# POST GRADUATE DIPLOMA IN
# APPLIED STATISTICS (PGDAST)

## Term-End Examination

01167          December, 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.                              *5×2=10*

    (a)   The standard deviation of 10 values of a data is 50. If each value is divided by 5, then the new standard deviation would become 10.

    (b)   If the correlation coefficient is 0, then the regression coefficient would also become 0.

(c) The geometric mean of regression coefficients is equal to twice the correlation coefficient.

(d) If $\dfrac{(A)}{N} = x$, $\dfrac{(B)}{N} = 2x$ and $\dfrac{(C)}{N} = 3x$ and $\dfrac{(AB)}{N} = \dfrac{(BC)}{N} = \dfrac{(AC)}{N} = y$, then neither x nor y can exceed $\dfrac{1}{4}$, if the data is consistent.

(e) If N = 100, (A) = 60, (B) = 40 and (AB) = 24, then the attributes A and B are positively associated.

2. (a) A toy factory has assigned a group of 4 workers to complete an order of 1400 toys of a certain type. The productive rates of the four workers are given below :

| Worker | Productive Rate |
|--------|-----------------|
| A | 4 minutes per toy |
| B | 6 minutes per toy |
| C | 10 minutes per toy |
| D | 15 minutes per toy |

Find the average minutes per toy devoted by the group of workers.

4

(b) The first four moments of a distribution about the value 5 of the variable are 2, 20, 40 and 50, respectively. Show that the mean is 7. Also check whether the distribution is symmetric and mesokurtic.     *6*

**OR**

For two firms, A and B, belonging to the same industry, the following details are available :

|  | Firm A | Firm B |
|---|---|---|
| Number of Employees | 100 | 200 |
| Average Wage | ₹ 240 | ₹ 170 |
| Standard Deviation of Wages | ₹ 6 | ₹ 8 |

(a) Which firm pays out a larger amount as weekly wages ?

(b) Which firm shows greater variability in the distribution of the wages ?

(c) Find the average weekly wages and standard deviation of the wages of all employees of both firms.     *10*

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

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

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) Two housewives, Mrs. Neena and Mrs. Meena, asked to express their preferences for different kinds of detergents, gave the following replies :

| Detergent | Meena | Neena |
|-----------|-------|-------|
| A | 1 | 1 |
| B | 4 | 2 |
| C | 2 | 4 |
| D | 3 | 3 |
| E | 5 | 7 |
| F | 7 | 8 |
| G | 6 | 6 |
| H | 8 | 5 |
| I | 9 | 9 |
| J | 10 | 10 |

To what extent do the preferences of these two ladies go together ? 3

**OR**

MST-002 4

(a) The profits of ₹ y of a certain company in the $x^{th}$ year of its existence are given by

X :  1   2   3   4   5

Y :  1250  1400  1650  1950  2300

Show that the best fit equation of the parabola on the given data is

$$Y = 1140 + 72X + 32 \cdot 15X^2.$$    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

4. (a) Two lines of regression are given by

$x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$ and $\sigma_x^2 = 12$.

Find (i) the mean values of x and y, (ii) the correlation coefficient between x and y, and (iii) the value of standard deviation of y.    7

(b) In a certain investigation, the following results were obtained for a given set of values of $X_1$, $X_2$ and $X_3$:

$r_{12} = 0 \cdot 8$, $r_{13} = 0 \cdot 2$ and $r_{23} = -0 \cdot 5$.

Do you think that the computations are free from error ?    3

**OR**

A company wants to assess the impact of R&D expenditure on its annual profit. The following table presents the information for the last eight years :

| Years | R&D Expenditure (in lakhs) | Annual Profit (in thousands) |
|-------|----------------------------|------------------------------|
| 2001  | 2                          | 20                           |
| 2002  | 3                          | 25                           |
| 2003  | 5                          | 34                           |
| 2004  | 4                          | 30                           |
| 2005  | 10                         | 60                           |
| 2006  | 5                          | 41                           |
| 2007  | 7                          | 42                           |
| 2008  | 9                          | 45                           |

Estimate the regression equation and predict the annual profit for 2012 for an allocated sum of ₹ 1 crore as R&D expenditure. *10*

5. (a) Among the adult population of a certain town, 50% of the population is male, 60% are wage-earners and 50% are 45 years of age or above. 10% of the males are not wage-earners and 45% of the males are under 45. Can we infer anything about what percentage of population of 45 years of age or above are wage-earners ? *4*

(b) In a sample of 1000 children, 400 came from higher income group and the rest from lower income group. The number of delinquent children in these groups was 50 and 200, respectively. Investigate if there is any association between delinquency and income groups. *6*

**OR**

For the given data, study the association between the temperament of Brothers and Sisters by computing the coefficient of contingency and interpret the results : *10*

| | | Temperament of Sisters | | | Total |
|---|---|---|---|---|---|
| | | Quick | Good Natured | Sullen | |
| Temperament of Brothers | Quick | 850 | 571 | 580 | 2001 |
| | Good Natured | 618 | 593 | 455 | 1666 |
| | Sullen | 540 | 456 | 457 | 1453 |
| | Total | 2008 | 1620 | 1492 | 5120 |
| | Given that $C_{max} = 0.816$. | | | | |