# POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST) <br> Term-End Examination <br> June, 2016 

## MSTL-001/S2 : BASIC STATISTICS LAB SET-2

Time : 3 Hours
Maximum Marks : 50
Note: (i) Attempt any two questions.
(ii) Solve the questions in Microsoft Excel.
(iii) Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.
(iv) Mention necessary steps, hypothesis, interpretation, etc.

1. (a) In 2008, a store operated more than 7200 outlets and reported revenues exceeding $\$ 373$ billion. Sales for the store are highly seasonal and therefore quarterly revenue was considered. The following table lists the quarterly revenues (in billions of \$) :

| Quarter | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| I | $43 \cdot 0$ | $48 \cdot 6$ | $55 \cdot 0$ | $56 \cdot 7$ | $64 \cdot 8$ | $71 \cdot 6$ | $79 \cdot 6$ | $85 \cdot 4$ |
| II | $46 \cdot 1$ | $53 \cdot 3$ | $59 \cdot 7$ | $62 \cdot 6$ | $69 \cdot 7$ | $76 \cdot 8$ | $84 \cdot 5$ | $91 \cdot 2$ |
| III | $45 \cdot 7$ | $51 \cdot 8$ | $58 \cdot 8$ | $62 \cdot 4$ | $68 \cdot 5$ | $75 \cdot 4$ | $83 \cdot 5$ | $90 \cdot 9$ |
| IV | $56 \cdot 6$ | $64 \cdot 2$ | $71 \cdot 1$ | $74 \cdot 5$ | $82 \cdot 2$ | $88 \cdot 6$ | $98 \cdot 1$ | $106 \cdot 3$ |

(i) Construct time-series plot of the given data and interpret.
(ii) Which quarter is most consistent? $4+6$
(b) There are several methods for calculating fuel economy. The following table indicates the mileage (in km /litre) as calculated by owners and by current Government standards for 9 different makes of a company :

| Makes : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Owner : | $14 \cdot 3$ | $15 \cdot 0$ | $27 \cdot 8$ | $27 \cdot 9$ | $48 \cdot 8$ | $16 \cdot 8$ | $23 \cdot 7$ | $32 \cdot 8$ | $37 \cdot 3$ |
| Government : | $16 \cdot 8$ | $17 \cdot 8$ | $26 \cdot 2$ | $34 \cdot 2$ | $47 \cdot 6$ | $18 \cdot 3$ | $28 \cdot 5$ | $33 \cdot 1$ | $56 \cdot 0$ |

(i) Compute the covariance and coefficient of correlation.
(ii) Represent the above data using a suitable diagram.
(iii) Compute the coefficients of skewness and kurtosis and interpret. $2+5+8$
2. (a) A problem with a telephone line that prevents a customer from receiving or making calls is disconcerting to both the customer and the telephone company. The following data represents a sample of 20 problems reported to two different offices of the time (in minutes) to clear these problems :

Office I: $\quad 1 \cdot 48,1 \cdot 75,0.78,2 \cdot 85,0.52,1 \cdot 60,4 \cdot 15,3 \cdot 97,1 \cdot 48,3 \cdot 10,1 \cdot 02$, $0.53,0.93,1 \cdot 60,0.80,1 \cdot 05,6 \cdot 32,3.93,5 \cdot 45,0.97$

Office II : $\quad 7 \cdot 55,3 \cdot 75,0 \cdot 10,1 \cdot 10,0 \cdot 60,0 \cdot 52,3 \cdot 30,2 \cdot 10,0.58,4 \cdot 02,3 \cdot 75$, $0.65,1.92,0.60,1.53,4.23,0.08,1.48,1 \cdot 65,0.72$
(i) Test the equality of population variances at $5 \%$ level.
(ii) Is there evidence of a significant difference in the mean waiting time between the two offices?
Assume that the time to clear a problem follows normal distribution. $5+5$
(b) A supermarket that has a chain of stores is concerned about its service quality reputation perceived by its customers. The data given below shows the perceived service quality with regard to politeness of the staff. The number in each cell of the data is the percentage of people who have said that the staff is polite.

| Day | Store |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | A | B | C | D | E |
| Monday | 79 | 81 | 74 | 77 | 66 |
| Tuesday | 78 | 86 | 89 | 97 | 86 |
| Wednesday | 81 | 87 | 84 | 94 | 82 |
| Thursday | 80 | 83 | 81 | 88 | 83 |
| Friday | 70 | 74 | 77 | 89 | 68 |

Perform suitable tests to examine whether significant differences at $1 \%$ level in politeness of the staff exist (i) day-wise, and (ii) store-wise. If there are significant differences, carry out pairwise comparisons.
3. (a) A sporting goods manufacturing company wanted to compare the distance travelled by golf balls produced across each of the 4 different designs. Ten balls were manufactured with each design and were brought to the local golf course for the club professionals to test. The order in which the balls were hit with the same club from the first tee was randomised so that the professional did not know which type of ball was being hit. All the 40 balls were hit in a short period of time, during which the environmental conditions were essentially the same. The results (distance travelled in metres) for the four designs were as follows :

| Design |  |  |  |
| :---: | :---: | :---: | :---: |
| 1 | 2 | 3 | 4 |
| 206.32 | 217.08 | 226.77 | 230.55 |
| 207.94 | 221.43 | 224.79 | 227.95 |
| 206.19 | 218.04 | 229.75 | 231.84 |
| 204.45 | 224.13 | 228.51 | 224.87 |
| 209.65 | 211.82 | 221.44 | 229.49 |
| 203.81 | 213.90 | 223.85 | 231.10 |
| 206.75 | 221.28 | 223.97 | 221.53 |
| 205.68 | 229.43 | 234.30 | 235.45 |
| 204.49 | 213.54 | 219.50 | 228.35 |
| 210.86 | 214.51 | 233.00 | 225.09 |

(i) At the 0.05 level of significance, is there evidence of any significant difference in the mean distances travelled by the golf balls with different designs?
(ii) If there are significant differences, carry out pairwise comparisons.
(b) A manufacturing company produces electric insulators. To test the strength of the insulators, destructive testing is carried out and force is measured in kilograms. The data of 30 values from this experiment are given below :
$1870,1728,1656,1610,1634,1784,1522,1696,1592,1662,1866,1764$, $1734,1662,1734,1774,1550,1756,1762,1866,1820,1744,1788,1688$, $1810,1752,1680,1810,1652,1736$
(i) Construct a continuous frequency distribution by computing suitable class width.
(ii) Calculate mean, median and standard deviation.
(iii) Construct box-plot.

