

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

Term-End Examination

June, 2018

01915

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 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.

5×2=10

- (a) If $x_1, x_2, x_3, \dots, x_n$ and $y_1, y_2, y_3, \dots, y_n$ are the variate values of two variables X and Y, and their geometric means are G_1 and G_2 , respectively, then geometric mean of $(x_i/y_i); i = 1, 2, \dots, n$ will be (G_1/G_2) .

(b) If each value of X is divided by 2 and of Y is multiplied by 2, then b'_{yx} will be same as b_{yx} .

(c) If X and Y are two independent variables and the variables $U = X + Y$ and $V = X - Y$ then the $r(U, V) = \frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$.

(d) If $(A) = 90$, $(AB) = 40$, $N = 150$ and $(\beta) = 80$ then $(\alpha\beta) = 30$.

(e) The mean and standard deviation of a set of values are 25 and 5, respectively. If a constant value 5 is added to each value, the coefficient of variation of the new set of values is equal to 10%.

2. (a) The numbers 3·2, 5·8, 7·9 and 4·5 have frequencies Y, (Y + 2), (Y - 3) and (Y + 6), respectively. If the arithmetic mean is 4·876, find the value of Y and write the whole series.

4

- (b) The following is the distribution of age (in years) of 800 workers :

<i>Age Group</i>	<i>No. of Workers</i>
20 – 25	50
25 – 30	70
30 – 35	100
35 – 40	180
40 – 45	150
45 – 50	120
50 – 55	70
55 – 60	60

Find (i) Median, (ii) Quartile Deviation, and (iii) Coefficient of Quartile Deviation.

6

OR

- (a) In the given data, two frequencies are missing and its mean is found to be 1.46.

<i>No. of Accidents (x)</i>	<i>Frequencies (f)</i>
0	46
1	?
2	?
3	25
4	10
5	5
Total	200

Find the missing frequencies.

6

- (b) Suppose 2, 6, 9, 5, 4 are 5 observations on a variable X with standard deviation 2.32. A new variable Y is obtained by multiplying each observation on X by 3. Further, another variable Z is obtained by dividing each observation on X by 2. Then find the effect of these changes in obtained variables in terms of standard deviation and explain the results.

4

3. (a) Suppose X and Y are the two variables having the correlation coefficient 0.85. The following are the values they have :

X	Y
10	40
30	30
50	70
60	80

If two new variables X' and Y' are obtained by adding 50 to each value of X and 100 to each value of Y, respectively, calculate the correlation coefficient between X' and Y' using the above data. Also compare the results.

5

- (b) Calculate the intensity of relation between the ranks of the marks, obtained by 10 candidates in an interview allotted by two experts, given as follows :

5

<i>Expert-1</i>	<i>Expert-2</i>
48	13
33	13
40	24
09	06
16	15
16	04
65	20
24	09
16	06
57	19

OR

- (a) The value of Spearman's rank correlation coefficient of a set of non-repeating values was found to be $\frac{2}{3}$. The sum of the squares of difference between the corresponding ranks was 55. Find the number of pairs.

4

- (b) Calculate Karl Pearson's coefficient of correlation between X and Y for the following data :

$$N = 12, \Sigma X = 120, \Sigma Y = 130, \Sigma(X - 8)^2 = 50, \\ \Sigma(Y - 10)^2 = 200 \text{ and } \Sigma(X - 8)(Y - 10) = 50.$$

6

4. (a) The following table shows the information as :

Statistical Measures	Advertisement Expenditure (X) (₹ Lakhs)	Sales (Y) (₹ Lakhs)
Mean	20	100
Standard Deviation	03	12

$r(X, Y) = 0.8$. Then find

- (i) the expected advertising expenditure of the company if sale is ₹ 125 lakhs, and
- (ii) the expected sales of the company if the advertising expenditure is ₹ 32 lakhs. 6

- (b) Given the following data : 4

$$r_{12} = 0.8, r_{13} = 0.6 \text{ and } r_{23} = 0.4$$

then find

- (i) $r_{12.3}$
- (ii) $r_{13.2}$
- (iii) $r_{23.1}$
- (iv) $R_{1.23}$

OR

- (a) In a statistical study relating to the prices (in ₹) of two shares, X and Y, the following two regression lines were found as

$$8X - 10Y + 70 = 0$$

$$20X - 9Y - 65 = 0$$

The standard deviation of $X = 3$,

then find

(i) the values of \bar{X} and \bar{Y} ,

(ii) $r(X, Y)$, and

(iii) standard deviation of Y.

7

- (b) Suppose a computer has found for a given set of values of X_1, X_2 and X_3 : $r_{12} = 0.90$, $r_{13} = 0.30$ and $r_{23} = 0.70$. Examine whether these computations are error free.

3

5. (a) An investigation of 23713 households was made in an urban and rural mixed locality. Of these 1618 were farmers, 2015 well to do and 770 families were having at least one graduate. Of these graduate families 335 were those of farmers and 428 were well to do; also 587 well to do families were those of farmers and out of them only 156 were having at least one graduate. Obtain all the ultimate class frequencies.

6

(b) Can vaccination be regarded as a preventive measure for smallpox from the given data :

- (i) Of 1482 persons in a locality exposed to smallpox, 368 in all were attacked, and
- (ii) Of 1482 persons, 343 had been vaccinated and of these only 35 were attacked.

4

OR

(a) Test whether there is any association between the heights of parents and that of their offsprings for the following data :

7

Offsprings Heights	Parents Heights				Total
	Above 5'8"	5'6" – 5'8"	5'3" – 5'5"	Below 5'3"	
Above 5'8"	20	30	20	02	72
5'6" – 5'8"	14	125	85	12	236
5'3" – 5'5"	03	140	165	125	433
Below 5'3"	03	37	68	151	259
Total	40	332	338	290	1000

(b) 50% of items have characteristics A and B both, 35% have A but not B, 25% have B but not A. Show that there must be some misprints in this report.

3