Indian Institute of Technology, Kharagpur
Mid-Autumn Semester Examination: 2018–2019

Date of Examination:........-09-2018 Session (FN/AN), *Duration: 2 Hrs,*
Subject. No. MA31020/MA41025 *No. of Registered Students (123+64)=187*
Subject Name: REGRESSION AND TIME SERIES MODEL
Department: Mathematics *TOTAL MARKS: 30*
Specific Chart, graph paper log book etc. required.... STATISTICAL TABLE...
Special Instruction: ***Begin to answer each question in a new page. Answer all parts of a question in a coherent place. Full credit will be given for the answers which are correct up to FOUR decimal places.*** ANSWER ALL THE QUESTIONS

1. (a) Define the orthogonal complement space ($S^{\perp}$) of a subspace S of $\mathbb{R}^3$.
   (b) Are S and $S^{\perp}$ disjoint?
   (c) Let $S = \{(x, y, 0)|2x + 3y = 0\}$. Find $S^{\perp}$. [2+1+2]

2. Consider the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ where ϵ_i s are independently and identically distributed $N(0, \sigma^2)$ variables. Here x variable stands for the length of a pendulum in $\log_{10}$ scale and y variable stands for the measured time period of it in the same scale.

x	1.04	1.09	1.16	1.12	1.17	1.25	1.26	1.29
y	0.818	0.845	0.899	0.865	0.890	0.946	0.938	0.935

   (a) Provide the least squared estimates of the model parameters β_0, β_1, σ. [3+2]
   (b) Find the 95% prediction interval of the time period for length 10 unit in usual scale.

3. Consider the multiple linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$ where, $\mathbf{Y} = (y_1, y_2, \cdots y_n)^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \cdots, \beta_k)^T$, $\mathbf{X} = (\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ and $\epsilon = (\epsilon_1, \epsilon_2, \cdots \epsilon_n)^T \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$.
   (a) What is the non-centrality parameter (λ,say) of the model sum of squared (SSModel).
   (b) Plot the above λ as a function of β_0. Justify the graph with a proof. [1+4]

4. Consider the multiple linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$ where, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)^T$ and $\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$. Sample size $n = 32.$, $SSRes = 193.1176$. Test for the null hypothesis $H_0 : \beta_2 = \beta_4 = 0$ Vs $H_1 : H_0$ is not true at 5% level when, $(\mathbf{X}^T\mathbf{X})^{-1} =$

$$\begin{bmatrix} 17.42309 & -0.159620 & 0.007268 & -0.014045 & -0.077966 \\ -0.159620 & 0.023686 & -0.001697 & -0.000985 & 0.000948 \\ 0.007268 & -0.001697 & 0.000778 & -0.000094 & -0.000085 \\ -0.014045 & -0.000985 & -0.000094 & 0.000543 & -0.000356 \\ -0.077966 & 0.000948 & -0.000085 & -0.000356 & 0.000756 \end{bmatrix}$$

   and $\hat{\boldsymbol{\beta}} = (84.26902, -3.06981, 0.00799, -0.11671, 0.08518)$. [5]

5. Consider the multiple linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$.
   (a) Define linear zero function (LZF).
   (b) Define the best linear unbiased estimator (BLUE) of a parametric function $\mathbf{p}^T\boldsymbol{\beta}$.
   (c) Prove that BLUE is unique for a parametric function $\mathbf{p}^T\boldsymbol{\beta}$. [1+1+3]

6. State TRUE or FALSE in complete words for the followings: [5]
   (a) An orthogonal projection matrix has eigen values $\{-1, +1\}$.
   (b) Simple linear regression line always passes through $(\bar{x}, \bar{y})$.
   (c) Diagonal elements of $(\mathbf{X}\mathbf{X}^T)^-$ are indicators of leverage.
   (d) Polynomial regression model is not a linear model.
   (e) Adjusted R^2 is a non-decreasing function of the number of regressor variables.

The following tabulated values may be of use:

$t_{0.025,7} = 2.365, t_{0.025,8} = 2.306, t_{0.05,6} = 1.943, t_{0.025,6} = 2.447, F_{0.05,2,26} = 3.37, F_{0.05,3,26} = 2.98$

****************** THE END ******************