

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

Term-End Examination

00042

December, 2018

MST-002 : DESCRIPTIVE STATISTICS

Time : 3 hours

Maximum Marks : 50

Note :

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

1. State whether the following statements are *True* or *False*. Give reasons in support of your answers.

$5 \times 2 = 10$

- (a) There are $(n + 1)$ observations in a sample. If \bar{x}_1 is the mean of the first n observations and \bar{x}_2 is the mean of the last n observations, then $\bar{x}_2 = \bar{x}_1$.

- (b) If the two lines of regression are coincident, the relation between the regression coefficients is $b_{YX} \cdot b_{XY} = 1$.
- (c) If X and Y are two independent variables with $\text{Var}(X) = 25$ and $\text{Var}(Y) = 15$, then correlation coefficient between $V = X + Y$ and $V = X - Y$ is equal to 0.25.
- (d) If $(A) = 90$, $(AB) = 40$, $N = 150$ and $(\beta) = 80$, then $(\alpha B) = 50$.
- (e) The mean and standard deviation of X are 25 and 5, respectively. If a constant 10 is multiplied to each value of X, the CV remains unchanged.
2. (a) The price of a commodity increased by 5% from 1999 to 2000, 8% from 2000 to 2001 and 77% from 2001 and 2002. The average price increase from 1999 to 2002 is quoted as 26% and not 30%. Why? 4
- (b) The mean and standard deviation of 20 items are found to be 10 and 2, respectively. At the time of checking, it was observed that one item 8 was incorrect. Find the mean and standard deviation if (i) the value is omitted, (ii) it is replaced by 12. 6

OR

- (a) The runs scored by two batsmen in eight matches are as follows :

| <i>Batsman A</i> | <i>Batsman B</i> |
|------------------|------------------|
| 27 | 0 |
| 16 | 100 |
| 39 | 80 |
| 45 | 5 |
| 101 | 60 |
| 80 | 40 |
| 40 | 10 |
| 52 | 121 |

Who is a better run scorer ? Also find which of the two batsmen is more consistent in scoring.

7

- (b) The geometric mean of 10 observations on a certain variable was calculated as 16.2. It was later discovered that one of the observations was wrongly read as 12.9, in fact it was 21.9. Apply the appropriate correction and calculate correct geometric mean.

3

3. (a) Fit an equation of the form $y = ab^x$ to the following data :

5

| | | | | | |
|---|-----|-------|-------|-------|-------|
| x | 2 | 3 | 4 | 5 | 6 |
| y | 144 | 172.8 | 207.4 | 248.5 | 298.5 |

- (b) Suppose X and Y are the two variables having the correlation 0.85 on the following values :

| | | | | |
|---|----|----|----|----|
| X | 10 | 30 | 50 | 60 |
| Y | 40 | 30 | 70 | 80 |

Then two new variables X' and Y' are obtained by multiplying each observation of X and Y by 3 and 4 , respectively. Find the correlation between X' and Y' using the given data. Also, interpret the result.

5

OR

- (a) The following data relating to the heights of the fathers and their sons (in inches) are as follows :

| <i>Height of Fathers</i> | <i>Height of Sons</i> |
|--------------------------|-----------------------|
| 68 | 65 |
| 68 | 64 |
| 69 | 67 |
| 72 | 69 |
| 65 | 64 |
| 59 | 60 |
| 62 | 59 |
| 67 | 68 |
| 61 | 60 |
| 71 | 64 |

What conclusion do you draw from this data in regard to the relation between the heights of fathers and sons ?

5

- (b) The coefficient of rank correlation of the marks obtained by 10 students in Statistics and Accountancy was found to be 0.2. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 9 instead of 7. Find the correct value of coefficient of rank correlation. 3
- (c) Explain the method of least squares. 2
4. (a) The equation of two regression lines are given as :
- $$Y = X + 5 \text{ and } 16X = 9Y + 95$$
- If standard deviation of Y is 4, then find the values of (i) Means of X and Y, (ii) $r(X, Y)$, and (iii) Standard deviation of X. 7
- (b) Examine if it is possible to get the following values of $r_{12} = 0.91$, $r_{13} = 0.33$ and $r_{23} = 0.81$ from some set of the experimental data. 3

OR

For the following data given in the table below :

| | | | | | | |
|---------|----|----|----|----|----|----|
| X_1 : | 4 | 6 | 7 | 19 | 13 | 15 |
| X_2 : | 15 | 12 | 8 | 6 | 4 | 3 |
| X_3 : | 30 | 24 | 20 | 14 | 10 | 4 |

find the regression equation of X_1 on X_2 and X_3 .
Also estimate the value of X_1 for $X_2 = 6$ and $X_3 = 8$. 10

5. (a) If in an urban district 817 per thousand of the women between 20 and 25 years of age were noted as occupied at a census and 263 per thousand were noted as married or widowed, what is the lowest proportion per thousand of the married or widowed that must have been occupied ?

3

(b) For 1000 school boys, test whether there is any association between the General ability and Mathematical ability from the following data :

7

| Mathematical Ability | General Ability | | | |
|----------------------|-----------------|------|------|-------|
| | Good | Fair | Poor | Total |
| Good | 44 | 22 | 04 | 70 |
| Fair | 265 | 257 | 178 | 700 |
| Poor | 41 | 91 | 98 | 230 |
| Total | 350 | 370 | 280 | 1000 |

OR

(a) A fundamental set of class frequencies are given below :

$$N = 275, (A) = 132, (B) = 112, (C) = 129$$

$$(AB\gamma) = 27, (A\beta C) = 39, (\alpha BC) = 37, (ABC) = 20.$$

Calculate all (i) positive, and (ii) ultimate class frequencies.

5

- (b) In a college of 1000 students, equal number of students came from each of higher and lower income groups. The number of boys in these groups was 100 and 400, respectively. Calculate the coefficient of association between gender and income group.

5
