



QP CODE: 25022335



Reg No :

Name :

M.Sc DEGREE (CSS) SPECIAL REAPPEARANCE EXAMINATION, APRIL 2025**Third Semester**

M.Sc BIOSTATISTICS

**CORE - ST020305 - STATISTICAL DATA MANAGEMENT USING
PYTHON,ADVANCED R AND SPSS**

2019 ADMISSION ONWARDS

8D29C7B7

Time: 3 Hours

Weightage: 30

(Answer any **three** questions. Each question carries a weight of **ten**)

1.

(i) A training programme to improve the total personality of young college students was conducted by an institute. The data presented below pertain to the scores of 10students with respect to their IQ(X) , Comprehension (Y), and Arithmetic Ability (Z) before and after undergoing the said training.

Students	Scores Before Training			Scores After Training		
	X	Y	Z	X	Y	Z
1	110	62	83	115	60	95
2	102	76	65	125	85	76
3	138	92	98	144	95	100
4	90	58	44	104	85	65
5	103	76	86	99	82	99
6	106	72	88	110	69	86
7	95	88	54	98	93	59
8	125	64	89	130	69	88
9	99	55	74	114	65	72
10	108	66	84	112	79	85

On the basis of the information furnished above, test whether you have any reason to believe that the training program is effective in improving the overall performance of the students.

(ii) The sum of squares and product matrix based on 50 observations from a 4 variate population is obtained as follows:





$$A = \begin{bmatrix} 13.0562 & 4.1740 & 8.9620 & 2.7332 \\ 4.1740 & 4.8250 & 4.0500 & 2.0190 \\ 8.9620 & 4.0500 & 10.8200 & 3.5820 \\ 2.7332 & 2.0190 & 3.5820 & 1.9162 \end{bmatrix}$$

Obtain the first 2 principal components and the percentage of total variance contained by them

2.

1. Hypothetical data of a new test for detecting Human Papilloma Virus (HPV) in a sample of 450 healthy women who attend for a cancer screening programme.

New Test	Gold standard test	
	HPV Positive	HPV Negative
HPV Positive	21	26
HPV Negative	2	401

Calculate sensitivity, specificity, positive and negative predictive value of the new test. Interpret the results.

2. The results from an experiment by Mc Leod (1944) on the effect of different concentrations of nicotine sulphate on *Drosophila melanogaster* are given below. Calculate LD₅₀.

Dose	No. of insects	No. killed	Mortality
0.50	1048	244	23
0.75	1202	451	38
1.10	1039	503	48
1.25	1102	638	58
1.50	1034	736	71
1.75	1167	880	75

3. The following stratified analysis has been constructed to illustrate a situation where cohort effects with regard to breastfeeding completely obscure a true protective association seen when age is controlled. Based on these hypothetical data:

	Age<60		Age>60		Total	
	Breastfed	Bottle-fed	Breastfed	Bottle-fed	Breastfed	Bottle-fed
Case	24	40	256	100	280	140
Control	79	86	204	54	280	140
OR	0.653		0.678		1.0	

- (a) Demonstrate that there is a cohort effect for breastfeeding.
- (b) Briefly explain (1-2 sentences referring to specific numbers or calculations for these tables) how failure to adjust for age interferes with finding a protective effect of breastfeeding.

3.

1. Four varieties A, B, C and D of a fertilizer are tested in LSD. The plot yields in pounds are as follows. Analyze the experimental yield.

A(22)	B(16)	C(12)	D(13)
B(19)	C(20)	D(17)	A(17)
C(18)	D(15)	A(18)	B(16)
D(19)	A(14)	B(19)	C(21)





2. Assuming the observation in the third row and fourth column to be missing. Estimate the missing value and carry out the Analysis.

4. 1. Let $X_0, X_1, \dots$ be a Markov sequence with initial distribution of X_0 given by: $(0.25, 0.25, 0.25, 0.25)$ and transition probability matrix

$$\begin{matrix} 0 & \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \end{bmatrix} \\ 1 & \begin{bmatrix} 0 & 0.5 & 0.5 & 0 \end{bmatrix} \\ 2 & \begin{bmatrix} 0 & 0 & 0.5 & 0.5 \end{bmatrix} \\ 3 & \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \end{bmatrix} \end{matrix}$$

(i) Find the probability distribution of X_1 and X_2

(iii) Determine $P(X_2=3/X_0=0)$, $P(X_3=2/X_0=1)$ and $P(X_3=0/X_0=0, X_1 \neq 0, X_2 \neq 0)$.

(iv) Find the probability of reaching state 3 from the state 1 after 3 steps for the first time.

(v) Find the expected number of times state 0 occurs in the first 100 transitions.

2. Suppose that the no. of calls per hr to an answering service follows a Poisson process with rate 4.

a) What is the probability that fewer than 2 calls came in the first hour

b) Suppose that 6 calls arrive in the first hour, what is the probability that 3 will be < 2 in the second hour?

c) Given that 6 calls arrive in the first 2 hours, what is the conditional probability that exactly 2 arrived in the first hour and exactly 4 in the second?

