

1

2

DURATION: 3 HOURS

MARKS: 100

59244

- N.B.: 1 All questions are compulsory.
 2 Use of simple calculator is allowed.

- Q.1 (a) **Correct the following if necessary.** 10
- (i) False. Secondary data is less reliable than primary data. 02
 - (ii) true The two attributes A and B are completely dissociated then coefficient of association is -1. 02
 - (iii) False. Harmonic mean is not reciprocal of arithmetic mean. 02
 - (iv) ~~False~~ ^{Standard} True. Mean deviation about ~~median~~ ^{mean} is minimum 02
 - (v) False. Percentiles divide the data into 100 equal parts. 02
- (b) **Answer in one sentence.** 10
- (i) Define secondary data. 02
 - (ii) ~~Stub, caption, body of table, foot note, source note, title.~~ ^{Def'n & example} 02
 - (iii) 19 02
 - (iv) For a symmetric distribution, $Q_1 = 30$ and $Q_2 = 35$ then $Q_3 = 40$ 02
 - (v) Define kurtosis. 02

- Q.2 (a) **Attempt any Two** 20
- Schedule** is the tool or instrument used to collect data from the respondents while interview is conducted. **Schedule** contains questions, statements (on which opinions are elicited) and blank spaces/tables for filling up the respondents. The **features of schedules** are :

- The schedule is presented by the interviewer. The questions are asked and the answers are noted down by him.
- The list of questions is a more formal document, it need not be attractive.
- The schedule can be used in a very narrow sphere of social research.

The main **purposes of schedule** are three fold :

- To provide a standardized tool for observation or interview in order to attain objectivity,
- To act as memory tickler i.e., the schedule keeps the memory of the interviewer/observer refreshed and keeps him reminded of the different aspects that are to be particularly observed, and
- To facilitate the work of tabulation and analysis. (6mk)

Merits ; ,Accurate screening, Capture verbal and non-verbal ques, Keep focus, Capture emotions and behaviors etc (2mk)

Demerits: Cost, quality of data, manual data entry, limitation on sample size etc (2mk)

- (b) (i) $(AB) \geq 0$ ii) $(AB) \geq (A) + (B) - N$ iii) $(AB) \leq (A)$ iv) $(AB) \leq (B)$ 10
- (1 mk each) (4mk)

2

(ii) $(A\beta\gamma)$, $(\alpha B\gamma)$, $(\alpha\beta C)$ and $(A\beta C)$. (6 mk)

- (c) (i) Explain Time series data and cross section data using example Def'n 2 mk each eg 1 mk each (6mk) 10
(ii) Define with a suitable example: Sample, attributes. Def'n along with eg 2 mk each (4mk)
- (d) Check for the consistency ^{of} for the following data: 10
(p) $(AB)=200$, $(\alpha\beta)=70$, $(\beta)=50$, $N=400$; not consistent
(q) $(AB)=12$, $(\alpha B)=26$, $(A\beta)=16$, $N=10$; not consistent

Q.3

Attempt any Two

- (a) (i) The weight (in grams) of 25 apples selected at random from a consignment are :- 20
05

106, 107, 76, 82, 107, 109, 115, 154, 95, 123, 125, 84, 111, 92, 86, 70, 126, 66, 130, 129, 149, 119, 115, 128, 100.

Form a grouped frequency table by taking class-intervals of equal size each of 20 grams such that the midpoint of first class-interval is 70 grams. Also find relative frequencies and percentage frequencies.

(3) 111 $\overline{||||} (5)$ $\overline{|||||} (9)$ $\overline{||||} (6)$ $|| (2)$
60-80, 80-100, 100-120, 120-140, 140-160

- (ii) Explain construction and use of stem and leaf diagram. 05
Construction (3mks)
Use of steam and leaf diagram illustration.(2mks)

- (b) Explain graphical methods for locating partition values Procedure (7 mk) 10
Illustration (3mk) *Ogive curves, Histogram*
Quartile mode

- (c) (i) If the value of each item is reduced by 15, what would be the effect on: 05

(p) The arithmetic mean; ^{reduced} by 15 (q) The range; *No change*
(r) Mean deviation about mean; (s) No effect on The standard deviatio *No change*.

- (ii) State any five requisites of a good measure of central tendency. 05
(any 5 points 1 mk each)

simple, easy to understand, rigidly defined, based on all observations, least affected by extreme observation, true representative of data etc.

- (d) For any two positive numbers a and b, show that 10
(p) $A.M. \geq G.M. \geq H.M.$ (4mk) (q) $G.M.^2 = A.M. \times G.M.$ (3mk)
(5) *(5)*

Q.4

Attempt any Two

- (a) (i) Explain concept of skewness. 20
04
(ii) State the different measures of skewness and their interpretations. 06
concept of skewness :The property of any deviation of frequency distribution from symmetry is skewness .(4mk)

3

figure to indicate different types of skewness and locate roughly the relative positions of mean, mode, median.(6mk)

- (b) (i) Explain effect of shift of origin and change in scale on central moments. 06
 (ii) What is dispersion? State the different relative measures of dispersion. 04

In **statistics, dispersion** (also called variability, scatter, or spread) is the extent to which a distribution is stretched or squeezed.(2mk)

Relative measures of dispersion:

- Coefficient of Range
- Coefficient of Quartile Deviation
- Coefficient of Mean deviation
- Coefficient of variation

- (c) Mean=2, variance=16, $\mu_3 = -64$, $\mu_4 = 672$, $\beta_1 = 1$, $\beta_2 = 2.625$. 10

- (d) i) Define raw moments about 'a' and central moment. 05
 ii) Compute the relationship between the central moments and the raw moments about zero (1mk each) 05

$\mu_1 = 0$ always

$\mu_2 = \mu_2' - \mu^2$

$\mu_3 = \mu_3' - 3\mu\mu_2' + 2\mu^3$

$\mu_4 = \mu_4' - 4\mu\mu_3' + 6\mu^2\mu_2'^2 - 3\mu^4$ (2mk)

Q.5

Attempt any Four

- (a) Calculate Yule's Coefficient of association between A and B and comment for the following data: $(AB)=200$, $(A\bar{B})=300$, $(\bar{A}B)=300$ and $(\bar{A}\bar{B})=200$. $(200 \times 200 - 300 \times 300) / (40000 + 30000) = -0.3846$ 20
 (b) Derive relationship between Yule's coefficient of association and coefficient of colligation. $Q = 2Y / (1 + Y^2)$ (Derivation 5mk) 05
 (c) Show that S.D. is not affected by shift of origin but get affected by change of scale. 05
 (d) Explain different components of table. 05
 Title of the table ii) Table number iii) Stubs and captions iv) Body of the table v) Source note vi) Foot note
 (e) Find geometric mean and arithmetic mean for the following data. 05

x	5	10	15	20
f	3	6	5	2

9

G.M = 10.86961 A.M = 11.875

- (f) How to find mean deviation about constant A? What should be choice of value of 'A' to have Mean deviation least? 05

Procedure: (2mk) choice of value of 'A' to have Mean deviation least = median (proof 3 mk)

- (g) State formulae for computing combined mean and combined standard deviation for two data series 05
combined standard deviation and mean

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$$

where, $d_1 = x_1 - \bar{x}$

$d_2 = x_2 - \bar{x}$

and $\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$ = combined AM