Name : ……………………………………………………………

Roll No. : ………………………………………………………

Invigilator's Signature : ………………………………………

CS/M.TECH(CHE)/SEM-3/CHE-18/2012-13

2012

ADVANCED STATISTICAL ANALYSIS

Time Allotted : 3 Hours *Full Marks* : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Answer any *five* questions taking at least *one* from each Module.

$5 \times 14 = 70$

MODULE – I

1. What do you mean by n-dimensional space ? Prove that Hyperplane is a convex set. State the central limit theorem and show how it is applied for Markov Chain. 3 + 4 + 3 + 4

2. If $\mu_x = E(x) = \int_{-\infty}^{+\infty} xf(x)\,dx$ where x is continuous and $\sum xp(x)$ where x is discrete, then prove that $\mu_{np} = E(r) = np$ and $\sigma_{np} = E(r^2) = \sqrt{np(1-p)}$. 7 + 7

MODULE – II

3. To determine whether the reject $H_0 : \mu_1 = \mu_2$, we would compare t_0 to the t distribution with $n_1 + n_2 - 2$ degrees of freedom. If $| t_0 | > t_{\alpha/2,\ n_1+n_2-2}$ degrees of freedom, we would reject H_0 and conclude that the mean strength of the two formulations of Portland cement mortar differ. Assume that the mean tension bond strength of the two mortar formulations are equal. Tension bond strength data for the Portland cement formulation experiment are given below. You are to analyze accordingly. 14

Sl. No.	Modified Mortar (y_{1j})	Unmodified mortar (y_{2j})
1	16·85	17·50
2	16·40	17·63
3	17·21	18·25
4	16·35	18·00
5	16·52	17·86
6	17·04	17·75
7	16·96	18·22
8	17·15	17·90
9	16·59	17·96
10	16·57	18·15

4. A textile company weaves a fabric on a large number of looms. They would like the looms to be homogeneous so that they can obtain a fabric of uniform strength. The process engineering suspect that, in addition to the usual variation in strength within samples of fabric from the same loom, they may also be significant variations in the strength between looms. To investigate this she selects four looms at random and makes four strength determinations on the fabric manufactured on each loom. This experiment is run in random order, and data obtained are shown in the following table :

	Observation				
Looms	1	2	3	4	y_i
1	98	97	99	96	390
2	91	90	93	92	366
3	96	95	97	95	383
4	95	96	99	98	388
				Total	1527

Analyse the data and draw your conclusion. 14

MODULE – III

5. A soft drink bottler is interested in obtaining more uniform fill heights in the bottles produced by his manufacturing progress. The filling machine theoretically fills reach bottle to the correct target height, but in practice, there is variation around the target, and the bottler would like to understand better the sources of this variability and eventually reduce it.

   The process engineer can control three variables during the filling process. The per cent carbonation (A), the operating pressure in the filter (B), and the bottles produced per minute or the line speed (C). The pressure and speed are easy to control, but the per cent carbonation is more difficult to control during actual manufacturing because it varies with product temperature. However, for purposes of an experiment, the engineer can control carbonation at three levels : 10, 12 & 14 per cent. He chooses two levels for pressure (25 and 30 psi) and two levels for line speed (200 and 250 bpm). He decides to run two replicates of a factorial design in these three factors, with all 24 runs taken in random order. The response variable observed is the average deviation from the target fill height observed in a production run of bottles at each set of conditions. The data that resulted from the experiment are shown below :

	Operating pressure (B)				
Per cent carbonation (A)	25 psi		30 psi		y_i
	Line speed (C)		Line speed (C)		
	200	250	200	250	
10	– 3	– 1	– 1	1	– 4
	– 1	0	0	1	
12	0	2	2	6	20
	1	1	3	5	
14	5	7	7	10	59
	4	6	9	11	

Choose appropriate statistical model compute ANOVA and draw your conclusions. 14

6. What do you mean by Nested design ? Develop a model for a response with three way nested design. Enumerate different aspects of split-plot design. 3 + 6 + 5

MODULE – IV

7. What do you mean by the Kolmogorov-Smirnov and Anderson-Darling Tests ?

It is desired to check whether pinholes in electrolytic tin plate are uniformly distributed across a plated coil on the basis of the following distances in inches of 10 pinholes from one edge of a long trip of tin plate 30 inches wide :

4·8	14·8	28·2	23·1	4·4	28·7	19·5	2·4	25·0	6·2

Test the null hypothesis at the 0·05 level of significance.

6 + 8

8. What do you mean by Multiple Regression ? The following are data on the number of twists required to break a certain kind of forged alloy bar and the percentages of two alloying elements present in the metal :

Number of twists (y)	Per cent of element A x_1	Per cent of element B x_2
41	1	5
49	2	5
69	3	5
65	4	5
40	1	10

Contd..

Number of twists (y)	Per cent of element A x_1	Per cent of element B x_2
50	2	10
58	3	10
57	4	10
31	1	15
36	2	15
44	3	15
57	4	15
19	1	20
31	2	20
33	3	20
43	4	20

Fit a least square regression plane and use its equation to estimate the number of twists required to break one of the bars when $x_1 = 2 \cdot 5$ and $x_2 = 12$. 14

=============