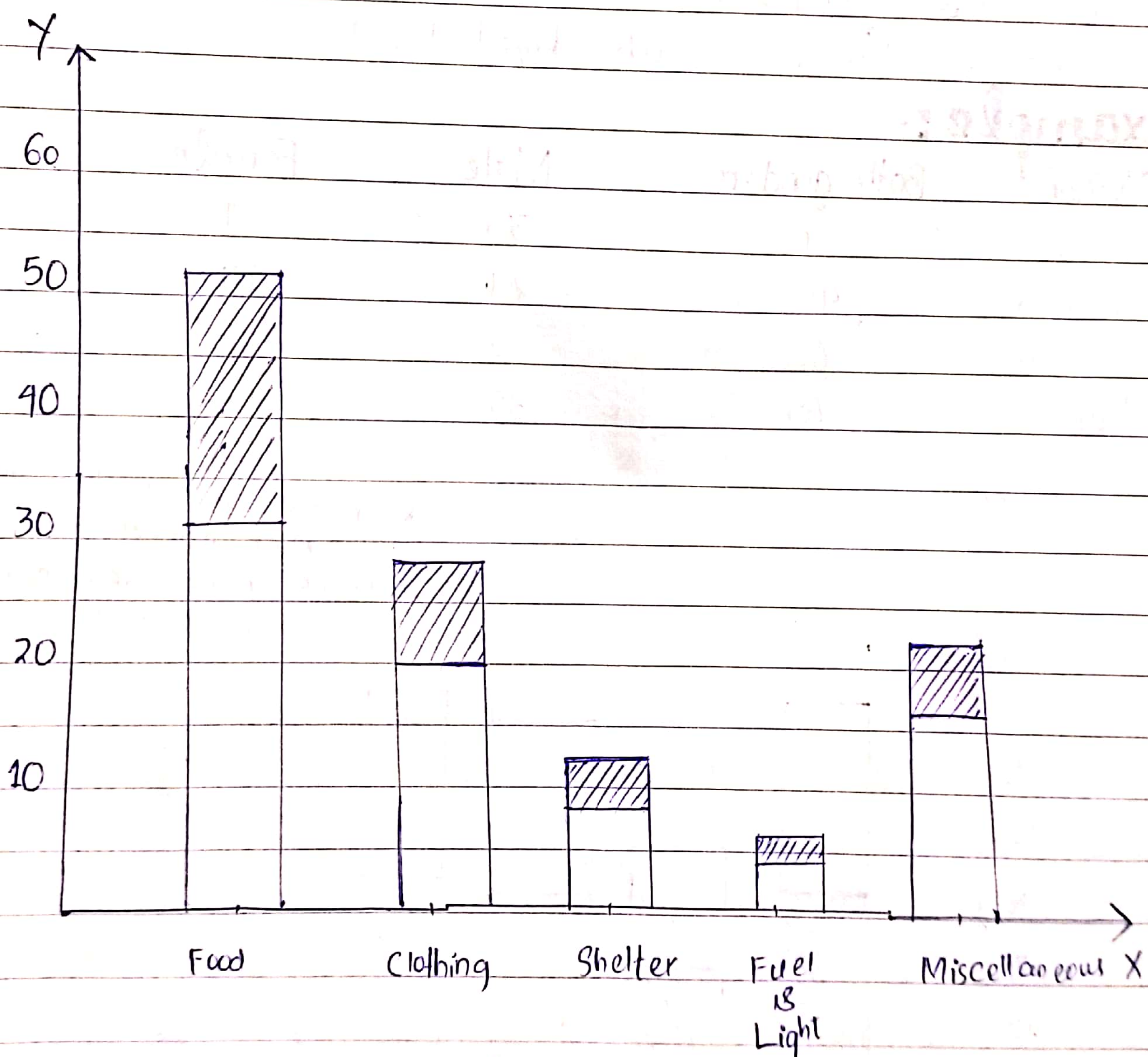
* shaded portion show the Male number.
* Blank portion show number of females.

Scanned with CamScanner

Date:

Example 2:-

Items of Expenditure	Income: Family A	Family B	Both
Food	32	20	52
Clothing	20	8	28
Shelter	8	4	12
Fuel & Light	4	2	6
Miscellaneous	16	6	22



* Blank portion show family A income.
* Shaded portion show family B income.

Scanned with CamScanner

Date:

Arithmetic Mean:-

Arithmetic mean is the sum of all observations in the given data set divided by the total number of observation in the data set.

Formula:

$$\text{Mean} = \frac{\text{sum of all observations}}{\text{Total no. of observations}}$$

$$\Sigma \bar{X} = \frac{X_1 + X_2 + X_3 \cdots X_n}{n}$$

Examples:-

1- Marks obtained by 9 students are as :-
45, 32, 37, 46, 39, 36, 41, 48, 36, Find Arithmetic Mean.

As we know that;

$$\bar{X} = \frac{X_1 + X_2 + X_3 \cdots X_n}{n}$$

$$\bar{X} = \frac{45 + 32 + 37 + 46 + 39 + 36 + 41 + 48 + 36}{9}$$

$$\bar{X} = \frac{360}{9} \rightarrow \boxed{\bar{X} = 40}$$

2- The number of cars crossing a certain bridge in a big city in 10 intervals of five each was recorded as:
25, 15, 18, 30, 20, 20, 12, 9, 16, 15, Find A.M.

As we know that;

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots X_n}{n}$$

$$\bar{X} = \frac{25 + 15 + 18 + 30 + 20 + 20 + 12 + 9 + 16 + 15}{10}$$

Date:

$\bar{X} = \frac{180}{10} \rightarrow \boxed{\bar{X} = 18}$

3- The monthly income of 10 families in rupees are as :-

Family:	A	B	C	D	E	F	G	H	I	J
Income:	85	70	10	75	500	8	42	250	40	36

As we know that ;

$$\bar{X} = \frac{X_1 + X_2 + X_3 \cdots X_n}{n}$$

$$\bar{X} = \frac{85+70+10+75+500+8+42+250+40+36}{10}$$

$$\bar{X} = \frac{1116}{10} \rightarrow \boxed{\bar{X} = 111.6}$$

4- Calculate the arithmetic mean of animal incomes of 15 families are as ;

60, 80, 90, 96, 120, 150, 200, 360, 480, 520, 1060, 1200, 1450, 2500, 7200

As we know that,

$$\bar{X} = \frac{X_1 + X_2 + X_3 \cdots X_n}{n}$$

$$\bar{X} = \frac{60+80+90+96+120+150+200+360+480+520+1060+1200+1450+2500+7200}{15}$$

$$\bar{X} = \frac{15566}{15} \rightarrow \boxed{\bar{X} = 1037.733}$$

Weighted Arithmetic Mean :-

The multipliers or a set of numbers which express more or less adequately the relative importance of various

Date:

observations in a set of data are technically called weight.

Formula:

$$\bar{X}_w = \frac{X_1 w_1 + X_2 w_2 + X_3 w_3 + \cdots X_n w_n}{w_1 + w_2 + w_3 + \cdots w_n}$$

$$\bar{X}_w = \frac{\Sigma x_i w_i}{\Sigma w_i}$$

Examples:-

1. Calculate the weighted mean;

Items	Expenditures	Weight
Food	290	7.5
Rent	54	2.0
Cloth	98	1.5
Fuel & Light	75	1.0
Other items	75	0.5

Items	Expenditures	Weights (w_i)	$X_i w_i$
Food	290	7.5	2175.0
Rent	54	2.0	108.0
Cloth	98	1.5	147.0
Fuel & Light	75	1.0	75.0
Other items	75	0.5	37.5
Total		$\Sigma w_i = 12.5$	$\Sigma x_i w_i = 2542.5$

As we know that,

$$\bar{X}_w = \frac{\Sigma x_i w_i}{\Sigma w_i}$$

$$\bar{X}_w = \frac{2542.5}{12.5} \rightarrow \boxed{\bar{X}_w = 203.4}$$

Scanned with CamScanner

Date:

2- Find weight average using the quantities in column 3 as weight.

Piece good	Price per meter (x)	Quantity (Σw) (million meters)	Σxw
Unbleached	8.37	286	2393.82
Bleached	9.50	255	2422.5
Printed flags	9.16	64	586.24
Other sorts	9.84	172	1692.48
Dyed in piece	13.65	162	2211.3
Dyed yarn	11.95	80	956
	$\Sigma x = 62.47$	$\Sigma w = 1019$	$\Sigma xw = 10262.34$

As we know that ;

$$\bar{X}w = \frac{\Sigma xw}{\Sigma w}$$

$$\bar{X}w = \frac{10262.34}{1019} \rightarrow \boxed{\bar{X}w = 10.0709}$$

$$\bar{X} = \frac{\Sigma x}{n}$$

$$\bar{X} = \frac{62.47}{6} \rightarrow \boxed{\bar{X} = 10.4117}$$

Scanned with CamScanner

Date:

Mean :

For grouped data :-

$$\bar{X} = \frac{\Sigma fx}{\Sigma f}$$

For ungroup data :-

$$\bar{X} = \frac{\Sigma x}{n}$$

For short cut Method :-

1 fixed value from x

$$\bar{X} = a + h\bar{u} \rightarrow u = \frac{X - a}{h} ; \bar{U} = \frac{\Sigma fu}{\Sigma f}$$

h = 20

Examples :-

1.

Classes	f	x (mid point)	fx	u	fu
65-84	9	74.5	670.5	-2	-18
85-104	10	94.5	945.0	-1	-10
105-124	17	114.5	1946.0	0	0
125-144	10	134.5	1345.0	1	10
145-164	5	154.5	722.5	2	10
165-184	4	174.5	698.0	3	12
185-204	5	194.5	972.5	4	20
	$\Sigma f = 60$	$\Sigma x = 941.5$	$\Sigma fx = 7350$		$\Sigma fu = 24$

① $\bar{X} = \frac{\Sigma fx}{\Sigma f} \Rightarrow \frac{7350}{60} \rightarrow \bar{X} = 122.5$

② $\bar{X} = \frac{\Sigma x}{n} \Rightarrow \frac{941.5}{7} \Rightarrow \bar{X} = 134.5$

③ $\bar{X} = a + h\bar{u}$

$\bar{u} = \frac{\Sigma fu}{\Sigma f} \Rightarrow \frac{24}{60} \Rightarrow 0.4$

Scanned with CamScanner

Date:

$\bar{X} = a + h\bar{u}$

$= 114.5 + 20(0.4)$

$= 114.5 + 8$

$\bar{X} = 122.5$

2.

h=5

Classes	f	x	fx	u	fu
35-39	15	37	555	-2	-30
40-44	13	42	546	-1	-13
45-49	17	47→a	799	0	0
50-54	29	52	1508	1	29
55-59	11	57	627	2	22
60-64	10	62	620	3	30
65-69	5	67	335	4	20
	$\Sigma f = 100$	$\Sigma x = 364$	$\Sigma fx = 4990$		$\Sigma fu = 58$

① $\bar{X} = \frac{\Sigma fx}{\Sigma f} \Rightarrow \frac{4990}{100} \rightarrow \boxed{\bar{X} = 49.9}$

② $\bar{X} = \frac{\Sigma x}{n} \Rightarrow \frac{364}{7} \longrightarrow \boxed{\bar{X} = 52}$

③ $\bar{X} = a + h\bar{u}$

$\bar{u} = \frac{\Sigma fu}{\Sigma f} \rightarrow \frac{58}{100} \longrightarrow \bar{u} = 0.58$

$\bar{X} = a + h\bar{u}$

$\bar{X} = 47 + 5(0.58)$

$\bar{X} = 47 + 2.9 \longrightarrow \boxed{\bar{X} = 49.9}$

Date: 3

Variable (Classes)	f	Mid value (x)	fx	u	fu
0-5	2	2.5	5	-1.67	-3.34
5-10	5	7.5	37.5	-0.83	-4.15
10-15	7	12.5 → a	87.5	0	0
15-20	13	17.5	227.5	0.83	10.79
20-25	21	22.5	472.5	1.67	35.07
25-30	16	27.5	440	2.50	40
30-35	8	32.5	260	3.33	26.64
35-40	3	37.5	112.5	4.17	12.51
	$\Sigma f = 75$	$\Sigma x = 160$	$\Sigma fx = 1642.5$		$\Sigma fu = 117.52$

h = 6

① $\bar{X} = \frac{\Sigma fx}{\Sigma f} \rightarrow \bar{X} = \frac{1642.5}{75} \rightarrow \boxed{\bar{X} = 21.9}$

② $\bar{X} = \frac{\Sigma x}{n} \rightarrow \bar{X} = \frac{160}{8} \rightarrow \boxed{\bar{X} = 20}$

③ $\bar{X} = a + h\bar{u}$

$\bar{u} = \frac{\Sigma fu}{\Sigma f} \rightarrow \frac{117.52}{75} \rightarrow \bar{u} = 1.57$

$\bar{X} = a + h\bar{u}$

$= 12.5 + 6(1.57)$

$= 12.5 + 9.42$

$\boxed{\bar{X} = 21.92}$

Scanned with CamScanner

Mode :-

It is defined as a value which occurs most repeated or frequent value in a set of data.

$$\text{Mode} = l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

Where, l = Lower class boundary
f_m = Frequency of model class
f_1 = perceeding value of f
f_2 = Associated value of f.

Median :-

It is defined as a value which divides a data set that have been ascending or descending ordered into two equal parts, then the middle value is called median.

For grouped data.

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c.f\right)$$

$c.f \rightarrow$ cumulative frequency.

e.g., For ungrouped data,

3, 5, 6, 9, 4, 7, 12, 13, 8

Arrange in ascending order.

$\underbrace{3, 4, 5, 6}, \underset{\downarrow \text{ median.}}{7}, \underbrace{8, 9, 12, 13}$

Mode :-

f	
2	
3	$\rightarrow f_1$
9	$\rightarrow f_m$ (model value / large value)
6	$\rightarrow f_2$
5	

Mode e.g., 2, 6, 5, 6, 3, 4, 6, 7

Mode = 6 $\rightarrow$ repeated value.

Date:

QNo.1:- Find mean, median, and mode.

Class	f	x	Class boundary	c.f
30-39	8	34.5	29.5 - 39.5	8
40-49	87	44.5	39.5 - 49.5	95
50-59	190 f_1	54.5	49.5 - 59.5	285
60-69	304 f_m	64.5	59.5 - 69.5	589
70-79	211 f_2	74.5	69.5 - 79.5	800
80-89	85	84.5	79.5 - 89.5	885
90-99	20	94.5	89.5 - 99.5	905
	$\Sigma f = n = 905$			

$$\text{Mode} = l + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 59.5 + \frac{304 - 190}{(304-190) + (304-211)} \times 10$$

$$= 59.5 + \frac{114}{114 + 93} \times 10$$

$$= 59.5 + \frac{114}{207} \times 10$$

$$= 59.5 + \frac{1140}{207}$$

$$= 59.5 + 5.507$$

$$\text{Mode} = 65.0$$

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c.f\right)$$

$$= 59.5 + \frac{10}{304}\left(\frac{905}{2} - 285\right)$$

$$= 59.5 + 0.032\,(452.5 - 285)$$

$$= 59.5 + 0.032\,(167.5)$$

$$= 59.5 + (\cancel{1.957})\ 5.36$$

$$= 64.86$$

$$= 65$$

Q No. 2 :-

Class	f	x	Class boundaries	C.f
118-126	3	122	117.5 - 126.5	3
127-135	5	131	126.5 - 135.5	8
136-144	9 f_1	140	135.5 - 144.5	17
145-153	12 f_m	149	144.5 - 153.5	29
154-162	5 f_2	158	153.5 - 162.5	34
163-171	4	167	162.5 - 171.5	38
172-180	2	176	171.5 - 180.5	40
	$\Sigma f = n = 40$	$\Sigma x = 1043$		

$$\text{Mode} = l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 144.5 + \frac{(12-9)}{(12-9)+(12-5)} \times 9$$

$$= 144.5 + \frac{3}{3+7} \times 9$$

$$= 144.5 + \frac{3}{10} \times 9$$

Date:

$$= 144.5 + \frac{27}{10}$$

$$= 144.5 + 2.7$$

Mode $= 147.2$

Median $= l + \frac{h}{f}\left(\frac{n}{2} - c.f\right)$

$$= 144.5 + \frac{9}{12}\left(\frac{40}{2} - 17\right)$$

$$= 144.5 + \frac{9}{12}(20 - 17)$$

$$= 144.5 + \frac{9}{12}(3) \rightarrow 144.5 + 0.75(3)$$

$$= 144.5 + 2.25$$

Median $= 147$

Q NO. 03:-

Class	f	x	Class boundary	C.F
5-24	4	14.5	4.5 - 24.5	4
25-44	6	34.5	24.5 - 44.5	10
45-64	14 f_1	54.5	44.5 - 64.5	24
65-84	22 f_m	74.5	64.5 - 84.5	46
85-104	14 f_2	94.5	84.5 - 104.5	60
105-124	5	114.5	104.5 - 124.5	65
125-144	7	134.5	124.5 - 144.5	72
145-164	3	154.5	144.5 - 164.5	75
	$\Sigma f = n = 75$	$\Sigma x = 676$		

Mode $= l + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$

Scanned with CamScanner

Date:

$$= 64.5 + \frac{22-14}{(22-14)+(22-14)} \times 20$$

$$= 64.5 + \frac{8}{8+8} \times 20$$

$$= 64.5 + \frac{8}{16} \times 20 \Rightarrow 64.5 + \times \frac{160}{16}$$

$$= 64.5 + 10$$

$$\boxed{\text{Mode} = 74.5}$$

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c.f\right)$$

$$= 64.5 + \frac{20}{22}\left(\frac{75}{2} - 24\right)$$

$$= 64.5 + \frac{20}{22}(37.5 - 24)$$

$$= 64.5 + \frac{20}{22}(13.5)$$

$$= 64.5 + 0.9091(13.5)$$

$$= 64.5 + 12.2729$$

$$\boxed{\text{Median} = 76.7729}$$

Date:

Range :-

It is defined as the difference between the largest and smallest observation in a set of data is called range.

$$R\ (\text{Range}) = X_{max} - X_{min}$$

$$R = X_m - X_o$$

Example :-

2, 3, 5, 9, 6, 11, 4, 7, 3

$$R = X_{max} - X_{min}$$

$$= 11 - 2$$

$$\boxed{R = 9}$$

$$\text{Coefficient of dispersion} = \frac{X_m - X_o}{X_m + X_o}$$

$$= \frac{11 - 2}{11 + 2}$$

$$= \frac{9}{13}$$

$$= 0.69$$

Variance :-

The mean of the squares of deviation of all observations from their mean, it is calculated from the entire population, then variance is called population variance.

Date:

For population data.
$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

For grouped data:-
$$\sigma^2 = \frac{\Sigma f x^2}{n} - \left(\frac{\Sigma f x}{n}\right)^2$$

For sample data.
$$S^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

Standard deviation:-

The positive square root of the variance called standard deviation.

For population data
$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$$

For grouped data:-
$$\sigma = \sqrt{\frac{\Sigma f x^2}{n} - \left(\frac{\Sigma f x}{n}\right)^2}$$

For sample data.
$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

Co-efficient of variation:-
$$C.V = \frac{S.D}{Mean} \times 100$$

QNO. 01:-

A population N= 10 has the observation. 7, 8, 10, 13, 14, 19, 20, 25, 26, 28. Find variance, S.D and C.V.

$$\mu = \frac{\Sigma x}{N}$$

$$\mu = \frac{170}{10} \rightarrow 17$$

For variance

$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} = 534/10$

$\sigma^2 = 53.4$

$S.D = \sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$

$\sigma = \sqrt{53.4} = \boxed{\sigma = 7.31}$

$C.V = \frac{S.D}{\mu} \times 100$

$C.V = \frac{7.31}{17} \times 100$

$\boxed{C.V = 43\,\%}$

Date:

x	$x-\mu$	$(x-\mu)^2$
7	-10	100
8	-9	81
10	-7	49
13	-4	16
14	-3	9
19	2	4
20	3	9
25	8	64
26	9	81
28	11	121
$\Sigma x = 170$		$\Sigma(x-\mu)^2 = 534$

Q No. 02 :-

Calculate variance and standard deviation from following marks obtained by 9 students.

x	$(x-\bar{x})$	$(x-\bar{x})^2$
45	5	25
32	-8	64
37	-3	9
46	6	36
39	-1	1
36	-4	16
41	1	1
48	8	64
36	-4	6
$\Sigma x = 360$		$\Sigma(x-\bar{x})^2 = 232$

$\bar{x} = \frac{\Sigma x}{n} \rightarrow \frac{360}{9} \rightarrow \bar{x} = 40$

For variance

$S^2 = \frac{\Sigma(x-x)^2}{n} \rightarrow \frac{232}{9} \Rightarrow \boxed{S^2 = 25.7}$

$S^2 = 26$

For standard deviation

$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$

$S = \sqrt{26} \Rightarrow \boxed{S = 5.0990}$

Coefficient of variation $= \frac{S.D}{Mean} \times 100$

$= \frac{5.0990}{40} \times 100$

$= 12.7475$

Scanned with CamScanner

Date:

QNO.03:

For a population of numbers 10, 8, 7, 9, 5, 12, 8, 6, 8, 2, calculate σ^2 and σ.

x	$x-\mu$	$(x-\mu)^2$
10	2.5	6.25
8	0.5	0.25
7	-0.5	0.25
9	1.5	2.25
5	-2.5	6.25
12	4.5	20.25
8	0.5	0.25
6	-1.5	2.25
8	0.5	0.25
2	-5.5	30.25
$\Sigma x = 75$		$\Sigma(x-\mu)^2 = 68.5$

$\mu = \frac{\Sigma x}{N} \rightarrow \frac{75}{10} \rightarrow \boxed{\mu = 7.5}$

For variance,

$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} = \frac{68.5}{10}$

$\boxed{\sigma^2 = 6.85}$

For standard deviation

$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \sqrt{6.85}$

$\boxed{\sigma = 2.617}$

Coefficient of variation $= \frac{S.D}{Mean} \times 100$

$C.V = \frac{2.617}{7.5} \times 100$

$\boxed{C.V = 34.8933\%}$

QNO.04:-

Classes	f_1	f_2	x	x^2	f_1x	f_2x	f_1x^2	f_2x^2
20 - 30	7	5	25	625	175	125	4375	3125
30 - 40	10	9	35	1225	350	315	12250	11025
40 - 50	20	21	45	2025	900	945	40,500	42525
50 - 60	18	15	55	3025	990	825	54,450	45375
60 - 70	7	6	65	4225	455	390	29,575	25350
			$\Sigma x = 225$		$\Sigma f_1x = 2870$	$\Sigma f_2x = 2600$	$\Sigma f_1x^2 = 141,150$	$\Sigma f_2x_2 = 127,400$

Scanned with CamScanner

Date:

For variance :-

$$\sigma^2 = \frac{\Sigma f x^2}{n} - \left(\frac{\Sigma f x}{n}\right)^2$$

$$\bar{x} = \frac{\Sigma x}{n} \rightarrow \frac{225}{5} \rightarrow 45$$

For f_1, then

$$\sigma^2 = \frac{\Sigma f_1 x^2}{n} - \left(\frac{\Sigma f_1 x}{n}\right)^2$$

$$\sigma^2 = \frac{141,150}{5} - \left(\frac{2870}{5}\right)^2$$

$$\sigma^2 = \frac{141,150}{5} - \frac{8,236,900}{25}$$

$$\sigma^2 = \frac{705,750 - 8,236,900}{25}$$

$$\sigma^2 = \frac{-7531,150}{25}$$

$$\sigma^2 = -301,246$$

$$\sigma^2 = \frac{141,150}{5} - \frac{8236,900}{25}$$

$$\sigma^2 = 28230 - 329,476$$

$$\sigma^2 = -301,246$$

Standard deviation = $\sigma = \sqrt{\frac{\Sigma f_1 x^2}{n} - \left(\frac{\Sigma f x}{n}\right)^2}$

$$\sigma = \sqrt{(-301,246)}$$

$$\sigma = 548.8588$$

Coefficient of variation = $\frac{S.D}{Mean} \times 100$

$$= \frac{548.8588}{45} \times 100$$

$$= 1219.6862\,\%$$

Scanned with CamScanner

Date:

For f_2

Variance $= \sigma^2 = \frac{\sum f_2 x^2}{n} - \left(\frac{\sum f_2 x}{n}\right)^2$

$\sigma^2 = \frac{127,400}{5} - \left(\frac{2600}{5}\right)^2$

$\sigma^2 = \frac{127,400}{5} - \frac{6,760,000}{25}$

$\sigma^2 = 25,480 - 270,400$

$\sigma^2 = -244,920$

Standard deviation $= \sigma = \sqrt{\frac{\sum f_2 x^2}{n} - \left(\frac{\sum f_2 x}{n}\right)^2}$

$\sigma = \sqrt{(-244,920)}$

$\sigma = 494.8939$

Coefficient of variation $= \frac{S.D}{Mean} \times 100$

$= \frac{494.8939}{45} \times 100$

$= 1099.7642\,\%$

Scanned with CamScanner

Date:

Binomial Distribution.

If the probability of each outcome - remains the same throughout the trials, then such trials are called Bernoulli trails and experiment is called Binomial distribution.

$$f(x) = P(X=x) = ({}^nC_x)p^x q^{n-x}, \quad x = 0,1,2,3,\ldots n$$

Q No. 01 :-

A fair coin is tossed 5 times. Find probability of various numbers.

$$x = 0,1,2,3, \ldots n$$

$$\star \; p = \frac{1}{2}$$

$$\star \; q = 1 - p$$

$$q = 1 - \frac{1}{2} \longrightarrow q = \frac{2-1}{2} \Rightarrow \boxed{q = 1/2}$$

i_ $P(X=x) = ({}^nC_x)p^x q^{n-x}$

$$P(X=0) = {}^5C_0 \, p^0 \, q^{n-x}$$

$$= {}^5C_0 (1/2)^0 (1/2)^{5-0}$$

$$= 1 \times 1 \times \left(\frac{1}{2}\right)^5$$

$$= 1 \times \frac{1}{32}$$

$$P(X=0) = \frac{1}{32}$$

Scanned with CamScanner

Date:

ii- $P(X=1) = {}^nC_x\, p^x (q)^{n-x}$

$= {}^5C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{5-1}$

$= 5 \times \frac{1}{2} \times \left(\frac{1}{2}\right)^4$

$= \frac{5}{2} \times \frac{1}{16}$

$= \frac{5}{32}$

iii- $P(X=2) = {}^nC_x\, p^x (q)^{n-x}$

$= {}^5C_2\, P^2 \left(\frac{1}{2}\right)^{5-2}$

$= 10 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^3$

$= 10 \times \frac{1}{4} \times \frac{1}{8}$

$= \frac{10}{32}$

iv- $P(X=3) = {}^nC_x\, p^x (q)^{n-x}$

$= {}^5C_3\, P^3\, q^{5-3}$

$= 10 \times \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$

$= 10 \times \frac{1}{8} \times \frac{1}{4}$

$= \frac{10}{32}$

Date:

v– $P(X=4) = {}^nC_x\, p^x\, q^{n-x}$

$= {}^5C_4 \left(\frac{1}{2}\right)^4 \times \left(\frac{1}{2}\right)^{5-4}$

$= 5 \times \frac{1}{16} \times \left(\frac{1}{2}\right)^1$

$= 5 \times \frac{1}{16} \times \frac{1}{2} \rightarrow \frac{5}{32}$

vi– $P(X=5) = {}^nC_x\, p^x\, q^{n-x}$

$= {}^5C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^{5-5}$

$= 1 \times \frac{1}{32} \times \left(\frac{1}{2}\right)^0$

$= 1 \times \frac{1}{32} \times 1$

$= \frac{1}{32}$

QNo. 02 :-

(a)– Find the probability of getting i– exactly 4 heads (ii) not more than 4 heads when 6 coins are tossed.

i)– exactly 4 heads

$x = 4$, $n = 6$, $p = \frac{1}{2}$

* $q = 1 - p \rightarrow q = 1 - \frac{1}{2} \Rightarrow \frac{2-1}{2} \rightarrow \boxed{q = 1/2}$

$P(X=4) = {}^nC_x\, p^x\, q^{n-x}$

Scanned with CamScanner

Date:

$$= {}^6C_4 \times \left(\frac{1}{2}\right)^4 \times \left(\frac{1}{2}\right)^{6-4}$$

$$= 15 \times \frac{1}{16} \times \left(\frac{1}{2}\right)^2$$

$$= \frac{15}{16} \times \frac{1}{4}$$

$$P(X=4) = \frac{15}{64} \Rightarrow 0.234375$$

ii) - Not more than 4 heads

$$P(X \le 4) = 1 - P(X > 4)$$

$$= 1 - \left[{}^6C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^{6-5} + {}^6C_6 \times \left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^{6-6}\right]$$

$$= 1 - \left[\left(6 \times \frac{1}{32} \times \frac{1}{2}\right) + 1 \times \frac{1}{64} \times 1\right]$$

$$= 1 - \left[\frac{6}{64} + \frac{1}{64}\right]$$

$$= 1 - \left(\frac{6+1}{64}\right)$$

$$= 1 - \frac{7}{64}$$

$$= 1 - 0.109375$$

$$P(x \le 4) = 0.890625$$

(b) - Find the probability of (i) 2 or more heads (ii) fewer than 4 heads in a single toss of 6 for coins.

Scanned with CamScanner

Date:

i) 2 or more heads

$n = 6, \quad p = 1/2, \quad q = 1/2$

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - \left[{}^6C_0 \times \left(\tfrac{1}{2}\right)^0 \times \left(\tfrac{1}{2}\right)^{6-0} + {}^6C_1 \times \left(\tfrac{1}{2}\right)^1 \times \left(\tfrac{1}{2}\right)^{6-1} \right]$$

$$= 1 - \left[\left(1 \times 1 \times \frac{1}{64}\right) + \left(6 \times \frac{1}{2} \times \frac{1}{32}\right) \right]$$

$$= 1 - \left(\frac{1}{64} + \frac{6}{64} \right)$$

$$= 1 - \frac{7}{64} \rightarrow 1 - 0.109375$$

$$P(X \geq 2) = 0.890625$$

ii) fewer than 4 heads

$x = 0, 1, 2, 3$

$$P(X < 4) = \left[P(x=0) + P(x=1) + P(x=2) + P(x=3) \right]$$

$$= \left[{}^6C_0 \times \left(\tfrac{1}{2}\right)^0 \times \left(\tfrac{1}{2}\right)^{6-0} + {}^6C_1 \times \left(\tfrac{1}{2}\right)^1 \times \left(\tfrac{1}{2}\right)^{6-1} + {}^6C_2 \times \left(\tfrac{1}{2}\right)^2 \times \left(\tfrac{1}{2}\right)^{6-2} + {}^6C_3 \times \left(\tfrac{1}{2}\right)^3 \times \left(\tfrac{1}{2}\right)^{6-3} \right]$$

$$= \left[\left(1 \times 1 \times \frac{1}{64}\right) + \left(6 \times \frac{1}{2} \times \frac{1}{32}\right) + \left(15 \times \frac{1}{4} \times \frac{1}{16}\right) + \left(20 \times \frac{1}{8} \times \frac{1}{8}\right) \right]$$

$$= \left[\frac{1}{64} + \frac{6}{64} + \frac{15}{64} + \frac{20}{64} \right]$$

$$= \frac{1 + 6 + 15 + 20}{64}$$

Date:

$$= \frac{42}{64}$$

$$P(X<4) = 0.65625$$

Q No·03 :-

A and B play a game in which A's probability of winning is $2/3$. In a series of 8 games, what is the probability that A will win (i) exactly 4 games, (ii) at least 4 games (iii) 6 or more games.

i)- exactly 4 games

$n = 8$, $p = \frac{2}{3}$, $q = 1 - p \rightarrow q = 1 - \frac{2}{3}$

$$q = \frac{3-2}{3} \rightarrow \boxed{q = 1/3}$$

$$P(X=4) = {}^8C_4 \times \left(\frac{2}{3}\right)^4 \times \left(\frac{1}{3}\right)^{8-4}$$

$$= 70 \times \frac{16}{81} \times \left(\frac{1}{3}\right)^4$$

$$= 70 \times \frac{16}{81} \times \frac{1}{81}$$

$$= \frac{1120}{6561}$$

$$P(X=4) = 0.1707$$

ii)- at least 4 games (means 4 or more)

Scanned with CamScanner

Date:

$P(X \geq 4) = 1 - P(x < 4)$

$= 1 - \left[{}^8C_0 \times \left(\frac{2}{3}\right)^0 \times \left(\frac{1}{3}\right)^{8-0} + {}^8C_1 \left(\frac{2}{3}\right)^1 \times \left(\frac{1}{3}\right)^{8-1} + {}^8C_2 \times \left(\frac{2}{3}\right)^2 \times \left(\frac{1}{3}\right)^{8-2} + {}^8C_3 \left(\frac{2}{3}\right)^3 \times \left(\frac{1}{3}\right)^{8-3} \right]$

$= 1 - \left[\left(1 \times 1 \times \frac{1}{6561}\right) + \left(8 \times \frac{2}{3} \times \frac{1}{2187}\right) + \left(28 \times \frac{4}{9} \times \frac{1}{729}\right) + 56 \times \frac{8}{27} \times \frac{1}{243} \right]$

$= 1 - \left[\frac{1}{6561} + \frac{16}{6561} + \frac{112}{6561} + \frac{448}{6561} \right]$

$= 1 - \left[\frac{1 + 16 + 112 + 448}{6561} \right]$

$= 1 - \frac{577}{6561}$

$= \frac{6561 - 577}{6561}$

$= \frac{5984}{6561}$

$P(X \geq 4) = 0.9120$

iii) 6 or more games.

$P(X \geq 6) = {}^8C_6 \times \left(\frac{2}{3}\right)^6 \times \left(\frac{1}{3}\right)^{8-6} + {}^8C_7 \times \left(\frac{2}{3}\right)^7 \times \left(\frac{1}{3}\right)^{8-7} + {}^8C_8 \times \left(\frac{2}{3}\right)^8 \times \left(\frac{1}{3}\right)^{8-8}$

$= \left(28 \times \frac{64}{729} \times \frac{1}{9}\right) + \left(8 \times \frac{128}{2187} \times \frac{1}{3}\right) + \left(1 \times \frac{256}{6561} \times 1\right)$

$$= \frac{1792}{6561} + \frac{1024}{6561} + \frac{256}{6561}$$

$$= \frac{1792 + 1024 + 256}{6561}$$

$$= \frac{3072}{6561}$$

$$P(X \geq 6) = 0.46822$$

Scanned with CamScanner

Date:

Poisson Distribution :-

Where events occur randomly over a specified interval of time or space or length, this distribution is called poisson distribution.

Formula :-

$$\lim_{\substack{n\to\infty \\ p\to 0}} b(x, n, p) = \frac{\mu^x e^{-\mu}}{x!}, \quad x = 0, 1, 2, 3, \ldots$$

Where $e = 2.71828$

$\mu = np$

Q No. 01 :-

$\mu = 2$

$x = 0, 1, 2, 3,$ or more

For $x = 0$

$$P(x, \mu) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P(0, 2) = \frac{(2)^0 (2.71828)^{-2}}{0!}$$

$$= \frac{1 \times 0.135335}{1}$$

$$P(0, 2) = 0.135335$$

For $x = 1$

$$P(1, 2) = \frac{(2)^1 (2.71828)^{-2}}{1!}$$

$$= \frac{2 \times 0.135335}{1}$$

$$P(1, 2) = 0.27067$$

Scanned with CamScanner

Date:

For $x = 2$

$$p(2,2) = \frac{(2)^2 (2.71828)^{-2}}{2!}$$

$$= \frac{4 \times 0.135335}{2}$$

$$p(2,2) = 0.27067$$

For $x \geq 3$

$$P(x \geq 3) = 1 - P(x < 3)$$

$$= 1 - [P(x=0) + P(x=1) + P(x=2)]$$

$$= 1 - [0.135335 + 0.27067 + 0.27067]$$

$$= 1 - 0.682705$$

$$P(x \geq 3) = 0.317295$$

Q NO. 02 :-

If X is a poisson random variable with parameter $n = 200$, $p = 0.01$. Find possibilities of $x = 0, 1, 2, 3$.

$$\mu = np$$

$$\mu = 200 \times 0.01$$

$$\mu = 2$$

For $x = 0$

$$p(x, \mu) = \frac{\mu^x e^{-\mu}}{x!}$$

$$p(0, 2) = \frac{(2)^0 \times (2.71828)^{-2}}{0!}$$

$$= \frac{1 \times 0.135335}{1}$$

Date:

$p(0,2) = 0.135335$

For $x = 1$

$$p(1,2) = \frac{(2)^1 \times (2.17828)^{-2}}{1!}$$

$$= \frac{2 \times 0.135335}{1}$$

$p(1,2) = 0.27067$

For $x = 2$

$$p(2,2) = \frac{(2)^2 \times (2.17828)^{-2}}{2!}$$

$$= \frac{4 \times 0.135335}{2}$$

$p(2,2) = 0.27067$

For $x = 3$

$$p(3,2) = \frac{(2)^3 \times (2.17828)^{-2}}{3!}$$

$$= \frac{8 \times 0.135335}{6}$$

$$= \frac{1.08268}{6}$$

$p(3,2) = 0.180446$

Q No. 03.

An event has probability $p = 3/8$. Find complete binomial distribution $n = 5$.

Date:

$p = \frac{3}{8}$, $q = 1 - p$

$q = 1 - \frac{3}{8}$

$q = \frac{8-3}{8} \rightarrow \boxed{q = \frac{5}{8}}$

$x = 0, 1, 2, 3, 4, 5$

For $x = 0$

$P(X=0) = {}^5C_0 \left(\frac{3}{8}\right)^0 \left(\frac{5}{8}\right)^{5-0}$

$= 1 \times 1 \times \frac{3125}{32768}$

$= \frac{3125}{32768}$

For $x = 1$

$P(X=1) = {}^5C_1 \left(\frac{3}{8}\right)^1 \left(\frac{5}{8}\right)^{5-1}$

$= 5 \times \frac{3}{8} \times \frac{625}{4096}$

$= \frac{9375}{32768}$

For $x = 2$

$P(X=2) = {}^5C_2 \left(\frac{3}{8}\right)^2 \left(\frac{5}{8}\right)^{5-2}$

$= 10 \times \frac{9}{64} \times \frac{125}{512}$

$= \frac{11250}{32768}$

Date:

For $x = 3$

$P(X=3) = {}^5C_3 \left(\frac{3}{8}\right)^3 \left(\frac{5}{8}\right)^{5-3}$

$= 10 \times \frac{27}{512} \times \frac{25}{64}$

$= \frac{6750}{32768}$

For $x = 4$

$P(X=4) = {}^5C_4 \left(\frac{3}{8}\right)^4 \left(\frac{5}{8}\right)^{5-4}$

$= 5 \times \frac{81}{4096} \times \frac{5}{8}$

$= \frac{2025}{32768}$

For $x = 5$

$P(X=5) = {}^5C_5 \left(\frac{3}{8}\right)^5 \left(\frac{5}{8}\right)^{5-5}$

$= 1 \times \frac{243}{32768} \times 1$

$= \frac{243}{32768}$

QNo. 03:-

If X is a random variable with parameter $\mu = 3$, Find probability $x = 3, 4, 5$ or more.

$\mu = 3$

$x = 3, 4, 5$ or more.

Date:

For $x = 3$

$$P(x,\mu) = \frac{\mu^{x} e^{-\mu}}{x!}$$

$$P(3,3) = \frac{(3)^{3}(2.71828)^{-3}}{3!}$$

$$= \frac{27 \times 0.049787}{6}$$

$$= \frac{1.34425}{6}$$

$$P(3,3) = 0.2240415$$

For $x = 4$

$$P(4,3) = \frac{(3)^{4}(2.71828)^{-3}}{4!}$$

$$= \frac{81 \times 0.049787}{24}$$

$$= \frac{4.03276}{24}$$

$$P(4,3) = 0.16803$$

For $x \geq 5$

$$P(x \geq 5) = 1 - [P < 5]$$

$$= 1 - [P(x=3) + P(x=4)]$$

$$= 1 - [0.2240415 + 0.16803]$$

$$= 1 - 0.39203$$

$$P(x \geq 5) = 0.60797$$

Q No. 04:

Scanned with CamScanner

Date:

If X is a random variable with parameter n= 100, p= 0.01. Find probability x= 2, 3, 4 and 5.

$$\mu = np$$
$$= (100)(0.01)$$
$$\mu = 1$$

For x=2

$$P(x, \mu) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P(2, 1) = \frac{(1)^2 (2.71828)^{-1}}{2!}$$

$$= \frac{1 \times 0.367879}{2}$$

$$P(2, 1) = 0.1839395$$

For x=3

$$P(3, 1) = \frac{(1)^3 (2.71828)^{-1}}{3!}$$

$$= \frac{(1)(0.367879)}{6}$$

$$= 0.0613$$

For x=4

$$P(4, 1) = \frac{(1)^4 (2.71828)^{-1}}{4!}$$

$$= \frac{(1)(0.367879)}{24}$$

$$= 0.015375$$

For x=5

Scanned with CamScanner

Date:

$$P(5,1) = \frac{(1)^5(2.71828)^{-1}}{5!}$$

$$= \frac{(1)(0.367879)}{120}$$

$$= 0.003065$$

Q NO. 05 :-

$x = 0, 1, 2, 3, 4$, $n = 100$, $p = .01$

$q = 1 - P \Rightarrow q = 1 - .001 \Rightarrow q = 0.99$

By Binomial Distribution :,

$P(X=x) = {}^nC_x \, p^x q^{n-x}$

For $x = 0$

$P(X=0) = {}^{100}C_0 (0.01)^0 (0.99)^{100-0}$

$= 1 \times 1 \times 0.36603$

$= 0.36603$

For $x = 1$

$P(X=1) = {}^{100}C_1 (0.01)^1 (0.99)^{100-1}$

$= 100 \times 0.01 \times (0.99)^{99}$

$= 0.3697$

For $x = 2$

$P(X=2) = {}^{100}C_2 (0.01)^2 (0.99)^{100-2}$

$= 4950 \times 0.0001 \times (0.99)^{98}$

$= 4950 \times 0.0001 \times 0.347346$

$= 0.18486$

For $x = 3$

$P(X=3) = {}^{100}C_3 (0.01)^3 (0.99)^{100-3}$

$= 161.700 \times 0.000001 \times (0.99)^{97}$

$= 0.1617 \times 0.37723$

$= 0.06099$

For $x = 4$

$P(X=4) = {}^{100}C_4 (0.01)^4 (0.99)^{100-4}$

$= 3921225 \times 0.00000001 \times (0.99)^{96}$

$= 3921225 \times 0.00000001 \times 0.38164$

$= 0.0149$

By Poisson Distribution :,

$P(x,\mu) = \frac{\mu^x e^{-\mu}}{x!}$, $\mu = np \Rightarrow \mu = 100 \times 0.01 \Rightarrow \boxed{\mu = 1}$

Date:

For $x=0$

$$P(0,1) = \frac{(1)^0 (2.71828)^{-1}}{0!}$$

$$= \frac{1 \times 0.36780}{1}$$

$$P(0,1) = 0.36780$$

For $x=1$

$$P(1,1) = \frac{(1)^1 (2.71828)^{-1}}{1!}$$

$$= \frac{1 \times 0.36780}{1}$$

$$P(1,1) = 0.36780$$

For $x=2$

$$P(2,1) = \frac{(1)^2 (2.71828)^{-1}}{2!}$$

$$= \frac{1 \times 0.36780}{2}$$

$$P(2,1) = 0.183989$$

For $x=3$

$$P(3,1) = \frac{(1)^3 (2.71828)^{-1}}{3!}$$

$$= \frac{1 \times 0.36780}{6}$$

$$P(3,1) = 0.0613$$

For $x=4$

$$P(4,1) = \frac{(1)^4 (2.71828)^{-1}}{4!}$$

$$= \frac{1 \times 0.36780}{24}$$

$$P(4,1) = 0.015325$$

x	Binomial Distribution	Poisson Distribution
0	0.36603	0.36780
1	0.3697	0.36780
2	0.18486	0.183989
3	0.06099	0.0613
4	0.0149	0.015325

Normal Distribution:-

Normal distribution, also known as Gaussian distribution is a probability distribution that is symmetric about the mean showing that data near the mean are more frequent in occurence, then data far from the mean. In a graphical form, the normal distribution appears as a "bell curve".

Formula:-

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} \qquad -\infty < x < +\infty$$

$\therefore$ Random variable = $z = \frac{x-\mu}{\sigma}$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}$$

$$f(x) = \frac{1}{\sigma\sqrt{2x}} e^{-\frac{z^2}{2}}$$

Q No. 01 :-

$x = 75$, $\sigma = 1.0$, $\mu = 3.5$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

$$f(75) = \frac{1}{1.0\sqrt{2\pi}} e^{-\left[\frac{(75-3.5)^2}{2(1)^2}\right]}$$

Scanned with CamScanner

Date:

$$f(x) = \frac{1}{1.0(2.5066)} e^{-\left[\frac{(71.5)^2}{2}\right]}$$

$$f(x) = \frac{1}{2.5066} e^{-\left(\frac{5112.25}{2}\right)}$$

$$f(x) = 0.3989 \times e^{-2556.125}$$

$$f(x) = 0.3989 \times (2.71828)^{-2556.125}$$

$$f(x) = 0.3989 \times 0$$

$$f(x) = 0$$

$$f(75) = 0$$

QNo.02:-

$x = 60$, $\sigma = 2.5$, $\mu = 4.0$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

$$f(60) = \frac{1}{2.5(2.5066)} e^{-\left[\frac{(60-4)^2}{2(2.5)^2}\right]}$$

$$f(60) = \frac{1}{6.2665} e^{-\left[\frac{(56)^2}{2(6.25)}\right]}$$

$$f(60) = \frac{1}{6.2665} e^{-\left[\frac{3136}{12.5}\right]}$$

$$f(60) = 0.15957 \times e^{-250.88}$$

$$f(6) = 0.15957 \times (2.71828)^{-250.88}$$

$$f(60) = 0.15957 \times 0$$

$$f(60) = 0$$

Scanned with CamScanner

Principles/Designs of Experiments:

The basic principles of experimental design are:

1) Randomization
2) Replication
3) Local control

Chi-square Distribution (χ^2)

A χ^2 random variable is defined as the sum of squares of n Independent standard normal random variables is called χ^2.

Formula:

$$f(\chi^2) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} (\chi^2)^{(n/2)-1} e^{-\chi^2/2}$$

Testing of hypothesis about attributes.

B \ A	A_1	A_2	A_3
B_1	O	O	O
B_2	O	O	O
B_3	O	O	O

→ rows

Frequency cells

columns

Steps of χ^2

null hypothesis   Alternative hypothesis

- Hypothesis → 1) H_0 (2) H_1
- Level of significance
  $\alpha = 0.05$
- Test statistics
  $$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

Scanned with CamScanner

Date:

- Critical region

$$\chi^2_{cal} \geq \chi^2_{\alpha\,(d.f)}$$

α → alpha

- Calculation
- Conclusion

→ d.f = (r - 1) (c - 1)

→ f_o = observed frequency

→ f_e = Expected frequency

→ $f_e = \dfrac{R_T \cdot C_T}{CrT}$

row → R_T, C_T → column

Q No. 01:~

	Math	Chemistry	Physics	Total
Music	24	83	17	124
Craftswork	11	62	28	101
Reading	32	121	34	187
Drama	10	26	44	80
Total	77	292	123	492

Solution ;~

1- Hypothesis

$H_0 = 0$

$H_1 \neq 0$

2- Level of significance

$\alpha = 0.05$

3- Test statistics

Scanned with CamScanner

Date:

$$\chi^2 = \frac{\Sigma (f_o - f_e)^2}{f_e}$$

4- Critical Region

$$\chi^2_{cal} \geq \chi^2_{\alpha(d.f)}$$

$$d.f = (R-1)(C-1) = (4-1)(3-1) = (3)(2) = 6$$

$$\chi^2_{cal} \geq \chi^2_{(0.05)6}$$

$$\chi^2_{Tab} = 12.59$$

5- Calculation

Fo	$f_e = \frac{R_T - C_T}{Cr_T}$	$(f_o - f_e)^2$	$(f_o - f_e)^2 / f_e$
24	$\frac{124 \times 77}{492} = 19.4$	21.16	1.09
11	15.8	23.04	1.46
32	29.3	7.29	0.29
10	12.5	6.25	0.50
83	73.6	88.36	1.20
62	59.9	4.41	0.07
121	11.10	100	0.90
26	47.5	462.25	9.73
17	31.0	196	6.32
28	25.3	7.29	0.29
34	46.7	161.29	3.45
44	20.0	576	28.80

$$\Sigma \frac{(f_o - f_e)^2}{f_e} = 54.04$$

Date:

$$\chi^2_{cal} = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

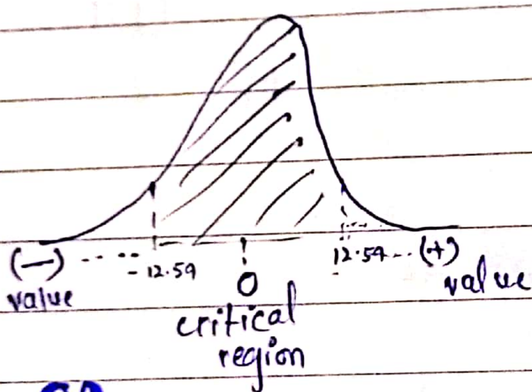
$$\chi^2_{cal} = 54.04$$

$$\chi^2_{cal} \geq \chi^2_{tab}$$

$$54.04 > 12.59$$

6- Conclusion :-

Hence, χ^2_{cal} is greater than χ^2_{tab}. It falls in critical region, then we reject H_o.



Q No. 02:-

Find χ^2 and test whether the two attributes are independent.

Let $\alpha = 0.05$.

Attributes	A_1	A_2	A_3	Total
B_1	215	325	60	600
B_2	135	175	90	400
Total	350	500	150	1000

Solution :-

1- Hypothesis

$H_o = 0$

$H_1 \neq 0$

Scanned with CamScanner

Date:

2- Level of significance.

$$\alpha = 0.05$$

3- Test Statistics

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

4- Critical Region

$$\chi^2_{cal} \geq \chi^2_{tab}$$

$$d.f = (R-1)(C-1)$$
$$= (2-1)(3-1)$$
$$= (1)(2)$$
$$d.f = 2$$

$$\chi^2_{cal} \geq \chi^2_{\alpha(d.f)}$$

$$\chi^2_{cal} \geq \chi^2_{(0.05)(2)}$$

$$\chi^2_{tab} = 5.99$$

5- Calculation :-

f_o	$f_e = RT \cdot CT / GT$	$(f_o - f_e)^2$	$(f_o - f_e)^2 / f_e$
215	$600 \times 350 / 1000 = 210$	25	0.11
135	$400 \times 350 / 1000 = 140$	25	0.17
325	$600 \times 500 / 1000 = 300$	625	2.08
175	$400 \times 500 / 1000 = 200$	625	3.13
60	$600 \times 150 / 1000 = 90$	900	10
90	$400 \times 150 / 1000 = 60$	900	15

$$\Sigma \frac{(f_o - f_e)^2}{f_e} = 30.49$$

$$\chi^2_{cal} = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

$$\chi^2_{cal} = 30.49$$

Scanned with CamScanner

Date:

$$\chi^2_{cal} > \chi^2_{tab}$$

$$30.49 > 5.99$$

6- Conclusion :-

Hence, χ^2_{cal} is greater than χ^2_{tab}. It falls in the critical region, hence we reject H_0.

Q No. 03:-

Test the null hypothesis that the two variables of classification are independ, using 0.05 level of significance

Classes	A_1	A_2	A_3	Total
B_1	337	291	302	930
B_2	225	207	238	670
Total	562	498	540	1600

Solution :-

1- Hypothesis

$H_0 = 0$

$H_1 \neq 0$

2- Level of significance

$\alpha = 0.05$

3- Test Statistics

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

4- Critical region

$$\chi^2_{cal} \geq \chi^2_{tab}$$

$$\chi^2_{cal} \geq \chi^2_{\alpha(d.f)}$$

$\therefore$ $d.f = (R-1)(C-1)$

$= (2-1)(3-1)$

$d.f = 2$

Scanned with CamScanner

Date:

$$\chi^2_{cal} \geq \chi^2_{(0.05)(2)}$$

$$\chi^2_{tab} = 5.99$$

5- Calculation

f_o	$f_e = R_T \cdot C_T / C_{rT}$	$(f_o - f_e)^2$	$(f_o - f_e)^2 / f_e$
337	$930 \times 562 / 1600 = 326.6$	108.16	0.33
225	$670 \times 562 / 1600 = 235.3$	106.09	0.45
291	$930 \times 498 / 1600 = 289.4$	2.56	0.008
207	$670 \times 498 / 1600 = 208.5$	2.25	0.01
302	$930 \times 540 / 1600 = 313.8$	139.24	0.44
238	$670 \times 540 / 1600 = 226.1$	141.61	0.63

$$\sum \frac{(f_o - f_e)^2}{f_e} = 1.868$$

$$\chi^2_{cal} = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$\chi^2_{cal} = 1.868$$

$$\chi^2_{cal} \geq \chi^2_{tab}$$

$$\chi^2_{cal} < \chi^2_{tab}$$

$$1.868 < 5.99$$

6- Conclusion

Hence, χ^2_{cal} is smaller than χ^2_{tab}. It do not fall do in critical region, Hence we do not reject H_o.

Date:

Student's t-test

A t-test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process actually has an effect on the population of interest, or whether two groups are different from one another.

Formula:~

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\therefore\ sp = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$$

Q No. 01:-

$n_1 = 12$	$n_2 = 18$
$\bar{x}_1 = 10$	$\bar{x}_2 = 25$
$s_1^2 = 1200$	$s_2^2 = 900$

Solution :~

1- Hypothesis

$H_0 : \mu_2 - \mu_1 = 10$

$H_1 : \mu_2 - \mu_1 > 10$

2- Level of significance

$\alpha = 0.05$

3- Test Statistics

Date:

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

4- Critical region

$t_{cal} > t_{tab}$

$t_{cal} > t_{\alpha(d.f)}$

$t_{cal} > t_{0.05\,(28)}$

$\therefore d.f = n_1 + n_2 - 2$

$= 12 + 18 - 2$

$= 30 - 2$

$= 28$

$t_{tab} = 1.701$

5- Calculation

$$sp = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{(12-1)\,1200 + (18-1)\,900}{28}}$$

$$= \sqrt{\frac{(11)\,1200 + 17\,(900)}{28}}$$

$$= \sqrt{\frac{13200 + 15300}{28}}$$

$$= \sqrt{1017.8571}$$

$sp = 31.90$

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$t = \frac{(25-10) - 10}{31.90\sqrt{\frac{1}{12} + \frac{1}{18}}}$$

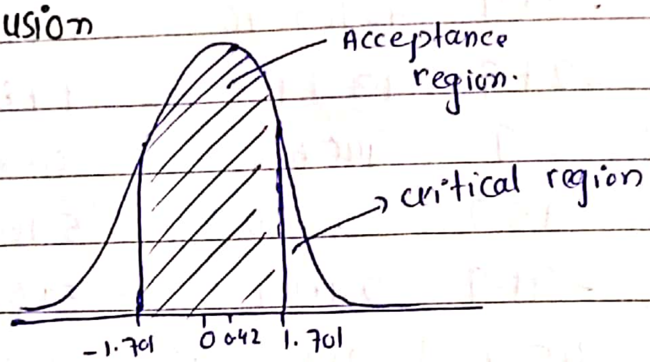
Date:

$$t = \frac{15 - 10}{31.90\sqrt{0.138}}$$

$$t = \frac{5}{11.889}$$

$$t = 0.42$$

6- Conclusion



Since, $t = 0.42$ does not fall in critical region, Hence we do not reject H_0.

Q No. 02:-

The weights in grams of 10 male and 10 female juvenile ring-necked pheasants are :

Male : 1293, 1380, 1614, 1497, 1340, 1643, 1466, 1627, 1383, 1711

Female : 1061, 1065, 1092, 1017, 1021, 1138, 1143, 1094, 1270, 1028

Test the hypothesis of a difference of 350 grams between population means in favour of males against alternative of a greater difference, using 0.05 level of significance. Assume that the weights are normally distributed.

ate:

$n_1 = 10$, $n_2 = 10$, $\alpha = 0.05$

$\bar{x}_1 = ?$, $\bar{x}_2 = ?$, $s_1^2 = ?$

$s_2^2 = ?$, $sp = ?$, $t = ?$

Solution:-

X_1	X_2	$X_1-\bar{X}_1$	$X_2-\bar{X}_2$	$(X_1-\bar{X}_1)^2$	$(X_2-\bar{X}_2)^2$
1293	1061	-202.4	-31.9	40965.76	1017.61
1380	1065	-115.4	-27.9	13317.16	778.41
1614	1092	118.6	-0.9	14065.96	0.81
1497	1017	1.6	-75.9	2.56	5760.81
1340	1021	-155.4	-71.9	24149.16	5169.61
1643	1138	147.6	45.1	21785.76	2034.01
1466	1143	-29.4	50.1	864.36	2510.01
1627	1094	131.6	1.1	17318.56	1.21
1383	1270	-112.4	177.1	12633.76	31364.41
1711	1028	215.6	-64.9	46483.36	4212.01
$\Sigma X_1 = 14954$	$\Sigma X_2 = 10929$			$\Sigma(X_1-\bar{X}_1)^2 = 191586.4$	$\Sigma(X_2-\bar{X}_2)^2 = 52848.9$

1- Hypothesis

$H_0 : \mu_1 - \mu_2 = 350$

$H_1 : \mu_1 - \mu_2$

2- Level of significance

$\alpha = 0.05$

3- Test Statistics

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

4- Critical Region

$\therefore \bar{X}_1 = \frac{\Sigma X_1}{n} \rightarrow \frac{14954}{10} \rightarrow \boxed{1495.4}$

$\therefore \bar{X}_2 = \frac{\Sigma X_2}{n} \rightarrow \frac{10929}{10} \rightarrow \boxed{1092.9}$

$\therefore S_1^2 = \frac{\Sigma(X_1-\bar{X}_1)^2}{n} \rightarrow \frac{191586.4}{10}$

$= 19158.64$

$\therefore S_2^2 = \frac{\Sigma(X_2-\bar{X}_2)^2}{n} \rightarrow \frac{52848.9}{10}$

$= 5284.89$

Scanned with CamScanner

Date:

$t_{cal} \geq t_{tab}$

$t_{cal} \geq t_{\alpha(d.f)}$

$t_{cal} \geq t_{0.05(18)}$

$t_{tab} = 1.734$

$\therefore df = n_1 + n_2 - 2$

$= 10 + 10 - 2 \Rightarrow 20 - 2$

$= 18$

5- Calculation

$$sp = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)^2 s_2}{n_1+n_2-2}}$$

$$sp = \sqrt{\frac{(10-1)\,19158.64 + (10-1)\,5284.89}{18}}$$

$$sp = \sqrt{\frac{(9)\,19158.64 + (9)\,5284.89}{18}}$$

$$sp = \sqrt{\frac{172427.76 + 47564.01}{18}}$$

$$sp = \sqrt{\frac{219,991.77}{18}}$$

$$sp = \sqrt{12221.765}$$

$$sp = 110.5$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$t = \frac{(1495.4 - 1092.9) - 350}{110.5\sqrt{\frac{1}{10} + \frac{1}{10}}}$$

Date:

$$t = \frac{402.5 - 350}{110.5\sqrt{\frac{2}{10}}}$$

$$t = \frac{402.5 - 350}{110.5\sqrt{0.2}}$$

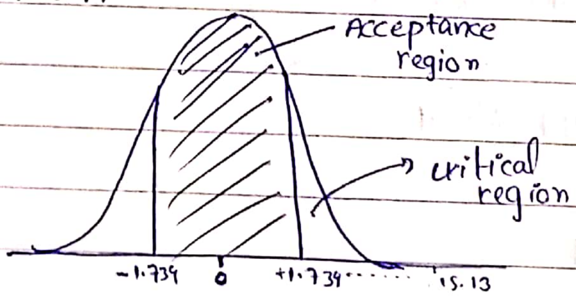
$$t = \frac{752.5}{110.5(0.45)}$$

$$t = \frac{752.5}{49.725} \longrightarrow t = 15.13$$

$$t_{cal} > t_{tab}$$

$$15.13 > 1.734$$

6- Conclusion



Since, $t = 15.13$ fall in the critical region, Hence we reject H_0.