

**POST GRADUATE DIPLOMA IN  
APPLIED STATISTICS (PGDAST)**

**Term-End Examination**

00107

June, 2016

**MST-002 : DESCRIPTIVE STATISTICS**

*Time : 3 hours*

*Maximum Marks : 50*

**Note :**

- (i) *Question no. 1 is compulsory. Questions no. 2 to 5 have internal choices.*
- (ii) *Use of scientific calculator is allowed.*
- (iii) *Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.*
- (iv) *Symbols have their usual meaning.*

1. State whether the following statements are *true* or *false*. Give reasons in support of your answers.  $5 \times 2 = 10$

(a) The weighted mean of  $n$  natural numbers is  $\frac{(2n+1)}{3}$ , if weights are the corresponding numbers.

(b) The values of regression coefficients  $b_{yx}$  and  $b_{xy}$  are  $\frac{1}{5}$  and 10 respectively.

- (c) If Y is proportional to X, then  $r(X, Y) = 1$ .
- (d) In case of consistent data, no class frequency can be negative.
- (e) If 25 is subtracted from each value of X and Y and then divided by 10, the new  $b'_{yx}$  is  $2\frac{1}{2}$  times of  $b_{yx}$ .

2. (a) The annual rates of growth of output of a factory in 5 years are 5.0, 7.5, 2.5, 5.0 and 10.0, respectively. What is the compound rate of growth of output per annum for the period ?
- (b) The number of employees, average daily wages and the variance of the wages for two factories are given as :

	Factory A	Factory B
Number of employees	50	100
Average daily wage (₹)	120	85
Variance of daily wages (₹)	09	16

- (i) In which factory is there a greater variation in the distribution of daily wages ?

- (ii) Suppose in factory B, the daily wages of an employee were wrongly noted as ₹ 120 instead of ₹ 100. What would the correct variance for factory B be? 6

OR

The data of monthly profit (in lacs) of various companies is given as follows :

Monthly Profit (in lacs)	Number of Companies
Less than 7.5	04
7.5 – 12.5	10
12.5 – 17.5	20
17.5 – 22.5	36
22.5 – 27.5	16
27.5 – 32.5	12
32.5 – 37.5	02

Determine whether the distribution is symmetric and mesokurtic by calculating the moments,  $\beta_1$  and  $\beta_2$ .

10

3. (a) A computer while calculating the correlation coefficient between the variables X and Y obtained the following results :

$$N = 30, \Sigma X = 120, \Sigma Y = 90, \Sigma X^2 = 600, \\ \Sigma Y^2 = 250, \Sigma XY = 335$$

It was, however, later discovered at the time of checking that it had copied wrongly, two pairs of observations (X, Y) as (8, 10) and (12, 7) while the correct values were (8, 12) and (10, 8) respectively. Obtain the correct value of the correlation coefficient between X and Y.

7

- (b) The coefficient of rank correlation of the marks obtained by 10 students in Statistics and Accountancy was found to be 0.8. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was taken as 7 instead of 9. Find the correct value of the rank correlation coefficient.

3

**OR**

- (a) In order to find the correlation coefficient between X and Y from 12 pairs of observations, the following calculations were made :

$$\Sigma X = 30, \Sigma Y = 5, \Sigma X^2 = 670, \Sigma Y^2 = 285, \Sigma XY = 334$$

On subsequent verification it was found that the pair (X = 11, Y = 4) was copied wrongly while the correct value was (X = 10, Y = 14). Find the correct value of the correlation coefficient.

7

- (b) An examination of eight applicants for a clerical post was taken by a firm. From the marks obtained by the applicants in the Accountancy and Statistics papers, compute the rank correlation coefficient to check how far the competence of the applicants in both subjects go together ?

3

Applicant :	A	B	C	D	E	F	G	H
Marks in Accountancy :	15	20	28	12	40	60	20	80
Marks in Statistics :	40	30	50	30	20	10	30	60

4. (a) The regression equation of Y on X and X on Y are  $Y = 2X$  and  $6X - Y = 4$ , respectively and the second moment of X about the origin is 3. Find (i) the mean values of X and Y; (ii) the value of the correlation coefficient, and (iii) the standard deviation of Y. 7

- (b) Suppose a data analyst has found  $r_{12} = 0.96$ ,  $r_{13} = 0.36$  and  $r_{23} = 0.78$  for a given set of values of  $X_1$ ,  $X_2$  and  $X_3$ . Examine whether their computations may be said to be free from error. 3

**OR**

Given the following bivariate data. :

X :	-1	5	3	2	1	1	7	3
Y :	-6	1	0	0	1	2	1	5

- (i) Fit a regression line of Y on X and find Y, if  $X = 10$ .
- (ii) Fit a regression line of X on Y and find X, if  $Y = 2.5$ . 10

5. (a) The following ultimate class frequencies are given as :

$$(ABC) = 60, (AB\gamma) = 250, (A\beta C) = 80,$$

$$(A\beta\gamma) = 350, (\alpha BC) = 75, (\alpha B\gamma) = 650,$$

$$(\alpha\beta C) = 55 \text{ and } (\alpha\beta\gamma) = 8200.$$

Find the frequencies of all positive classes. 6

- (b) What is meant by a class frequency of (i) first order, and (ii) third order ? How would you express a class frequency of first order in terms of class frequencies of third order ?

4

**OR**

A company is interested in determining the strength of association between the communicating time of their employees and the level of stress-related problems observed on job. A study of 116 assembly workers reveals the following :

<i>Communicating Time</i>	<i>Stress</i>			<i>Total</i>
	<i>High</i>	<i>Moderate</i>	<i>Low</i>	
Under 20 minutes	9	5	18	32
20 – 50 minutes	17	8	28	53
Over 50 minutes	18	6	7	31
Total	44	19	53	116

Use the coefficient of contingency to determine the amount of association between the communicating time and stress, given  $C_{\max} = 0.816$ .

10