POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

February, 2021

MST-002 : DESCRIPTIVE STATISTICS

Time : 3 hours *Maximum Marks : 50*

Note :

(i) *Question no. **1** is **compulsory**.*

(ii) *Attempt any **four** questions from the remaining questions no. 2 to 7.*

(ii) *Use of scientific (non-programmable) calculator is allowed.*

(iii) *Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.*

(iv) *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 the arithmetic mean of the numbers 3·2, 5·8, 7·9 and 4·5 with their corresponding frequencies Y, (Y + 2), (Y – 3) and Y + 6 is 4·876, then the value of Y is 5.

(b) The sum of squares of deviations for 10 observations taken from their mean 50 is 250. The coefficient of variation is 10%.

(c) If the sum of the product of deviations of X and Y values from their respective means is zero, then the r(x, y) will be –1.

(d) If $b_{xy} = -\frac{4}{3}$ and $b_{yx} = -\frac{1}{12}$, the value of r will be $+\frac{1}{3}$.

(e) The data is consistent if N = 1000, (A) = 600 (B) = 500 and (AB) = 50.

2. (a) The following is the distribution of age of 80 workers : *3*

Age Group	*No. of Workers*
20 – 25	5
25 – 30	7
30 – 35	10
35 – 40	18
40 – 45	15
45 – 50	12
50 – 55	7
55 – 60	6

Find Quartile deviation.

(b) The number of runs scored by two batsmen in consecutive eight matches are given below :

Batsman A :	27	16	39	45	101	80	40	52
Batsman B :	0	100	80	5	60	40	10	121

Find who is a better run scorer. Also find which of the two batsmen is more consistent in scoring. *7*

3. With 10 observations, each on two variables X and Y, the following data were observed :

$$\bar{X} = 12,\ \sigma_x = 3,\ \bar{Y} = 15,\ \sigma_y = 4 \text{ and } r = 0{\cdot}5$$

However, on subsequent verification, it was found that one value of X (=15) and one value of Y (=13) were wrongly taken as 16 and 18 respectively. Find the correct value of correlation coefficient. *10*

4. Find the multiple linear regression equation of X_1 on X_2 and X_3 from the data relating to three variables given below :

X_1 :	4	6	7	9	13	15
X_2 :	15	12	8	6	4	3
X_3 :	30	24	20	14	10	4

Also estimate the best value of X_1 for $X_2 = 4$ and $X_3 = 10$. *10*

5. 800 candidates comprising both boys and girls appeared in an examination. The boys outnumbered the girls by 15% of the total. The number of candidates who passed exceeded the number failed by 480. Equal number of boys and girls failed in the examination. Prepare a 2×2 table and find the coefficient of association. *10*

6. (a) The mean annual salary of all employees in a company is ₹ 25,000. The mean salary of male and female employees is ₹ 27,000 and ₹ 17,000 respectively. Find the percentage of males and females employed by the company. *3*

(b) For a bivariate data, the two regression equations are $8Y = 6X$ and $Y = 3X$. Find (i) means of X and Y, (ii) r(X,Y), and (iii) value of σ_y if value of $\sigma_x = 4$. *7*

7. (a) Calculate the coefficient of rank correlation for the following data : *5*

X :	80	78	75	75	68	57	60	59
Y :	110	111	114	114	114	116	115	117

(b) Check whether A and B are independent, positively associated or negatively associated in case of the following data : *5*

$(A) = 490$, $(AB) = 294$, $(\alpha) = 570$ and $(\alpha\beta) = 380$.