No. of Printed Pages : 5

MST-002

POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

00232

December, 2017

MST-002 : DESCRIPTIVE STATISTICS

Time : 3 hours *Maximum Marks : 50*

Note :

(i) *Attempt **all** questions.*

(ii) *Questions no. 2 to 5 have internal choices.*

(iii) *Use of scientific calculator is allowed.*

(iv) *Formulae and 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 D_i denotes the i^{th} decile, i = 1, 2, ..., 9, then 70% observations are greater than or equal to D_7 and 30% observations are less than or equal to D_7.

(b) The range of the data given below is 60.

CI	f
10 – 20	0
20 – 30	5
30 – 40	10
40 – 50	10
50 – 60	5
60 – 70	0

(c) For the data 4, 7, 9, 5, 10, the value of A is 5·5, such that $\sum_{i=1}^{5} |x_i - A|$ is minimum.

(d) If X is measured in metres and Y is measured in kg, then the unit of r (X, Y) is kg/m.

(e) If (AB) = 150, (αB) = 260, (Aβ) = 230, then (A) = 410.

2. (a) For the data 4, 7, 9, 5, 8, 15, 11, 10, 12, 18, if we are interested in a value which is greater than or equal to 40% observations and less than or equal to 60% observations, which measure should be applied and what is the required value ?

(b) For the data 10, 27, 40, 60, 33, 30, 10; find the square root of the mean of squares of the deviations from the mean of this data. Also find out the unit of this measure, if the unit of the given data is km.

(c) Assume that the data given below is in ascending order :

$$4, 7, 9, x + 3, 2x, 17, 19, 20$$

Find the value of x if the median of the data is 13·5. *4+4+2*

OR

Weekly wages of workers (in "000" rupees) are given in the the following table :

Weekly wages	Frequency
10 – 12	1
12 – 14	3
14 – 16	7
16 – 18	12
18 – 20	12
20 – 22	4
22 – 24	3

By using appropriate measures, comment on the skewness and kurtosis of the distribution of the data. *10*

3. For the data

x	6	7	8	9	11
y	5	4	3	2	1

fit a linear curve between y (response) and x (predictor), using the principle of least squares. Also estimate the response if predictor is 5. *10*

OR

X and Y are two characteristics which cannot be measured directly. Using a criterion, scores are given to these characteristics as

X	70	70	80	80	80	90	100
Y	90	90	90	80	70	60	50

Find the association between qualitative characteristics X and Y. *10*

4. Height of fathers and their sons in inches are given below :

Height of Fathers	65	66	67	67	68	69	70	71
Height of Sons	66	68	65	69	74	73	72	70

What is the estimated average height of the son corresponding to the height 68·5 inches of the father ? *10*

OR

From the given data in the following table, estimate the value of X_1 for $X_2 = 45$ and $X_3 = 8$: *10*

X_1	1	2	3	4	5
X_2	3	4	5	6	7
X_3	4	5	6	7	8

5. (a) Given the following class frequencies, do you find any inconsistency in the data ?

$(A) = 300, (B) = 150, (\alpha\beta) = 110, N = 500$

(b) A number of school children were examined for the presence or absence of certain defects of which three chief descriptions were noted. Let A denote development defects; B, nerve sign and C, low nutrition. Given the following ultimate frequencies, find the frequencies of the class defined by the presence of the defects :

$(ABC) = 60, (\alpha BC) = 75, (AB\gamma) = 250,$

$(\alpha B\gamma) = 650, (A\beta C) = 80, (\alpha\beta C) = 55,$

$(A\beta\gamma) = 350, (\alpha\beta\gamma) = 8200$ *5+5*

OR

(a) In a survey of 1000 children, 811 liked pizza, 752 liked chowmein and 418 liked burger; 570 liked both pizza and chowmein; 356 liked both pizza and burger; 348 liked both chowmein and burger; 297 liked all three. Test the consistency of the data.

(b) 1660 candidates appeared for a competitive examination. Of these, 425 were successful; 252 had attended a coaching class and of them, 150 were successful. Is there any association between success and utility of the coaching class ? *5+5*