Answer all questions                    Time : Two hours

Calculator and Statistical tables will be provided

1. (a) Draw a box plot for the data given below and interpret it.

   14, 36, 39, 41, 17, 22, 19, 16, 21, 21, 18, 33, 25, 31, 18, 60.

   (b) Marks of students in a class for a certain course unit, have been summarized as shown in table below.

   | Marks | $0 < X \le 10$ | $10 < X \le 20$ | $20 < X \le 30$ | $30 < X \le 40$ | $40 < X \le 50$ |
   |---|---|---|---|---|---|
   | Number of students | 10 | 20 | 30 | 25 | 15 |

   Comment on the symmetry of the distribution of marks of students by using the Pearson's first coefficient of skewness.

2. With the intension of finding the relationship between two variables $X, Y$. Following data have been collected from a sample of 10 units.

   | X | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
   |---|---|---|---|---|---|---|---|---|---|---|
   | Y | 37 | 63 | 81 | 115 | 140 | 155 | 195 | 207 | 245 | 268 |

   (a) Test the significance of the relationship between $X$ and $Y$ at 5% significance level and interpret the results.

   (b) Fit a regression model of the form of $Y = \beta_0 + \beta_1 X$.

   (c) Test the significance of the $\beta_0$ and $\beta_1$ at 5% significance level.

   (d) Find the 95% confidence interval for $\beta_1$ and interpret it.

3. (a) Show that Spearman's rank correlation coefficient $r_s$ is given by

$$r_s = 1 - \frac{6 \sum\limits_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

where $n$ is the number of observations and $d_i$ is the difference between ranks assigned to the $i^{\text{th}}$ individuals.

(b) Find the rank correlation coefficient between the midterm marks and the IQ ranking of a random sample of 10 students in a large class.

| Students | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Midterm Marks | 77 | 78 | 65 | 84 | 85 | 88 | 67 | 92 | 68 | 96 |
| IQ Ranking | 7 | 6 | 8 | 5 | 4 | 3 | 9 | 1 | 10 | 2 |

4. A farmer wants to find the relationship between the amount of fertilizer used and the yield of corn. He selected several acres of his land on which he used different amount of fertilizer to grow corn. The following table gives the amount of fertilizer (in pounds) used and the yield of corn (in pounds) for each of the seven acres.

| Amount of fertilizer | 120 | 80 | 100 | 70 | 88 | 75 | 110 |
|---|---|---|---|---|---|---|---|
| Yield of corn | 138 | 112 | 129 | 96 | 119 | 104 | 134 |

(a) Draw a scatter diagram for those data. Does the scatter diagram show a linear relationship between fertilizer used and yield of corn?

(b) Fit the estimated regression line. giving the estimated model for the data.

(c) Give a brief interpretation of the estimated slop calculated in part (b).

(d) Test at 5% significance level if the true slop is different from zero.

(e) Compute the coefficient of determination and interpret it.

(f) Find a 95% confidence interval for the true slop.

(g) What is the estimated value of the yield of corn if the farmer used 125 pounds of fertilizer to grow?