[This question paper contains 2 Printed Pages]

Royal Education Society's

COCSIT LATUR

FACULTY OF COMPUTER STUDIES

[CBCS PATTERN]

B.C.A. FY (SEM I)

Model Question Paper (November- 2022)

Subject – Element of Statistics

Date: / /

Time: Three Hours ***Maximum Marks - 75***

Instructions to the candidates:

1. All questions are Compulsory.

2. Figures to the right indicate full marks.

3. Assume suitable data, if required.

Q.1 Attempt any FIVE of the following (3 Marks each) 15

a) Write Merits and Demerits of the Mode ?

b) Define Statistics by according to Webster's and Secrists .

c) Calculate the range and coefficient of range of the following data.

10, 20, 15, 5, 25, 32, 45, 30.

d) Explain Correlation.

e) Write a note on primary and secondary data .

f) Define probability.

g) Calculate the median from the following data

22, 18 , 6, 10, 4, 3, 9 .

Q. 2 Attempt any three of the following (5 Marks each) 15

a) From the following data , find the mean

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students	6	5	8	15	7	6	3

b) Explain Scope of statistics.

c) Explain Dispersion in detail.

d) Prove that the Addition theorem of two events

$P(A\ B) = P(A) + P(B) - P(A\ B)$

e) Calculate Correlation Coefficient between X and Y of the Data

X	7	2	6	8	4	3
Y	3	8	7	9	5	4

P. T .O.

Q. 3 Attempt any three of the following (5 Marks each) 15

a) Explain advantages of statistics.

b) Explain types of Correlation in detail.

c) Calculate the Standard deviation of the following data

Values	15	24	52	34	40	16
Frequency	3	8	14	10	5	3

d) Explain Permutation and Combination.

e) Find the means of x and y from two regression line is

$2X + 5Y + 8 = 0$, $6X + 3Y + 6 = 0$.

Q. 4 Attempt any three of the following (5 Marks each) 15

a) Define events and explain its types .

b) Explain regression equation?

c) Calculate the mode from the following data :

Wages	200-250	250-300	300-350	350-400	400-450	450-500
Workers	3	5	9	12	8	4

d) Find Variance from the following data :

Class	0-20	20-40	40-60	60-80	80-100
Frequency	5	16	20	15	10

e) If P (A) = 0.6 P(B) = 0.5 , ~~P(AB)~~ = P(A∩B) = 0.3

Compute

$P(A')$, $P(A \cup B)$, $P(A \cap B')$, $P(A' \cap B')$, $P(A' \cup B')$.

Q .5 Attempt any three of the following (5 Marks each) 15

a) Explain Limitation of Statistics .

b) Calculate the mean and median from the following data

Class	300-500	500-700	700-900	900-1100	1100-1300	1300-1500
Frequency	25	55	30	20	14	6

c) A bag contain 30 balls numbered from 1 to 30. One ball is drawn at random . find the probability that the numbered ball drawn will be multiple of i) 5 or 7

ii) 3 or 7 .

d) Obtain regression equation of x on y from following data

X	7	2	5	8	4
Y	3	8	7	9	5

e) Explain Scatter diagram method in detail.

******

Q.4)

e) If $P(A) = 0.6$, $P(B) = 0.5$, $P(A \cap B) = 0.3$

Compute.

$P(A')$, $P(A \cup B)$, $P(A \cap B')$, $P(A' \cap B')$, $P(A' \cup B')$.

Page No.:
Date. / /

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

Model Question Paper - Answer key.

Class - BCA - F.Y. (sem-I) Sub - Element of Statistics.

Q.1) Attempt any Five of the following. (3-Mark each) 15-M.

a) Write merits and De-merits of the mode ?

→ Merits of the mode -

i) It is easy to calculate and simple to understand.
ii) It is applicable for qualitative as well as quantitative data.
iii) It is unaffected for extreme observation
iv) It is graphically calculated.

De-merits of the Mode -

i) It is not rigidly defined.
ii) It is not applicable for further mathematical treatment.
iii) It is not based on all observation.

b) Define statistics by according to Webster's and secrists ?

→ By according Websters following defination of statistics

The classified facts representing the condition of people in the specially those fact which can be stated in tables or tables of numbers or any tabular form is called as statistics.

By according to secrists defination of statistics

By the statistics we mean aggregates of facts affected to a mark extend by multiplicity causes numerically expressed or estimated according to reasonable standard accuracy and placed in a relation to each other is called as statistics.

c) Calculate the range and coefficient of range from following data.

10, 20, 15, 5, 25, 32, 45, 30.

→

10, 20, 15, 5, 25, 32, 45, 30.

Smallest number = 5

largest number = 45.

Range = ~~Smallest number~~ - ~~largest number~~

Range = Largest number - Smallest number

= 45 - 5

= 40.

Coefficient of range = $\frac{L - S}{L + S}$.

= $\frac{45 - 5}{45 + 5}$

= $\frac{40}{50}$

Coefficient of range = 0.8.

Page No.:
Date. / /

d) Explain correlation.

→ Many a times we come a cross the situation where two variables are interrelated such as the height of the son that of the father, and Income and expenditure of the family, etc. in this situtation we study the two variables.

Correlation analysis deals with association between two or more than two variables then it is said to be correlation. and the measure of correlation is called as the correlation coefficient or coefficient of correlation and it is denoted by r. and its value is -1 to 1.

Some types of correlation are given below such as.

i) positive correlation
ii) Negative correlation.
iii) simple correlation.
iv) Multiple or partial correlation.
v) Linear correlation.
vi) Non-linear correlation.

e) Write a note on primary and secondary data.

→ The information Collected through Serveys, experiments and censues in routien manner is called as collection of data. Data mean information the important data in statistics is collection of statistical data. Such as the following type is

i) Primary data
ii) Secondary data

i) Primary data -

It is data collected by particular person for his own use the primary sources like as the interviews and experiment.

ii) Secondary data -

It is data collected by some other person for their own use but the investigator also gets its for this use like as newspaper, journals, and books, etc.

f) Define probability.

If a random experiment or a trial results in 'n' exhaustive mutually exclusive and equally likely outcome out of which m are favourable to occurence of event A, then the probability 'p' of occurrence of happening of A usually denoted by P(A) is given by

$$P = P(A) = \frac{\text{Number of favourable cases}}{\text{Total number of exhaustive event}}$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{m}{n}$$

g) Calculate median for following data.

22, 18, 6, 10, 4, 3, 9.

→ Arranging in ascending order

3, 4, 6, 9, 10, 18, 22.

No. of observation $n = 7$ (odd number)

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{term} = \frac{7+1}{2} = \frac{8}{2} = 4\text{th term}.$$

Ans = median = 9.

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

Page No.:
Date. / /

Q.2) Attempt any three of the following (15-M)

a) From the following data, find the mean.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Student	6	5	8	15	7	6	3

Marks	No. of Student (f)	Mid Point (x)	$x_i \cdot f_i$
0-10	6	5	30
10-20	5	15	75
20-30	8	25	200
30-40	15	35	525
40-50	7	45	315
50-60	6	55	330
60-70	3	65	195

$N = \Sigma f_i = 50$ $\quad \Sigma x_i f_i = 1670$

$$\text{Mean} = \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{\Sigma f_i x_i}{N}$$

$$\bar{x} = \frac{1670}{50}$$

$$\text{Mean} = \bar{x} = 33.4$$

b) Explain scope of statistics.

→ The use of statistics indispensable in the agriculture, business, economics, education, government agencies, industries, social sciences, biological sciences, medical sciences, managment sciences, etc.

i) Statistics and Computer Science -

Development of computer technology have enabled statistics to integrate their models into information systems and thus make statistics a part of decision of many organization.

Statistics and computer science both are together useful in providing solution to the problems in various fields. Statistical software packages like MINITAB, MATHLAB, SAS, etc used for data analysis.

ii) Statistics and Industry -

In many industries, statistical quality control division is operating. Industry makes use of several places such as administration, planning, production, growth and development. Manufactured goods possess a desirable standard or not are examined using various control charts.

iii) Statistics and Economics -

Statistical data and Techniques of statistical analysis useful in solving a vanity of economics problems such as wages, prices, consumption

Page No.:
Date. / /

production, probability distribution of income and wealth.

c) Explain dispersion in detail.

→ In earlier we have learn about the measure of central Tendancy namely as mean, median, mode. Such an average tell us only about the central part of the data. but it does not give any information spread about the data.

The degree to which numberical data tends to about an average value is called as dispersion. following terms are the measure of dispersion.

Measure of dispersion -.
i) Range
ii) Standard deviation.
iii) Mean deviation.
iv) variance.

Range -
it is difference between largest observation to smallest observation.
Range = L − S.

Coefficient of range = $\frac{L-S}{L+S}$.

Standard deviation - (S.D.) or σ^2.
The positive square root of mean of the square of deviation taken from airthmetic mean is called as standard deviation. denoted by σ.

Page No.:
Date. / /

Standard deviation for ungrouped data.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

Where n = no. of observation.

Standard deviation for grouped data.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{N}}$$

where $N = \sum f_i$

Coefficient of Variation. (C.V.)

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

---

d) Prove that the addition theorem of two events.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

→ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Let, $n(S) = n$ Total number of distinct possiable outcome in the sample space (S).

$n(A) = x$ Number of favourable outcome for the event A.

$n(B) = y$ Number of favourable outcome for the event B.

$n(A \cap B) = z$ Number of favourable outcome for both event A and B.

Page No.:
Date. / /

Paper No.....Subject.....
Date.....
Invigilators Sign.....

$$P(A) = \frac{n(A)}{n(S)} = \frac{x}{n},$$

$$P(B) = \frac{n(B)}{n(S)} = \frac{y}{n},$$

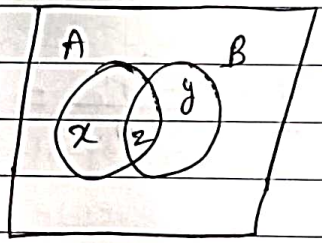
$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{z}{n},$$

as all outcomes are equally likely.

$n(A \cup B) = x - z + z + y - z.$

$= x - \cancel{z} + \cancel{z} + y - z.$

$= x + y - z$



$n(A \cup B) = n(A) + n(B) - n(A \cap B)$

dividing both side by $n(S)$

$$\frac{n(A \cup B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A \cap B)}{n(S)}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

hence proved that.

e) Calculate Correlation coefficient between x and y of the following data.

x	7	2	5	8	4	3
y	3	8	7	9	5	4.

→

X	Y	$(x-\bar{x})=dx$	$(y-\bar{y})=dy$	$dx \cdot dy$	dx^2	dy^2
7	3	7-5 = 2	3-6 = -3	-6	4	9
2	8	2-5 = -3	8-6 = 2	-6	9	4
6	7	6-5 = 1	7-6 = 1	1	1	1
8	9	8-5 = 3	9-6 = 3	6	9	9
4	5	4-5 = -1	5-6 = -1	1	1	1
3	4	3-5 = -2	4-6 = -2	4	4	4
				$\Sigma dx dy = 0$	$\Sigma dx^2 = 28$	$\Sigma dy^2 = 28$

$$\bar{x} = \frac{\Sigma x_i}{n.} = \frac{7+2+5+8+4+3}{6} = \frac{30}{6} = 5$$

$$\bar{y} = \frac{\Sigma y_i}{n.} = \frac{3+8+7+9+5+4}{6} = \frac{36}{6} = 6$$

$\Sigma dx.dy = 0$, $\Sigma dx^2 = 28$ $\Sigma dy^2 = 28$

$$r = \frac{\Sigma dx \cdot dy}{\sqrt{\Sigma dx^2 \times \Sigma dy^2}}$$

$$r = \frac{0}{\sqrt{28 \times 28}}$$

$$r = \frac{0}{\sqrt{784}} = \frac{0}{28}$$

$$r = 0.$$

There is no correlation between x and Y.

Q.3) Attempt any three of the following. (15-M)

a) Explain advantages of statistics.

→

→ i) Statistics are treated as a branch of science dealing with collection, presentation, analysis, and interpretation of data.

ii) Statistics helps in presenting large quantity of data in simple and classified form.

iii) It facilities several function apart from summarization.

iv) It gives the method of comparison of data.

v) It helps in planning, controlling, decision making etc.

vi) Estimation, prediction is also possiable using Statistical tools.

vii) Inter relation between two or more variable can be measured using statistical techniques.

viii) It proves useful in a number of fields like Railway, Banks, Army, etc.

ix) It is important because considerable amount of time, money, and manpower can be saved.

b) Explain types of correlation in detail.

→ Some types of correlation are defined as following

on the bases of direction of variables following correlation are defined.

i) Positive Correlation ii) Negative Correlatio

i) Positive correlation -

Increase in value of one variable is associated with increase in value of other variable or decrease in value of one variable is associated with decrease in value of other variable. Correlation between these variable is said to positive.

Ex:- Height and weight of the person.

ii) Negative correlation -

Increase in value of one variable is associated by decrease in value of other variable or vice versa. Correlation between these variable is said to be negative.

Ex:-

Demand and price of the product.

on the bases of no. of variables following types of correlation are defined as

i) Simple correlation
ii) multiple or partial correlation.

i) Simple correlation -

When only two variables are studied it is problem of Simple correlation.

Ex:- Income and expenditure of the family.

ii) Multiple or Partial correlation -

When ~~two~~ three or more than three variables are

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

studied at simultaneously then the correlation is said to be multiple correlation.

Ex:-

when we study the relationship between the yield of price per acre and both the amount of rainfall and amount of fertilizers are used.

on the bases of constancy of ratio of change between the variables are following types defined

i) Linear correlation -
ii) Non-Linear correlation.

i) Linear correlation.

If the amount of change in one variable tends to same constant ratio to the amount of change in the other variable. said to linear correlation.

Ex:

x	10	20	30	40	50
y	70	140	210	280	350.

$x : y = 1 : 7.$

Ratio betn x and y are the same.

ii) Non-Linear correlation.

If the amount of change in one variable tends to different in constant ratio to amount of change in other variable then correlation is non-linear.

Ex:- If we double the amount of rainfall the production of rice or wheat etc would not necessarily doubled

Page No.:
Date. / /

4) Calculate standard deviation of the following data

Values	15	24	52	34	40	16
frequency	3	8	14	10	5	3

→

Values (x)	frequency (f)	fixi	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
15	3	45	15-36.11 = -21.11	445.63	1336.89
24	8	192	24-36.11 = -12.11	146.65	1173.2
52	14	728	52-36.11 = 15.89	252.49	3534.86
34	10	340	34-36.11 = -2.11	4.45	44.5
40	5	200	40-36.11 = 3.89	15.13	75.65
16	3	48	16-36.11 = -20.11	404.41	1213.23

$\Sigma f_i = N = 43$ $\Sigma f_i x_i = 1553$ $\Sigma (x_i - \bar{x})^2 = 1268.76$ $\Sigma f_i (x_i - \bar{x})^2 = 7378.33$

$$\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{1553}{43} = 36.11$$

S.D. or

$$\sigma = \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{N}}$$

$$= \sqrt{\frac{7378.33}{43}}$$

$$= \sqrt{171.58}$$

S.D. or $\sigma = 13.098$.

$$\boxed{\sigma = 13.098}$$

Page No.:
Date. / /

d) Explain permutation and combination.

→ Permutation refer to the number of different arrangment of given objects when order of the object is important then then arrangment of such of object is called as the permutation, and it is denoted by $P(n,r)$, ${}^nP_r$,

$${}^nP_r = \frac{n!}{(n-r)!}, \quad \text{for } r \leq n.$$

The number of distinct permutation of r distinct object choosen from a given collection of n-distinct object. denoted by ${}^nP_r$.

Combinations are related to the number of different selection from a given set of object when the order of object in the selection is immateral then the arrangment is called combination and it is denoted by $C(n,r)$, ${}^nC_r$.

$${}^nC_r = \frac{n!}{r!(n-r)!}$$

A combination of n-different object taken r-at a time is selection of r-out of the n-objects with no attention given to order of arrangment.

e) Find the means of x and y from two regression line is $2x + 5y + 8 = 0$, $6x + 3y + 6 = 0$.

→ Given regression line is

$$2x + 5y + 8 = 0$$
$$6x + 3y + 6 = 0.$$

Page No.:
Date. / /

$2x + 5y = -8$ —— (1)

$6x + 3y = -6$ —— (2)

Multiplying eqn (1) by 3

$6x + 15y = -24$ —— (3)

Subtracting the (3) by (2) we get.

$$\begin{array}{r} 6x + 15y = -24 \\ - \quad 6x + 3y = -6 \\ \hline - \quad + \end{array}$$

$12y = -18$

$y = \frac{-18}{12} = \frac{-3}{2} = -1.5$

$y = -1.5$ or $\bar{y} = -1.5$

Substituting the value of $y = -1.5$ in eqn (1) we get

$2x + 5y = -8$

$2x + 5(-1.5) = -8$

$2x - 7.5 = -8$

$2x = -8 + 7.5$

$2x = -0.5$

$x = \frac{-0.5}{2} = -0.25$

$x = -0.25$ or $\bar{x} = -0.25$

The mean of x and y is $\bar{x} = -0.25$ & $\bar{y} = -1.5$.

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

Page No.:
Date. / /

Q.4 Attempt any three of the following. (15-M)

a) Define event and explain its types.

→ Event -

A subset of a sample space is called an event. if two coin is toss. and event getting only head.

$S = \{HH, HT, TH, TT\}$.

$E_1 = \{HH\}$

$E_1 \subseteq S$.

$\therefore E_1$ is event of the sample space S.

Following are the some types of event.

i) simple event / Elementary event
ii) Compound event / certain event
iii) Sure event / impossiable event.

1) simple event -

An event consisting of single outcome is called simple event.

Ex: Tossing a coin.

$S = \{H, T\}$.

event occurs at only head.

$E_1 = \{H\}$

2) Compound event -

If an event has more than one sample point is called compound event.

Ex: Tossing a coin thrice.

$S = \{HHH, HHT, HTH, HTT, TTH, THT, THH, TTT\}$

event occurs at head appears.

$E_1 = \{HTH, HHH, HHT, HTT, TTH, THT, THH\}$

4) sure / impossiable event -

If event occurs empty or null sample point or no ~~one~~ the point possiable outcome is favourable is called as impossiable event.

Ex: A die is throwon.

$S = \{1, 2, 3, 4, 5, 6\}$

Event occurs greater than 6.

$E_1 = \{\phi\}$ or $E_1 = \{\ \}$

b) Explain regression equation.

→ If the variables in a bivariate distribution are related, we will find that the points in the scatter diagram will cluster round some curve called the curve of regression.

If the curve is straight line it is called the line of regression and there is said to be linear regression between the variables otherwise regression is said to be curvilinear.

Following are the two types of equation - i.e. x on y and y on x.

i) Regression line/equation of x on y.

$$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

where

$$r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma xy}{\Sigma y^2} \qquad x = x - \bar{x} \qquad y = y - \bar{y}.$$

ii) Regression line/equation of y on x.

$$y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

where

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2} \qquad x = x - \bar{x} \qquad y = y - \bar{y}.$$

c) Calculate the mode from the following data.

Wages -	200 - 250	250-300	300-350	350-400	400-450	450-500
Workers -	3	5	9	12	8	4

→

→ 7.

Wages	Workers	
200-250	3	
250-300	5	
300-350	9	→ f_1
350-400	12	→ fm
400-450	8	→ f_2
450-500	4	

from the above data maximum frequency is 12 and it lies in between 350-400.

$l = 350,\ fm = 12,\ f_1 = 9,\ f_2 = 8,\ h = 50$

$$\text{Mode} = l + \frac{fm - f_1}{2fm - f_1 - f_2} \times h.$$

$$= 350 + \frac{12 - 9}{2 \times 12 - 9 - 8} \times 50.$$

$$= 350 + \frac{3}{24 - 9 - 8} \times 50$$

$$= 350 + \frac{3}{24 - 17} \times 50$$

$$= 350 + \frac{150}{7}$$

$$= 350 + 21.42$$

Mode = 371.42

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

Page No.:
Date. / /

Q. 4) ——→ 2.

d) find the variance from following data.

Class	0-20	20-40	40-60	60-80	80-100
frequency	5	16	20	15	10.

→

class	frequency (f)	Mid-point (xi)	xifi	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
0-20	5	10	50	$10 - 52.72 = -42.72$	1824.9	9124.5
20-40	16	30	480	$30 - 52.72 = -22.72$	516.19	8259.04
40-60	20	50	1000	$50 - 52.72 = -2.72$	7.398	147.96
60-80	15	70	1050	$70 - 52.72 = 17.28$	298.59	4478.85
80-100	10	90	900	$90 - 52.72 = 37.28$	1389.79	13897.9

$N = \Sigma f = 66$ $\quad \Sigma f_i x_i = 3480$ $\quad \Sigma f_i (x_i - \bar{x})^2 = 35908.25$

$$\bar{x} = \frac{\Sigma x_i f_i}{N} = \frac{3480}{66} = 52.72$$

$$\text{Variance } (\sigma^2) = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{N}$$

$$= \frac{35908.25}{66}$$

$$\sigma^2 = 544.06$$

$\therefore$ Variance $(\sigma^2) = 544.06$.

Page No.:
Date. / /

Q 4) → 2.

e) If $P(A) = 0.6$, $P(B) = 0.5$, $P(A \cap B) = 0.3$
Compute.
$P(A')$, $P(A \cup B)$, $P(A' \cap B)$, $P(A' \cap B')$, and $P(A' \cup B')$

→ If given that

$P(A) = 0.6$, $P(B) = 0.5$, $P(A \cap B) = 0.3$

i)

$P(A') = 1 - P(A)$

$= 1 - 0.6$

$P(A') = \underline{0.4}$

ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= 0.6 + 0.5 - 0.3$

$P(A \cup B) = 0.8$

$P(A \cup B) = 0.8$

iii)

$P(A' \cap B) = P(B) - P(A \cap B)$

$= 0.5 - 0.3$

$P(A' \cap B) = 0.2$.

Page No.:
Date. / /

iv) $P(A' \cap B') = 1 - P(A \cup B)$

$= 1 - 0.8$

$P(A' \cap B') = 0.2$

v) $P(A' \cup B') = 1 - P(A \cap B)$

$= 1 - 0.3$

$P(A' \cup B') = 0.7$

Q.5) Attempt any three of the following. (15-M)

a) Explain limitation of statistics.

→ i) statistical laws are held to be true on the average.

ii) Statistical data always is treated as approximation.

iii) Statistical does not take individual cases.

iv) statistical method can not be applied to facts that can not be measured quantitatively such as health, culture, poverty, etc.

v) Statistical results are depend upon sample

vi) statistical results might lead to faciliate the conclusion.

Date. / /

vii) The greatest limitation of statistics is that only one who has an expert knowledge of statistical methods can efficiently handle statistical data.

b) Calculate the mean and median from the following data.

class	300-500	500-700	700-900	900-1100	1100-1300	1300-1500
frequency	25	55	30	20	14	6

→

class	frequency (fi)	midpoint (xi)	fixi	Cumulative freq c.f.
300-500	25	400	10,000	25
500-700	55	600	33,000	80
700-900	30	800	24,000	110
900-1100	20	1000	20,000	130
1100-1300	14	1200	16,800	144
1300-1500	6	1400	11,200	150.

$N = \Sigma f_i = 150$ $\Sigma f_i x_i = 115000$

Mean ($\bar{x}$)

$$\bar{x} = \frac{\Sigma f_i x_i}{N}$$

$$\bar{x} = \frac{115000}{150}$$

$$\bar{x} = 766.66$$

Class.....Roll No.....
Paper No.....Subject.....
Date.....
Invigilators Sign.....

Page No.:
Date. / /

Median -

$$\frac{N}{2} = \frac{150}{2} = 75.$$

Median class = 500 - 700

$l = 500, \frac{N}{2} = 75, f = 55, h = 200, C.f. = 25.$

$$\text{Median} = l + \left(\frac{\frac{N}{2} - C.f.}{f}\right) \times h.$$

$$= 500 + \frac{75 - 25}{55} \times 200$$

$$= 500 + \frac{50}{55} \times 200$$

$$= 500 + \frac{10,000}{55}$$

$$= 500 + 181.81.$$

Median = 681.81.

c) A bag contain 30 balls. numbered from 1 to 30. one ball is drawn at random. find the probability that the numbered ball drawn will be multiple of

i) 5 or 7

ii) 3 or 7.

Date. / /

A bag contain 30 balls. therefore number of ball is 30.

$$n(S) = 30 = {}^{30}C_1 = 30.$$

i) 5 or 7.

Suppose event A is the numbered ball will be multiple of 5.

$$A = \{5, 10, 15, 20, 25, 30\}$$

$$n(A) = 6$$

Suppose event B is the numbered ball will be multiple of 7.

$$B = \{7, 14, 21, 28\}$$

$$n(B) = 4.$$

~~Sup~~ $A \cap B$ the numbered ball will not be common in both event A and B.

$$n(A \cap B) = 0.$$

∴ the event A and B are mutually exclusive then the probability will be.

$$P(A \cup B) = P(A) + P(B)$$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)}$$

$$= \frac{6}{30} + \frac{4}{30}$$

$$P(A \cup B) = \frac{10}{30} = \frac{1}{3}$$

ii) 3 or 7.

Suppose the event A is the numbered ball will be multiple of 3.

$$A = \{3, 6, 9, 12, 15, 18, 21, 24, 27, 30\}$$

$$n(A) = 10$$

Suppose the event B is the numbered ball will be multiple of 7.

$$B = \{7, 14, 21, 28\}$$

$$n(B) = 4.$$

$A \cap B$ is the numbered ball will be common in both event A and B. i.e. $3 \times 7 = 21$

$$A \cap B = \{21\}$$

$$n(A \cap B) = 1.$$

then the probability of multiple of 3 or 7.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A \cap B)}{n(S)}$$

$$= \frac{10}{30} + \frac{4}{30} - \frac{1}{30}$$

$$P(A \cup B) = \frac{14}{30} - \frac{1}{30}$$

$$= \frac{13}{30}$$

$$P(A \cup B) = \frac{13}{30}.$$

d)

Obtain regression equation of x on y from the following data.

x	7	2	5	8	4
y	3	8	7	9	5

→

x	y	$(x-\bar{x})$	$(y-\bar{y})$	x^2	y^2	$x.y$
7	3	$7-5.2=1.8$	$3-6.4=-3.4$	3.24	11.56	-6.12
2	8	$2-5.2=-3.2$	$8-6.4=1.6$	10.24	2.56	-5.12
5	7	$5-5.2=-0.2$	$7-6.4=0.6$	0.04	0.36	0.12
8	9	$8-5.2=2.8$	$9-6.4=2.6$	7.84	6.76	7.28
4	5	$4-5.2=-1.2$	$5-6.4=-1.4$	1.44	1.96	1.68

$\Sigma x^2 = 22.8$ $\Sigma y^2 = 23.2$ $\Sigma xy = -2.16$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{26}{5} = 5.2$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{32}{5} = 6.4.$$

Regression equation of x on y

$$(x - \bar{x}) = r \frac{\sigma x}{\sigma y} (y - \bar{y})$$

$$r \frac{\sigma x}{\sigma y} = \frac{\Sigma x.y}{\Sigma y^2} = \frac{-2.16}{23.2} = -0.093$$

Class..... Roll No.....
Paper No..... Subject.....
Date.....
Invigilators Sign.....

Page No.:
Date. / /

$$(x - \bar{x}) = r \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

$$x - 5.2 = -0.093 (y - 6.4)$$

$$x - 5.2 = -0.093y + 0.595$$

$$x = -0.093y + 0.595 + 5.2$$

$$x = -0.093y + 5.795$$

or

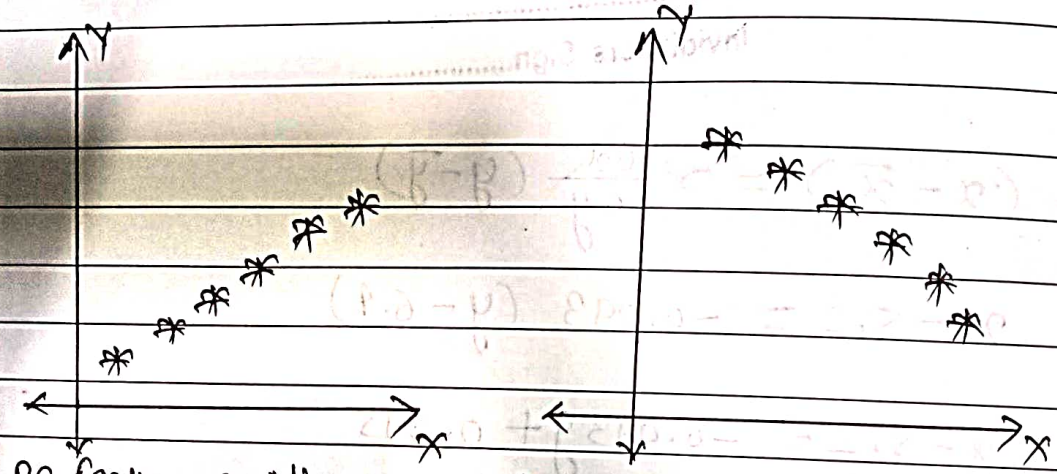
$$\boxed{x = 5.795 - 0.093y}$$

e) Explain Scatter diagram method.

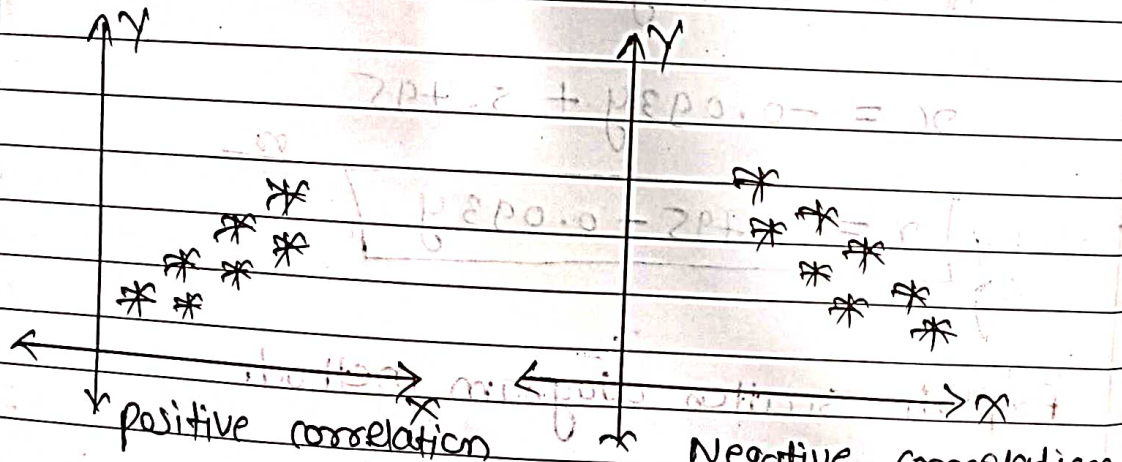
→

The scatter diagram method is the simplest method to study the correlation between two variables in the values for each pair of variables is plotted on a graph in the form of dots thereby obtaining as many points as the number of observation.

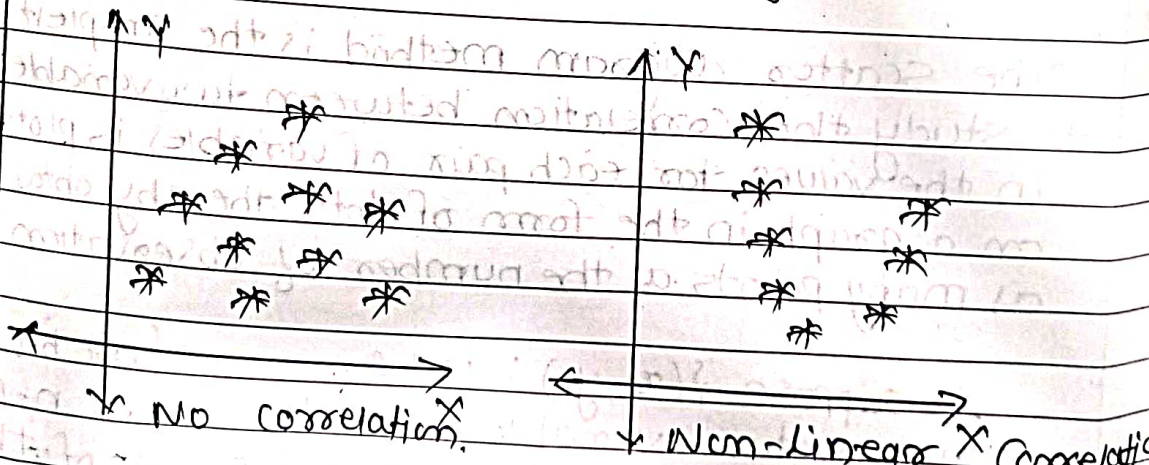
Suppose $\{(x_i, y_i);\ i = 1, 2, \ldots n\}$ are bivariate data on two variables x and y. if these n-pairs are plotted on a graph paper taking one of the variable on x-axis and other y-axis. we get a diagram called as scatter diagram with the help of scatter diagram we get a general idea about the existence of correlation and the type of correlation.



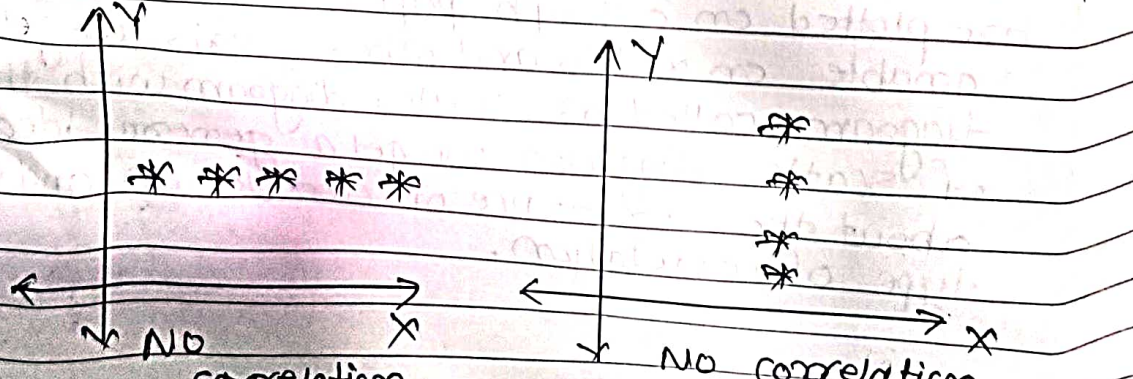
Perfectly positive correlation Perfectly negative correlation



positive correlation Negative correlation



No correlation. Non-Linear Correlation.



No correlation. No correlation.