

**INDIAN INSTITUTE OF TECHNOLOGY, KHARAGPUR**  
**END SEMESTER EXAMINATION**

Date: 27-11-2012 FN

Time: 3 hours

Full Marks: 50

Autumn Semester : 2012-2013

Department: Mathematics

Subject No.: MA31020

Subject Name: Regression and Time Series Models

Course: B.Tech. 3<sup>rd</sup> Year (AG, BT, HS, IM, MI)

No. of students: 46

**Instructions:** Answer any five questions. Statistical tables may be used. Choose appropriate levels of significance in different problems.

1. (i) Using the method of least squares analyze a simple linear regression model. Derive estimates of parameters of the model and show that they are unbiased.  
(ii) A chemical engineer is investigating the effect of process operating temperature on product yield. The study results in the following data. Temperature is in  $^{\circ}\text{C}$  and increase in yield is in percentage.

Temp (X)	100	110	120	130	140	150	160	170	180	190
Yield (Y)	45	51	54	61	66	70	74	78	85	89

Draw the scatter diagram. Fit a simple linear regression model. Find 95% confidence intervals for  $\beta_0$ ,  $\beta_1$  and  $\sigma^2$ . Further test for the significance of the regression coefficient. Find the coefficient of determination. How much regression does the model explain? Find a 95% confidence interval for the expected percentage increase in yield at temperature  $^{\circ}145$ . Also determine a 95% prediction interval for the percentage increase in yield at temperature  $^{\circ}200$ .

2. The following data is collected to relate the damage (Y measured in mm) susceptibility of peaches to the height at which they are dropped (drop height  $X_1$  measured in mm) and the density of the peach ( $X_2$  measured in  $\text{g/cm}^3$ ):

Y	3.62	7.27	2.66	1.53	4.91	10.36	5.26	6.09	6.57	4.24
$X_1$	303.7	366.7	336.8	304.5	346.8	600	369	418	269	323
$X_2$	0.9	1.04	1.01	0.95	0.98	1.04	0.96	1	1.01	0.94

Fit a multiple regression model for the data. Find 95% confidence intervals for the coefficients of the model and  $\sigma^2$ . Test for the significance for the regression coefficients. Find  $R^2$  and find the percentage of regression that the model explains. Predict the damage when  $X_1 = 460$  and  $X_2 = 0.99$ . Give your comments on the suitability of the model.

3. (i) Using the method of least squares show that we can analyze a polynomial regression model. Give the necessary forms of estimates and tests, confidence intervals.  
(ii) Sidewall panels for the interior of an airplane are formed in a 1500-ton press. The unit manufacturing cost varies with the production lot size. The data shown below give the average cost (Y) per unit (in thousands rs.) for the product and the production lot size (X).

Y	1.81	1.7	1.65	1.55	1.48	1.4	1.3	1.26	1.24	1.21	1.2	1.18
X	20	25	30	35	40	50	60	65	70	75	80	90

Show with the help of a scatter diagram that a second degree curve is appropriate for the data and fit this model. Test for the significance of the model.

4. Give the method of analyzing a nonlinear models  $y = ae^{bx}$ . Find the estimates, tests and confidence intervals for a and b. Apply these to the following data:

X	1	2	3	4	5	6	7	8
Y	15.3	20.5	27.4	36.6	49.1	65.6	87.8	117.6

5. Discuss the methods for diagnosing multicollinearity. What are standard methods for treating the problem?
6. Explain various models of time series with the help of examples. What are trend and seasonality. Write appropriate models for these. Define stationary time series, autocorrelation function and its sample analogue.

7. (a) What is the basic assumption in time series analysis modelling.

- (b) Do you think that the assumption is being satisfied for the following model viz.,

$$(i) \tilde{Z}_t = a_t,$$

$$(ii) \tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t,$$

where  $\{a_t\}$  is a white noise process or shocks and  $\tilde{Z}_t$  is a mean deviated observation at time t ?

- (c) What do you mean by a partial autocorrelation function ?

What are the characteristics of autocorrelation and partial autocorrelation functions for an Autoregressive process of order 2 (AR(2)) ?

- (d) How do estimate the parameter  $\theta$  of the following model as given by

$$\tilde{Z}_t = a_t - \theta a_{t-1} ?$$

- (e) Suppose that a mean deviated time series  $\{\tilde{Z}_t\}$  has the following underlying model

$$\tilde{Z}_t = -0.8\tilde{Z}_{t-1} + a_t$$

If the value of  $\tilde{Z}_t = 3.8$  at  $t = 100$ , would you expect  $\tilde{Z}_{101}$  to be above the mean of the process or below ?