

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.

5 × 14 = 70

1. a) Show that

$$\begin{aligned} f(x) &= x, & 0 < x < 1 \\ &= k - x, & 1 < x < 2 \\ &= 0, & \text{otherwise} \end{aligned}$$

is a density function for a suitable value of k .

Calculate $P(1/3 < x < 1/2)$.

b) In a certain factory blades are manufactured in packets of 10. There is a 0.2% probability for any blade to be defective. Using Poisson distribution calculate approximately the number of packets containing two defective blades in a consignment of 20,000 packets.

$$(e^{-0.02} = 0.9802)$$



- c) In a normal distribution, 31% of the terms are under 45 and 8% are above 64. Find the mean and standard deviation.

(Given that $P (0 < Z < 1.405) = 0.42$ and

$$P (- 0.496 < Z < 0) = 0.19)$$

4 + 5 + 5

2. a) If the equation of two Regression lines obtained in a correlation analysis are $3x + 12y - 19 = 0$ and

$9x + 3y = 46$ determine which one is Regression equation of y on x and which one is the regression equation of x on y . Find the means of x on y and correlation coefficient between x and y .

- b) Show that $- 1 \leq r_{xy} \leq 1$, for any bivariate data given by (x, y) , where r_{xy} is correlation coefficient of x and y .

- c) The bivariate data (x, y) results the followings :

$$\sum x_i y_i = 414, \quad \sum x_i = 120, \quad \sum y_i = 90, \quad \sum x_i^2 = 600, \\ \sum y_i^2 = 300.$$

Calculate the correlation coefficient between x and y if the number of data is 30.

5 + 5 + 4



3. a) The following marks have been obtained by **Mechanics** and **statistics** (out of 100) :

<i>Mechanics</i> :	55	56	58	60	65	68	70	75	80	85
<i>Statistics</i> :	50	48	60	62	64	65	70	74	82	90

Compute the coefficient of correlation for the above data. Find also the equations of lines of regression.

b) In a certain experiment to compare two types of pig foods *A* and *B*, the following results of increase in weights were observed in pigs :

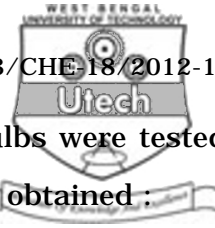
<i>Pig number</i>		1	2	3	4	5	6	7	8	Total
<i>Increase in weight in lb</i>	Food A	49	53	51	52	47	50	52	53	407
	Food B	52	55	52	53	50	54	54	53	423

- i) Assuming that the two samples of pigs are independent, can we conclude that food *B* is better than food *A* ?
- ii) Also examine the case when the same set of eight pigs were used in both the foods. 7 + 7

4. a) Random samples of 400 men and 600 women were asked whether they would like to have a flyover near their residence. 200 men and 325 women were in favour of the proposal. Test the hypothesis that proportions of men and women in favour of the proposal are same against that they are not at 5% level of significance.



- b) i) What is null hypothesis ?
- ii) Define critical region.
- c) The mean weekly sale of soap bars in departmental store was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful ? 5 + 4 + 5
5. a) In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 ozs with a standard deviation of 12 ozs, while the corresponding figures in a sample of 400 items from the other process are 124 ozs and 14 ozs. Obtain the standard error of difference between the two sample means. Is this difference significant ? Also find the 99% confidence limits for the difference in the average weights of items produced by the two processes respectively.



- b) Samples of two types of electric light bulbs were tested for length of life and following data were obtained :

	<i>Type-I</i>	<i>Type-II</i>
Sample No.	$n_1 = 8$	$n_2 = 7$
Sample means	$\bar{x}_1 = 1234$ hrs	$\bar{x}_2 = 1036$ hrs
Sample Standard deviations	$S_1 = 36$ hrs	$S_2 = 40$ hrs

Is the difference in the means sufficient to warrant that Type-I is superior to Type-II regarding length of life ?

[Given $t_{0.10; 13} = 1.77$ (two tail tables)] 7 + 7

6. A study of iron deficiency among infants compared samples of infants following different feeding regiments. One group contained breast-fed infants, while the children in other group were fed a standard baby formula without any iron supplements. Here are summary results on blood hemoglobin levels at 12 months of age.

<i>Group</i>	<i>n</i>	<i>Mean</i>	<i>S</i>
Breast-fed	23	13.3	1.7
Formula	19	12.4	1.8



- a) Is there significant evidence that the mean hemoglobin level is higher among breast-fed babies? State null hypothesis and alternate hypothesis and conduct a t -test.
- b) Give a 95% confidence interval for the mean difference in hemoglobin level between the two populations of infants. 14

7. a) Let five specimens of a certain kind of cold drinks have an average fat content of 12.1% and a standard deviation of 0.4%. Find out the 95% confidence interval for the population standard deviation.

$$[\text{Given } \chi^2_{0.025;4} = 11.143, \chi^2_{0.975;4} = 0.484]$$

- b) In measuring the radius of ball bearing manufactured by a machine an engineer estimates that the standard deviation is 0.05 cm. How large a sample of measurements must be taken in order to be 95% confident that the error of his measurement of average radius will not exceed 0.01 cm ?

$$[\text{Given } \int_0^{1.96} \phi = 0.475] \quad 7 + 7$$



8. a) Write short notes on the following :

- i) Standard error
 - ii) Type-I and Type-II errors
 - iii) Central limit theorem
 - iv) Random effects.
- b) A cigarette manufacturing firm claims that its brand A of the cigarettes outsells its brand B by 8%. If it is found that 42 out of a sample of 200 smokers prefer brand A and 18 out of another random sample of 100 smokers prefer brand B. Test whether the 8% difference is a valid claim.
- (Use 5% level of significance). 6
-