Name : ....

Roll No. : ....

Invigilator's Signature : ....

CS/M.TECH(CHE)/SEM-3/CHE-18/2011-12

2011

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 Group.

GROUP - A

1. 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

2. The Poisson distribution can be used to approximate the Binomial distribution for the situations in which n is large and p is very small – prove this statement mathematically.

GROUP - B

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 $| t_0 | > t_{a/2}$ with $(n_1 + n_2 - 2)$ 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 :

j	*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. What do you mean by coefficient of contingency ?

The following table shows the observed and expected frequencies in tossing a die 120 times. Test the hypothesis that the die is fair, using a significance level of 0·05

Die Face	1	2	3	4	5	6
Observed Frequency	25	17	15	23	24	16
Expected Frequency	20	20	20	20	20	20

4 + 10

GROUP - C

5. A soft drink bottler is interested in obtaining more uniform fill heights in the bottles produced by his manufacturing process. The filling machine theoretically fills each 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 in 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	

Analyse the data and draw your conclusions.

6. The folloing table shows the yields in bushels per acre of a certain variety of wheat grown in a particular type of soil treated with urea, potash and super-phosphate. Find

a) the mean yield for the different treatments

b) the grand mean for all the treatments

c) the total variation

d) the variation between treatments

e) the variation within treatments.

Urea	48	49	50	49
Potash	47	49	48	48
Super-phosphate	49	51	50	50

GROUP - D

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

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 hypthesis 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	Percent of element B x_2
41	1	5
49	2	5
69	3	5
65	4	5
40	1	10
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 $x_2 = 12$.