# POST GRADUATE DIPLOMA IN <br> APPLIED STATISTICS (PGDAST)/CERTIFICATE IN CONDITION MONITORING (CCOMO) 

Term-End Examination

December, 2022

## MST-005 : STATISTICAL TECHNIQUES

Time: 3 hours
Maximum Marks : 50
Note: Question no. 1 is compulsory. Attempt any four questions from the remaining questions no. 2 to 7. Use of scientific (Non-programmable) calculator is allowed. Use of Formulae and Statistical Tables Booklet for PGDAST is allowed. Symbols have their usual meanings.

1. State whether the following statements are True or False. Give reasons in support of your answers.

$$
5 \times 2=10
$$

(a) In cluster sampling, $\mathrm{s}_{\mathrm{w}}^{2}>\mathrm{s}_{\mathrm{b}}^{2}$, where $\mathrm{s}_{\mathrm{w}}^{2}$ represents the variance within clusters and $\mathrm{s}_{\mathrm{b}}^{2}$ between clusters.
(b) Consecutive 3 random numbers starting from 8937 by 'middle square method' are 8937, 8699 and 6726.
(c) If the sum of squares of errors in a two-way ANOVA having 4 rows and 5 columns is given as 48 , the mean sum of squares will be 4 for the same.
(d) If we draw all possible samples of size 2 from the population $\{2,3,4\}$, then $\mathrm{E}(\overline{\mathrm{x}})=\overline{\mathrm{X}}$.
(e) In $2^{3}$ factorial design, the number of factors are two each at three levels.
2. The table given below presents the summary of data of complete census of all the 450 farms of wheat in a region. The farms were stratified according to farm size (in acres) into 2 strata. The population values of strata means ( $\overline{\mathrm{X}}_{\mathrm{i}}$ ) and standard deviation $\left(\sigma_{i}\right)$ for the area under wheat are given as follows :

| Strata <br> No. | Farm <br> size (in <br> Acres) | No. of <br> Farms | Strata <br> Means | Standard <br> Deviation |
| :---: | :---: | :---: | :---: | :---: |
| 1. | $0-100$ | 300 | 45 | 15 |
| 2. | $100-200$ | 150 | 90 | 60 |

How would you draw the sample of size 45 using
(i) Proportional allocation,
(ii) Neyman allocation?

Also, obtain the variance of the estimate of the population mean for the proportional allocation and compare its efficiency with SRSWOR.
3. A researcher wants to test four diets $A, B, C$ and D on growth rate in mice. These animals are divided into 3 groups according to their weights. Heaviest 4, next 4 and lightest 4 are put in Block I, Block II and Block III, respectively. Within each block, one of the diets is given at random and after 15 days, increase in weight is noted as given in the following table :

| Block | Treatment/Diet |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | A | B | C | D |
| I | 12 | 8 | 6 | 5 |
| II | 15 | 12 | 9 | 6 |
| III | 14 | 10 | 8 | 5 |

Perform a two-way ANOVA to test whether the data indicates any significant difference between
(i) four diets,
(ii) three blocks.
4. Identify the design given in the following table and then carry out the analysis :

10

| Row | Column | II | III | IV |
| :---: | :---: | :---: | :---: | :---: |
| I | A 8 | C 18 | B 11 | D 8 |
| II | C 16 | B 10 | D 7 | A 4 |
| III | B 12 | D 10 | A 6 | C 20 |
| IV | D 10 | A 9 | C 28 | B 16 |

5. The following table gives the frequency distribution of 100 random numbers generated from $N(0,1)$ distribution :

| Class Interval | Frequency |
| :---: | :---: |
| $\leq(-2.5)$ | 02 |
| $(-2.5)-(-1.5)$ | 04 |
| $(-1.5)-(-1.0)$ | 08 |
| $(-1.0)-(-0.5)$ | 18 |
| $(-0.5)-0$ | 19 |
| $0-0.5$ | 12 |
| $0.5-1.0$ | 14 |
| $1.0-1.5$ | 14 |
| $1.5-2.0$ | 05 |
| $2.0-2.5$ | 02 |
| $2.5-3.0$ | 02 |

Use chi-square test of randomness to determine whether the fit of distribution is satisfactory or not.
6. (a) (i) Distinguish between linear and circular systematic sampling, with an example.
(ii) The data for production of wheat (in thousand kg ) in 25 districts is given by $23,20,30,37,76,36,13,36,16,58,53$, $83,10,15,13,17,12,16,17,21,20,13$, 61, 31, 71.
Select all possible systematic random samples of 7 units.
(b) The following data is related to the production (in kg ) of 3 varieties of wheat :

| P | 14 | 16 | 18 |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Q | 14 | 13 | 15 | 22 |  |
| R | 18 | 16 | 19 | 15 | 20 |

Is there any significant difference among the three varieties at $5 \%$ level of significance?
7. (a) For the data given below, the yield of treatment C in the second block is missing. Estimate the missing value and analyse the data.

| Block | Treatment |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | A | B | C | D |
|  | 105 | 114 | 108 | 109 |
| II | 112 | 113 | X | 112 |
| III | 106 | 114 | 105 | 109 |

(b) The distribution function of Pareto distribution is given by

$$
\mathrm{f}(\mathrm{x})=1-\left(\frac{\mathrm{k}}{\mathrm{x}}\right)^{\mathrm{a}} ; \quad \mathrm{a}>0,0<\mathrm{k}<\mathrm{x}
$$

Given a $\mathrm{U} \sim \mathrm{U}(0,1)$, generate a random number from the above distribution, when $a=2$ and $k=1$. Suppose $U=0 \cdot 5$, then find $x$.

