

**POST GRADUATE DIPLOMA IN
APPLIED STATISTICS (PGDAST)**

Term-End Examination

02862

June, 2017

MST-005 : STATISTICAL TECHNIQUES

Time : 3 hours

Maximum Marks : 50

Note :

- (i) *Attempt all questions. Questions no. 2 to 5 have internal choices.*
- (ii) *Use of scientific calculator is allowed.*
- (iii) *Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.*
- (iv) *Symbols have their usual meaning.*

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

$5 \times 2 = 10$

- (a) The total number of all possible samples of size 2 without replacement from a population of size 7 is 21.

(b) Consecutive 3 random numbers starting from 8937 by 'middle square method' are 8937, 8699, 6726.

(c) RBD is suitable in situations where it is not possible to divide the experimental material into a number of homogeneous blocks.

(d) As we increase the sample size, representativeness of the population by the sample decreases.

(e) In a big hall, there are 50 rows and each row has 60 students. A research scholar selects 10 rows randomly and then randomly selects 15 students from each selected row. It is an example of cluster sampling procedure.

2. (a) Draw all possible samples of size 2 from the population {2, 3, 4} and verify that $E(\bar{x}) = \bar{x}$. Also find variance of \bar{x} .

5



- (b) A sample of 60 students is to be drawn from a population consisting of 600 students belonging to two villages, A and B. The means and standard deviations of their marks are give below :

2

Villages	Stratum sizes (N_i)	Means (\bar{x}_i)	Standard deviations (σ_i)
Village A	400	60	20
Village B	200	120	80

What are the sample sizes for the two villages using proportional allocation technique ?

- (c) Give three examples where only sampling techniques can be recommended instead of census.

3

OR

To determine the yield rate of wheat in a district of Punjab, 6 groups of 6 plots each were constructed. The data are given in the following table :

Plot No.	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
1	8	6	18	13	17	12
2	13	5	8	7	15	15
3	11	16	6	13	10	11
4	26	5	10	6	21	17
5	13	16	16	7	20	8
6	31	5	20	2	25	10

Select a cluster sample of 3 clusters from the above data and find its sample mean. Further, explain the procedure of two-stage sampling if we want to draw a sample of 6 plots. Which are the 6 plots in your sample ?

10

3. (a) The following data relate to production in kg of three varieties P, Q, R 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

- (b) Write any three assumptions of ANOVA.

3

OR

A researcher wants to test four diets A, B, C, 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 to the animals. After 15 days, increase in weight is noted, which is given in the following table :

Blocks	Treatments/Diets			
	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 the four diets due to different blocks.

10

4. In the following data, two values are missing. Estimate these values by Yates method and analyse the data by suitable technique.

10

Treatments	Blocks		
	I	II	III
A	12	14	12
B	10	y	8
C	x	15	10

OR

Identify the design given in the following table and then carry out the analysis :

10

Column Row	I	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. (a) The distribution function of Pareto distribution is given by $f(x) = 1 - \left(\frac{k}{x}\right)^a$, $a > 0$, $0 < k \leq x$.

Given a $U \sim U(0, 1)$, generate a random number from the above distribution, when $a = 2$ and $k = 1$. Suppose $U = 0.5$, then find x .

5

- (b) Generate a complete cycle for the LCG given below :

$x_i = (5x_{i-1} + 3) \bmod 16$, with $x_0 = 5$. A man tosses an unbiased coin ten times. Using the first ten random numbers generated above, obtain a sequence of heads and tails by taking Head (H) as $u \geq 0.5$.

5

OR



Times between successive crashes of a computer system were generated for a 6-month period and are given in increasing order as follows (time in hours) :

1 10 20 30 40 52 63 70 80 90 100 102 130
140 190 210 266 310 530 590 640 1340

The parameter $\alpha = 0.00435$, mean = $\frac{1}{\alpha} = 230$ hrs.

Use the Kolmogorov-Smirnov test to examine the goodness of fit of exponential distribution.

10
