

3 (Sem-1) STS M 1

2017

STATISTICS

(Major)

Paper : 1.1

(Descriptive Statistics)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following questions as directed (reasoning is not necessary) : $1 \times 7 = 7$
 - (a) Write whether true or false :
"Data obtained from physical experiments are secondary data."
 - (b) Is it true that population is defined as a set of human beings?
 - (c) Is mode a partition value?
 - (d) Is coefficient of variation invariant of change of scale?

8A/391

(Turn Over)

- (e) State the range of multiple correlation coefficient.
- (f) Define geometric mean of the values x_1, x_2, \dots, x_n in terms of arithmetic mean.
- (g) State the values of β_1 and β_2 for a symmetric distribution.

2. Answer the following questions : $2 \times 4 = 8$

- (a) Mention two limitations of statistics.
- (b) If x_i / f_i ($i = 1, 2, \dots, n$) is a frequency distribution and $u_i = \frac{x_i - a}{h}$, then show that $\bar{x} = a + h\bar{u}$. (Symbols have their usual meanings.)
- (c) If r_{XY} is the coefficient of correlation between X and Y , then interpret the cases where—
- $r_{XY} = +1$;
 - $r_{XY} = -1$.
- (d) State Sheppard's corrections for moments for grouped data.

3. Answer any three of the following : $5 \times 3 = 15$

- (a) What is a statistical table? Mention with explanation the main parts of a statistical table.
- (b) Define arithmetic mean of a discrete frequency distribution. Show that the algebraic sum of the deviations of observations for the frequency distribution is minimum when taken about mean.
- (c) Define the following :
- Coefficient of correlation
 - Regression coefficients
 - Partial correlation coefficient
 - Multiple correlation coefficient
 - Correlation index
- (d) What are partition values of a distribution? Explain (with definition) median and quartiles as partition value.
- (e) Obtain the normal equations for fitting of the 2nd-degree parabola $Y = a + bX^2$ on the basis of the n pairs $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ of values of (X, Y) .

(4)

4. Answer either (a) or (b) :

(a) Define the following with suitable examples : $2 \times 5 = 10$

- (i) Qualitative and quantitative data
- (ii) Normal and ordinal data
- (iii) Cross-sectional and time-series data
- (iv) Discrete and continuous data
- (v) Frequency and non-frequency data

(b) What are primary data and secondary data? Clearly mention various sources of secondary data. $2+2+6=10$

5. Answer either (a) or (b) :

(a) Define raw moments, standard moments and factorial moments of a set of non-frequency numerical data. Express the 4th-order standard moment in terms of raw moments. $2+2+2+4=10$

(b) Define standard deviation of the observed values x_1, x_2, \dots, x_n . If $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$ are the variances of p different sets containing n_1, n_2, \dots, n_p observed values respectively, then find out the variance of all the $n_1 + n_2 + \dots + n_p$ observed values. $2+8=10$

(5)

6. Answer either (a) or (b) :

(a) Explain the principle behind the method of least squares of fitting a mathematical curve $y = f(x)$ to a set of numerical data viz. $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ on (x, y) . Find the normal equations for fitting of the mathematical curve

$$y = a + bx + cx^2 + dx^3$$

to the data on (x, y) mentioned above.

$$4+6=10$$

(b) Write notes on any two of the following :

$$5 \times 2 = 10$$

- (i) Skewness and Kurtosis
- (ii) Orthogonal polynomials
- (iii) Graphic representation of data
