POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

June, 2021

MST-005 : STATISTICAL TECHNIQUES

Time : 3 hours *Maximum Marks : 50*

Note :

(i) *Question no. **1** is **compulsory**.*

(ii) *Attempt any **four** questions from the remaining questions no. 2 to 7.*

(iii) *Use of scientific calculator (non-programmable) is allowed.*

(iv) *Use of Formulae and Statistical 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) Sampling error arises at the stage of data processing.

(b) The units belong to a group that are homogeneous among themselves in cluster sampling.

(c) The error degrees of freedom for an Analysis of Variance of population means of 3 and 4 levels of two factors with total 12 observations will be 6.

(d) If there is a two-way heterogeneity, we use Randomised Block Design.

(e) For the given LCG

$x_i = (1478\ x_{i-1} + 21) \bmod (10^3)$,

the value of x_1 will be 71, when $x_0 = 75$.

2. (a) Differentiate between SRSWR and SRSWOR with examples. *4*

(b) The data of heights (in cm) of 8 students from each of 5 different classes are given as follows :

Student	Class 1	Class 2	Class 3	Class 4	Class 5
1	146	161	167	178	180
2	153	185	163	164	175
3	173	164	167	168	173
4	162	170	172	160	161
5	165	176	158	149	155
6	165	152	159	160	160
7	167	173	148	154	150
8	156	169	180	163	153

(i) How many first stage samples of size 2 and second stage samples of size 6 can be selected ?

(ii) Select one sample of size as described in (i).

(iii) Determine the sample mean under this sampling scheme. *6*

3. (a) Write a short note on the importance of the assumption of Normality in Analysis of Variance. *3*

(b) In order to test whether there is any significant difference in the average number of units produced per week by each of three production methods, the following data were collected :

Method I	Method II	Method III
18	19	17
16	18	18
17	19	19
18	17	19
17		16

Use $\alpha = 0{\cdot}05$ to test the difference among the units produced by the three methods. *7*

4. An experiment was conducted to investigate the effect of four different diets (A, B, C and D) on milk production of cows. There are four cows in the study. Loctation period and cows are considered as blocking variables. The data are summarised in the following table :

Period	Cow			
	1	2	3	4
1	A = 29	B = 36	C = 39	D = 38
2	B = 30	C = 35	D = 32	A = 28
3	C = 33	D = 38	A = 35	B = 30
4	D = 31	A = 33	B = 33	C = 30

Analyse the data at 1% level of significance and give your conclusions. *10*

5. The inter-arrival times of the customers arriving in a customer care centre of a mobile company has exponential distribution with rate $\alpha = 0{\cdot}5$ per minute. Simulate the time of six customers arriving at the centre. It is given that LCG is : *10*

$$x_i = (17\, x_{i-1} + 3) \bmod 16 \text{ with } x_0 = 15.$$

6. (a) A sample survey for estimating the number of orchards of apple was conducted in a district divided into four strata of the villages. A simple random sample without replacement of villages in each stratum was selected. The data of the number of villages (N_i), the average number of orchards ($\bar{y}_i$) and standard deviation of the number of orchards (S_i) are given below :

Stratum Number	N_i	$\bar{y}_i$	S_i
1	300	5	8
2	500	16	12
3	200	24	18

(i) Determine the sample sizes for 100 villages in each stratum under Neyman's allocation.

(ii) Also estimate the sample mean and the variance of the sample mean under the given sampling scheme. *7*

(b) Describe the Monte Carlo Simulation method. *3*

7. Four treatment combinations of nitrogen and phosphate were applied to a grain variety in a randomised complete block experiment. The yields (in kilograms) per plot are given as follows :

Block	(1)	n	p	np
1	4·5	6·7	6·6	9·0
2	4·9	9·5	8·2	8·2
3	5·7	8·1	8·4	6·6
4	5·7	6·5	5·6	5·0

Perform the analysis of variance on the given design at $\alpha = 0{\cdot}01$. *10*