INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

End-Spring Semester 2016-17

Date of Examination : 24/04/2017(AN) Duration 3hours Marks: 50

Subject No. : MA60056 (Regression and Time Series Models)

Department: Mathematics

Students: PGDBA core + BTech/MSc elective (92)

Questions:

1. State whether the following statements are true or false. In each case, justify your answer. [10 marks]

   1.1. Let $\{Z_t\}$ be a sequence of independent normal random variables with mean 0 and variance σ^2. Then for any real constants a and b, the time series defined by $X_t = a + bZ_t$ is a stationary series.

   1.2. Let $\{X_t\}$ be a time series where all the random variables X_t are identically distributed. Then the series $\{X_t\}$ is strictly stationary.

   1.3. Every iid sequence of random variables is a stationary series but not necessarily strictly stationary.

   1.4. Let $\{X_t\}$ and $\{Y_t\}$ are uncorrelated stationary processes, that is, X_r is uncorrelated with Y_s for every r and s. Then the process $\{X_t + Y_t\}$ is also stationary.

   1.5. Consider a standard regression problem $y_i = \alpha_0 + \alpha_1 x + \epsilon_i$ where $i = 1,2, \dots, n$ and $\epsilon_i \sim N(0, \sigma^2)$ are iid for $i = 1,2, \dots, n$. The maximum likelihood estimators for the parameters α_0 and α_1 are same as those of the least squares estimators.

2. Let $Y = X\beta + \epsilon$ be a multiple linear regression model where $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$, $\beta \in \mathbb{R}^p$ and $\epsilon \sim N(0, \sigma^2 I_n)$. Derive the principal component regression estimator of β. Is this estimator an unbiased estimator of β? Show that the variance of this PCA estimator is smaller than that of the least squares estimator. [8 marks]

3. Compute the auto covariance function of the time series $X_t = Z_t + 0.3Z_{t-1} - 0.4Z_{t-2}$ where $\{Z_t\}$ is a white noise process with mean zero and variance 1. [4 marks]

4. Let $\{Z_t\}$ be an iid $N(0,1)$ process.

$$\text{Define } X_t = \begin{cases} Z_t & for\ t\ even \\ (Z_t^2 - 1)/\sqrt{2} & for\ t\ odd \end{cases}$$

Then show that $\{X_t\}$ is a white noise process with mean 0 and variance 1. [4 marks]

5. Let X_1, X_2, X_4, X_5 be observations from the time series model $X_t = Z_t + \theta Z_{t-1}$ where Z_t is a white noise process with mean 0 and variance σ^2 and θ is a real constant.

   5.1. Compute the best linear estimate of the missing value X_3 from X_1 and X_2. [2 marks]

   5.2. Compute the best linear estimate of the missing value X_3 from X_4 and X_5. [2 marks]

   5.3. In each of the above cases, compute the mean squared errors of the estimates. [4 marks]

6. Consider $y_i = x_i^T\beta + \epsilon_i$ for $i = 1,2,\dots,n$ be a regression model where $x_i \in \mathbb{R}^p$ are the observed values of regressors, $\beta \in \mathbb{R}^p$ is the parameter vector to be estimated and noise $\epsilon_i \sim N(0, \sigma_i^2)$ are independent. In such a setting, pose the weighted regression problem stating the need for doing so. Derive the expression for the estimator of β in the weighted regression case. Discuss whether this estimator is unbiased. Further, compute the variance of this estimator. [8 marks]

7. Tests for checking if a given time series data $x_1, x_2, \dots, x_n$ is a realization of an iid sequence of length n.

   7.1. The turning point test: A turning point at time i is defined as a point if $x_{i-1} < x_i$ and $x_i > x_{i+1}$ or $x_{i-1} > x_i$ and $x_i < x_{i+1}$. Let T denotes the number of turning points in the given time series data. What is the distribution of this random variable T in case of an iid data of length n. Using this construct a test to check if the given time series data is a realization of an iid sequence of length n. [4 marks]

   7.2. The difference-sign test: Let S denote the number of time instances i such that $x_i - x_{i-1}$ is positive. Construct a test using this random variable S to check if the given time series data is a realization of an iid sequence of length n. [4 marks]