## BACHELOR OF COMPUTER APPLICATIONS (Revised) (BCA)

## **Term-End Practical Examination**

June, 2017

## **BCSL-044(P)/S2: STATISTICAL TECHNIQUES LAB**

| Time : 1 Hour |      | Maximum Marks:  |  |  |
|---------------|------|---|--|--|
| Note:         | (i)  | There are two compulsory questions in this paper of 20 marks each. Rest 10 marks are for viva-voce                  |  |  |
|               | (ii) | Use any spreadsheet package for solving the problems. For programming (if asked), you may use any $C/C++$ compiler. |  |  |

1. Salary of 20 different persons of a city was recorded in the following table:

## Monthly Salary in INR

| 75,000 | 10,000 | 6,000  | 71,000 | 15,000 |
|--------|--------|--------|--------|--------|
| 85,000 | 35,000 | 25,000 | 15,000 | 17,000 |
| 27,000 | 45,000 | 57,000 | 67,000 | 74,000 |
| 38,000 | 49,000 | 59,000 | 61,000 | 63,000 |

Perform the following tasks for the data given above:

8+4+4+4=20

- (a) Enter the data in a spreadsheet and create a frequency distribution in 5 equal ranges. Use array formula for finding the frequency distribution.
- (b) Draw the histogram of the data.
- (c) Find the relative frequency distribution for the frequency distribution obtained in part (a).
- (d) Find the mean and standard deviation for the data.

2. To find a relationship between infrastructure availability level and number of industries, the following data was collected:

| Infrastructure Availability Level (1 – 10) (1-minimum, 10-maximum) | No. of Industries in that town |  |
|--|--------------------------------|--|
| 5  | 27                             |  |
| 1  | 11                             |  |
| 9  | 39                             |  |
| 10   | 45                             |  |
| 6  | 27                             |  |
| 3  | 15                             |  |
| 8  | 35                             |  |
| 7  | 30                             |  |

- (a) Draw a scatter plot for the given data using a spreadsheet package.
- (b) Find the best linear regression line assuming that number of industries is a dependent variable and infrastructure availability level is an independent variable. Is this line a good fit? Explain.

  10+10=20