Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Spatial Statistics (MTH643A)

Quiz 1, Date: August 28, 2023, Monday

Time: 10:05–10:50 AM (45 minutes) | Fall 2023-24 | Max point: 10(+2 bonus)

---

1. Write down the category for each of the following data structures. (2 points)

   a The coordinates of the wildfire locations during July 2023 across Madhya Pradesh.

   b A processed image of polarimetric measurements of Earth obtained by SHAPE from the Propulsion Module of Chandrayaan 3 from the lunar orbit.

   c At 10 AM on August 28, 2023, relative humidity data collected by 26 weather stations across Delhi.

   d Number of functioning colleges across the districts of Uttar Pradesh as of today.

2. Write down the four different types of point-referenced spatial data and give an example in each case. (2 points)

3. We have T Gaussian processes $X_t(\cdot) \overset{IID(over\ t)}{\sim} \mathcal{GP}[\mu(\cdot), K(\cdot,\cdot)]$ defined over the same domain $\mathcal{D} \subset \mathbb{R}^2$. Show that a stochastic process $\overline{X}(\cdot)$ constructed by $\overline{X}(s) = T^{-1}\sum_{t=1}^{T} X_t(s)$ for all $s \in \mathcal{D}$ will also be a Gaussian process. You are only allowed to use the definition of a Gaussian process. (2 points)

4. Suppose $X \sim \mathrm{N}(0, \sigma^2)$ and $W(s) \overset{IID}{\sim} \mathrm{N}(0, \tau^2)$ for all $s \in \mathcal{D} \subset \mathbb{R}^2$. We define $Y(s) = X + W(s)$ for all $s \in \mathcal{D}$. We know that $Y(\cdot)$ is a valid Gaussian process. Suppose $Y(\cdot) \sim \mathcal{GP}[0, K(\cdot,\cdot)]$. Write down the mathematical expression of the kernel $K(s, s')$. Show that it is a valid covariance kernel. (1 + 2 points)

5. For a Gaussian process $Y(\cdot) \sim \mathcal{GP}[0, K(\cdot,\cdot)]$, where $K(s, s') = \sigma^2 \exp[-(\|s - s'\|/\phi)^2] + \tau^2\delta(s = s')$ and the ratio of the partial sill to nugget is 1:4. Write down the expression of the semivariogram. Find the effective range of $Y(\cdot)$. (1+2 points)

$\gamma(h) = E(Y(s) - Y(s'))^2$

$=$

Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Spatial Statistics (MTH643A)

Quiz 2, Date: September 11, 2023, Monday

Time: 10:05–10:50 AM (45 minutes) | Fall 2023-24 | Max point: 10(+2 bonus)

1. Suppose we know the Cholesky factors of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. Write down an efficient algorithm to draw a sample (multivariate) from $\boldsymbol{Y} \sim \mathrm{MVN}_n(\boldsymbol{X\beta}, \boldsymbol{A}'\Sigma_1\boldsymbol{A} + \boldsymbol{B}'\Sigma_2\boldsymbol{B} + \tau^2\boldsymbol{I}_n)$. You should justify your algorithm. You are allowed to draw a maximum of $4n$ IID samples from Normal$(0,1)$ distribution and necessary basic matrix operations. The algorithm and the justification should be clearly and separately written. (1.5+1.5 points)

2. We have a hierarchical spatial model constructed as

$$\boldsymbol{Y}|\boldsymbol{Z} \sim \mathrm{MVN}_n(\boldsymbol{X\beta} + \boldsymbol{BZ}, \tau^2\boldsymbol{I}_n),\ \boldsymbol{Z} \sim \mathrm{MVN}_p(0, \sigma^2\Sigma).$$

Assuming a frequentist setting, calculate $\pi(\boldsymbol{Z}|\boldsymbol{Y})$. (2 points)

3. Suppose we have spatial data $Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n)$. We can safely assume that the data follows a spatial Gaussian process $Y(\cdot) \sim \mathcal{GP}[0, K(\cdot,\cdot)]$ defined over $\mathcal{D} \subset \mathbb{R}^2$, with $K(\boldsymbol{s}, \boldsymbol{s}') = \sigma^2 \exp\left[-\frac{(x-x')^2}{4\phi^2} - \frac{4(y-y')^2}{\phi^2}\right]$, where $\boldsymbol{s} = (x,y)'$ and $\boldsymbol{s}' = (x',y')'$. Write down the maximum likelihood estimation steps via profile likelihood. Please be to the point. (3 points)

4. We have spatial data $Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n)$ from a stochastic process with constant unknown mean μ but known covariance kernel $K(\cdot,\cdot)$. The data vector is denoted by $\boldsymbol{Y}$. Suppose $\boldsymbol{\Sigma}$ is given by its $(i,j)^{th}$ element $\Sigma_{i,j} = K(\boldsymbol{s}_i, \boldsymbol{s}_j), i,j = 1,\ldots,n$. We estimate μ ensuring the Euclidean norm $\|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{Y} - \mu\boldsymbol{1}_n)\|$ is minimized. Write down the estimator for μ. Then, write down the kriging equation for predicting $Y(\boldsymbol{s}_0)$. Show that the sum of the kriging weights is one. (1+1+2 points)

$\frac{1' \Sigma^{-1} y}{1' \Sigma^{-1} 1}$ $\quad Y(s_0) = \hat{\mu} + \Sigma_0' \Sigma^{-1}(y - \hat{\mu} 1)$

Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Spatial Statistics (MTH643A)

Quiz 3, Date: October 18, 2023, Wednesday

Time: 11:05–11:50 AM (45 minutes) Fall 2023-24 Max point: 10(+2 bonus)

Under Vision 2030, Saudi Arabia has proposed a project called "The Line". The idea is to build a futuristic city that has only a single horizontal direction and a vertical direction. Thus, our common assumption of the spatial domain $\mathcal{D} \subset \mathbb{R}^2$ would not apply to reasonable inferential problems if we are at a fixed elevation, i.e., we can think of $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 : ax + by = c\}$ for some known a, b, c. Suppose we are in the year 2100. This is the setup for the following questions.

1. We have a total of 10^6 temperature and humidity sensors across the basement of "The Line", and at a specific time point, we have observed data from each of them. Thus, we have a total of 2×10^6 observations. Write down a valid statistical model for this dataset that assumes a separable covariance structure across sensors and also across temperature and humidity). All assumptions should be stated. (2 points)

   Write down the covariance matrix (in a clean format) of the data vector (the stacking of components should be shown) and the corresponding precision matrix. For a wrong stacking, you will lose marks in both components. (1+1 = 2 points)

2. Suppose we now focus on a valid covariance kernel over "The Line" for humidity (after removing the mean) that is weakly stationary. Give an example of a valid spectral density in this context. Only fully correct answers will be given marks. (1 point)

   If the spectral density is radially symmetric, then show that the resulting covariance function is isotropic. All notations should be appropriate and in the context of "The Line". (2 points)

3. Now suppose we have annual average temperature data from all sensors between the years 2030 and 2099 and we want to assume a weakly stationary spatiotemporal covariance function for the underlying spatiotemporal process. Write down a version of the spectral representation theorem here. Give an example of a valid spectral density here. Only fully correct answers will be given marks. (1.5+1=2.5 points)

   Show that if the spectral density is separable across space and time, the resulting covariance function will also be separable across space and time. All notations should be appropriate and in the context of "The Line". (2.5 points)

Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Spatial Statistics (MTH643A)

Quiz 4, Date: November 6, 2023, Monday

Time: 10:05–10:50 AM (45 minutes) Fall 2023-24 Max point: 10(+2 bonus)

Suppose the data at n areal units are denoted by $Y_1, \ldots, Y_n$. The adjacency matrix is denoted by $\boldsymbol{A}$ and $\boldsymbol{M}$ denotes the diagonal matrix with its diagonal entries being the row-sums of $\boldsymbol{A}$. Corresponding to the i-th areal unit, we have some covariate information available, which we denote by the P-length vector $\boldsymbol{X}_i$. Follow the usual notations (σ, ρ, etc.) used in the class. Let $\boldsymbol{X}$ denote the $n \times P$-dimensional design matrix and the corresponding vector of regression coefficients be $\boldsymbol{\beta}$.

1. Write down a conditional autoregressive model, i.e., the conditional distribution of Y_i given $\{Y_j : j \sim i\}$ that takes care of the covariates (a linear regression model). Then, write down the joint distribution of $\boldsymbol{Y}$ using Brook's lemma. (1+1=2 point)

2. In Question 1, write down the maximum likelihood estimator of $\boldsymbol{\beta}$ (assuming other parameters to be known). Write down the closed form expression (without matrices) of the maximum likelihood estimator of $\boldsymbol{\beta}$ when $P = 1$ and $\boldsymbol{X} = \mathbf{1}_n$. (2+1=3 points)

3. Obtain an upper bound of the spatial autocorrelation parameter ρ assuming the covariance matrix of $\boldsymbol{Y}$ to be positive definite, in terms of the eigenvalues of $\boldsymbol{M}^{-1/2}\boldsymbol{A}\boldsymbol{M}^{-1/2}$. (2 points)

4. Show that the covariance matrix of $\boldsymbol{Y}$ is singular/non-invertible if $\rho = 1$. (1 point)

5. Write down a simultaneous autoregressive model that takes care of the covariates (a linear regression model). Then, derive the joint distribution of $\boldsymbol{Y}$. (1+1=2 point)

6. In Question 5, write down the closed form expression (without matrices) of the maximum likelihood estimator of $\boldsymbol{\beta}$ when $P = 1$ and $\boldsymbol{X} = \mathbf{1}_n$. You can use the standard notations used in the class. (2 points)

Indian Institute of Technology Kanpur

Department of Mathematics and Statistics

Spatial Statistics (MTH643A)

Mid-semester exam, Date: September 20, 2023, Wednesday

Time: 8:00–9:30 AM (90 minutes) Fall 2023–24 Max point: 20(+5 bonus)

Result1: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is an invertible square matrix and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ are column vectors. Then $\boldsymbol{A}+\boldsymbol{u}\boldsymbol{v}'$ is invertible if and only if $1+\boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u} \neq 0$. In this case, $(\boldsymbol{A}+\boldsymbol{u}\boldsymbol{v}')^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}'\boldsymbol{A}^{-1}}{1+\boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u}}$.

Result2: $(\boldsymbol{A}+\boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1}+\boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}$ for conformable matrices.

Result3: $|\boldsymbol{A}+\boldsymbol{U}\boldsymbol{C}\boldsymbol{V}| = |\boldsymbol{C}^{-1}+\boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}| \times |\boldsymbol{A}| \times |\boldsymbol{C}|$.

1. Write down the category for each of the following data structures (choose one of the three categories only and a sub-category should be considered as its main category). (4 points)

   a The coordinates of the landslide locations during June 2023 across Himachal Pradesh.

   b A processed climate model output for annual average temperature in 2050 available over a $1° \times 1°$ grid.

   c Precipitation data from 1218 weather stations in the US on September 19, 2023.

   d Constituency-wise projected winning probabilities (treat them as data) of a political party in 2024 assessed by the Indian Political Action Committee (IPAC).

   e Daily gridded rainfall data from NASA-JAXA joint mission Tropical Rainfall Measuring Mission (TRMM) during 2000–2020 available over a $0.5° \times 0.5°$ grid.

   f Binary gridded daily rainfall occurrence data available during 1950–2020 over a $0.25° \times 0.25°$ grid across India.

   g Annual maxima of daily average temperature data available during 1950–2020 over a $0.25° \times 0.25°$ grid across India.

   h The year-wise coordinates of the cyclogenesis locations during 1900–2020 across the North Indian Ocean.

2. We construct a low-rank spatial model by

$$Y(\boldsymbol{s}) = \mu + \sum_{l=1}^{L} \cos[\boldsymbol{\omega}'\boldsymbol{s}]\alpha_l + \sum_{l=1}^{L} \sin[\boldsymbol{\omega}'\boldsymbol{s}]\beta_l + \varepsilon(\boldsymbol{s}),$$

where $\boldsymbol{\omega}$ is a known frequency, $\alpha_l, \beta_l \stackrel{IID}{\sim} \mathrm{N}(0, \sigma^2 g(\omega))$ and $\varepsilon(\boldsymbol{s}) \stackrel{IID}{\sim} \mathrm{N}(0,\tau^2)$, and $\alpha_l, \beta_l, \varepsilon(\boldsymbol{s})$ are independent of each other. Comment on whether the process is weakly stationary or not. Justify your answer. No point will be given without justification. (3 points)

3. Suppose we have spatial data $\boldsymbol{Y} = [Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n)]'$. We assume that the data follow a spatial stochastic process (not necessarily a Gaussian process) with unknown but constant mean μ and covariance kernel $K(\boldsymbol{s}, \boldsymbol{s}') = \sigma^2 + \tau^2\delta(\boldsymbol{s} = \boldsymbol{s}')$, where $\sigma^2/\tau^2 = \rho$. Suppose $\boldsymbol{\Sigma}$ is given by its $(i,j)^{th}$ element $\Sigma_{i,j} = K(\boldsymbol{s}_i, \boldsymbol{s}_j), i,j = 1,\ldots,n$. We estimate μ ensuring the Euclidean norm $\|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{Y} - \mu\boldsymbol{1}_n)\|$ is minimized. Derive the estimator for μ in terms of $\boldsymbol{Y}$ and ρ only (in clean form, without any matrix). Write down the kriging equation in terms of $\boldsymbol{Y}$ and ρ only (in clean form, without any matrix). (2+2=4 points)

4. We have time series data $Y(t_1), \ldots, Y(t_n)$ from a zero-mean Gaussian process with covariance kernel $K(\cdot,\cdot)$ given by $K(t,t') = \sigma^2 \exp(-|t-t'|/\phi)$. Here $t_{i+1} = t_i + h, i = 2, \ldots, n-1$. The data vector is denoted by $\boldsymbol{Y}$ and $n = 10^6$. Write down the maximum likelihood estimation steps via profile likelihood. Please be to the point. No matrix computation is allowed within the optimization algorithm. Any closed form of the estimator σ^2 for a fixed ϕ should be derived. The objective function should be shown without unnecessary terms. (4 points)

5. In Question 4, suppose we know ϕ (or plug in an empirical Bayes estimate, irrelevant to this question). Choose a conjugate prior for σ^2 and derive the posterior distribution of σ^2. No matrix computation is allowed. (0.5+1.5=2 points)

6. A multi-resolution fixed rank kriging model is given by

$$Y(s) = \boldsymbol{X}(s)'\boldsymbol{\alpha} + \sum_{r=1}^{R}\sum_{l=1}^{L_r} B_{r,l}(s)\beta_{r,l} + \varepsilon(s), \ \forall s \in \mathcal{D}$$

where $B_{r,l}(\cdot), l = 1, \ldots, L_r$ are known spatial radial basis functions at r^{th} resolution, $\varepsilon(s) \overset{iid}{\sim} \mathrm{N}(0,\tau^2)$, $\boldsymbol{\beta}_r = (\beta_{r,1}, \ldots, \beta_{r,L_r})' \sim \mathrm{MVN}_{L_r}(\boldsymbol{0}, \Sigma_r), r = 1, \ldots, R$. Here, $\boldsymbol{\beta}_r$ and $\boldsymbol{\beta}_{r'}$ are independent for $r \neq r'$. Let $\boldsymbol{s}^*_{r,1}, \ldots, \boldsymbol{s}^*_{1,L_r}$ be the knot locations at the r^{th} resolution and $\mathrm{Cov}(\beta_{r,l}, \beta_{r,l'}) = \sigma_r^2 \exp[-\|\boldsymbol{s}^*_{r,l} - \boldsymbol{s}^*_{r,l'}\|/\phi_r]$.

Write down the joint distribution of $\boldsymbol{Y} = [Y(s_1), \ldots, Y(s_n)]'$. The covariance matrix of $\boldsymbol{Y}$ should be written in such a way that it can be calculated as efficiently as possible (without unnecessary matrix multiplications). Write down an efficient algorithm to draw a (multivariate) sample from $\boldsymbol{Y}$. Justify your algorithm. (2+1.5+1.5=5 points)

Note that you can write the model as a hierarchical spatial model as

$$\boldsymbol{Y}|\boldsymbol{\beta} \sim \mathrm{MVN}_n(\boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{B}\boldsymbol{\beta}, \tau^2\boldsymbol{I}_n), \ \boldsymbol{\beta} \sim \mathrm{MVN}_p(\boldsymbol{0}, \Sigma)$$

for certain forms of $\boldsymbol{X}$, $\boldsymbol{B}$, p, and Σ. State their forms if required. Assuming a frequentist setting, calculate $\pi(\boldsymbol{\beta}|\boldsymbol{Y})$. The necessary matrices should be written in such a way that they can be calculated efficiently (without unnecessary matrix multiplications). (3 points)

Indian Institute of Technology Kanpur
Department of Mathematics and Statistics
Spatial Statistics (MTH643A)
End-semester exam, Date: November 19, 2023, Sunday

Time: 8:00–9:30 AM (90 minutes) Fall 2023–24 Max point: 20

Points will be deducted for writing unnecessarily long answers

1. We have time series data $Y(t_1), \ldots, Y(t_n)$ from a zero-mean Gaussian process with covariance kernel $K(\cdot, \cdot)$ given by $K(t, t') = \sigma^2 \exp(-|t - t'|/\phi)$. Here $t_{i+1} = t_i + h, i = 1, 2, \ldots, n-1$. The data vector is denoted by $\boldsymbol{Y}$. Write down the likelihood function of σ^2 and ϕ without involving any matrices. (2 points)

2. Suppose we have spatial data $Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n)$. However, we realize that there is no spatial structure and we model them using $Y(\cdot) \sim \mathcal{GP}[\mu(\cdot), K(\cdot, \cdot)]$ defined over the domain $\mathcal{D} \subset \mathbb{R}^2$, with $\mu(\boldsymbol{s}) = \mu$ and $K(\boldsymbol{s}, \boldsymbol{s}') = \delta(\boldsymbol{s} = \boldsymbol{s}')$. We choose the prior $\mu \sim \text{Normal}(0, 10^2)$. Calculate the posterior mean and variance of μ. (2 points)

3. Suppose we know the Cholesky factors of $\boldsymbol{\Sigma}_i, i = 1, \ldots, R$. Write down an efficient algorithm (point-by-point) to draw a sample (multivariate) from $\boldsymbol{Y} \sim \text{MVN}_n(\boldsymbol{X\beta}, \sum_{r=1}^{R} \boldsymbol{A}_r' \boldsymbol{\Sigma}_r \boldsymbol{A}_r + \tau^2 \boldsymbol{I}_n)$. You should justify your algorithm, i.e., you should show that the final sample you draw is actually from this multivariate normal distribution with the same mean vector and covariance matrices. You are allowed to draw a maximum of $(R+3)n$ IID samples from Normal$(0, 1)$ distribution and necessary basic matrix operations. (2 points)

4. We have spatial data $Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n)$ from a zero-mean stochastic process with covariance kernel $K(\cdot, \cdot)$ given by $K(\boldsymbol{s}, \boldsymbol{s}') = \exp(-\|\boldsymbol{s} - \boldsymbol{s}'\|/\phi)$. The data vector is denoted by $\boldsymbol{Y}$. We use a tapering function $C_T(h) = 1$ if $h < \gamma$ and zero otherwise. Write down the mathematical expression of the L_1 and L_2 distance between the original and approximated covariance matrices. (1+1=2 points)

5. For a L-variate Gaussian process $\boldsymbol{Y} \sim \text{GP}_L(\boldsymbol{\mu}(\cdot), \boldsymbol{K}(\cdot, \cdot))$ defined over $\mathcal{D} \subset \mathbb{R}^2$, write down the domain and range of the functions $\boldsymbol{\mu}(\cdot)$ and $\boldsymbol{K}(\cdot, \cdot)$. (1+1=2 points)

6. Show that a zero-mean bivariate normal spectral density with covariance matrix $\phi^2 \boldsymbol{I}_2$ leads to a squared exponential correlation structure of the underlying Gaussian process. (2 points)

gauss fn

7. Show that a Kernel convolution leads to a isotropic squared exponential covariance structure when we choose a squared exponential kernel with a fixed bandwidth. (2 points)

8. For a discrete-time spatiotemporal CAR model, where the observations are independent across time, given by $Y_{i,t} | \boldsymbol{Y}_{-i,t} \sim \text{Normal}(\mu + \rho[\bar{Y}_{i,t} - \mu], \sigma^2/m_i), i = 1, \ldots, n, t = 1, \ldots, T$, calculate the MLE of μ and σ^2 assuming ρ to be known. Here the notations are standard from the class notes except that t denotes the t-th time point. (1+1=2 points)

9. Suppose we have a non-homogeneous Poisson process (NHPP) defined over the domain $\mathcal{D} = [0, 1]^2$. Let us denote the intensity function by $\lambda(\boldsymbol{s}) = 5 + 2x + 2y$, where $\boldsymbol{s} = (x, y)'$. Write down the spatial density function. Calculate the probability of a random location lying within the area $\mathcal{D} = [0, 1/2] \times [0, 1]$. (1+1=2 points)

$e(\cdot x(s)' \beta +$

10. Write down a sampling scheme for the NHPP in Question 9. (2 points)

MTH-515a: Inference-II

2023-2024: I Semester
Quiz I

Time Allowed: 45 Minutes **Maximum Marks: 25**

1. Let X be a single random observation from a distribution having the pdf

$$g_\theta(x) = \begin{cases} 3(x-\theta)^2, & \text{if } \theta \le x \le \theta+1 \\ 0, & \text{otherwise} \end{cases},$$

where $\theta \in (-\infty, \infty) = \Theta$ is an unknown parameter. Consider problem of estimating θ under the loss function $L(\theta, a) = |a - \theta|$, $a \in \mathcal{A} = (-\infty, \infty)$, $\theta \in \Theta$. Show that the estimation problem is invariant under the additive group of transformations $\mathcal{G} = \{g_c : c \in \mathbb{R}\}$, where $g_c(x) = x + c$, $x \in \mathbb{R}, c \in \mathbb{R}$. Find the MRIE of θ under the additive group of transformations. $\boxed{\text{5+8=13 Marks}}$

2. Let X be a single random observation from $\mathrm{U}(\theta, 2\theta)$ distribution, where $\theta \in \Theta = (0, \infty)$ is an unknown parameter. Consider the problem of estimating θ^2 under the loss function $L(\theta, a) = (\frac{a}{\theta^2} - 1)^2$, $a \in \mathcal{A} = (0, \infty)$, $\theta \in \Theta$. Show that the estimation problem is invariant under the multiplicative group of transformations $\mathcal{G} = \{g_c : c > 0\}$, where $g_c(x) = cx$, $x \in \mathbb{R}, c > 0$. Find the MRIE of θ under the multiplicative group of transformations. $\boxed{\text{5+7=12 Marks}}$

MTH-515a: Inference-II

2023-2024: I Semester
Mid Semester Examination

Time Allowed: 120 Minutes | Maximum Marks: 50

NOTE: (i) Start answer of every question on a new page. Moreover, attempt all the parts of a question at one place.
(ii) Answer each question legibly, clearly and concisely.
(ii) For each question, provide details of all the steps involved in arriving at the final answer(s)/conclusion(s) and box your final answer(s)/conclusion(s).

1. (a) Let $X_1, \ldots, X_n$ be a random sample such that X_1 has a p.d.f.

$$g_{\underline{\theta}}(x) = \begin{cases} \frac{e^{-(x-\theta)}}{1-e^{-1}}, & \text{if } \theta \leq x \leq \theta + 1 \\ 0, & \text{otherwise} \end{cases},$$

where $\theta \in \mathbb{R} = \Theta$, say. Consider estimation of θ under the loss function $L(\theta, a) = (a-\theta)^2$, $\theta \in \Theta, a \in \mathcal{A} = \mathbb{R}$. Find the MRIE of θ.

(b) Let $X_1, \ldots, X_n$ be a random sample such that X_1 has a p.d.f.

$$g_{\underline{\theta}}(x) = \begin{cases} \frac{1}{2} e^{-\frac{x-\theta}{2}}, & \text{if } x \geq \theta \\ 0, & \text{otherwise} \end{cases},$$

where $\theta \in \mathbb{R} = \Theta$, say. Consider estimation of θ under the loss function $L(\theta, a) = |a-\theta|$, $\theta \in \Theta, a \in \mathcal{A} = \mathbb{R}$. Find the MRIE of θ. $\boxed{\text{7+7=14 Marks}}$

2. (a) Let $X_1, \ldots, X_n$ be a random sample from from $U(-\theta, \theta)$, where $\theta \in (0, \infty) = \Theta$, say. Consider estimation of θ^2, under the loss function $L(\theta, a) = |\frac{a}{\theta^2} - 1|, \theta \in \Theta, a \in \mathcal{A} = (0, \infty)$. Find the MRIE of θ^2. (Hint: $T = \max\{|X_1|, \ldots, |X_n|\}$ is a complete-sufficient statistic).

(b) Let $X_1, \ldots, X_n$ be a random sample from $U(\theta, 2\theta)$, where $\theta \in (0, \infty) = \Theta$, say. Consider estimation of θ under the loss function $L(\theta, a) = (\frac{a}{\theta} - 1)^2, \theta \in \Theta, a \in (0, \infty) = \mathcal{A}$, say. Find the MRIE of θ. $\boxed{\text{7+5=12 Marks}}$

3. Let $X_1, \ldots, X_n$ $(n \geq 2)$ be a random sample such that X_1 has a p.d.f.

$$g_{\underline{\theta}}(x) = \begin{cases} \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}}, & \text{if } x \geq \mu \\ 0, & \text{if } x < \mu \end{cases},$$

where $\underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty) = \Theta$, say. Consider estimation of $\eta = \mu + \sigma$ under the loss function $L(\underline{\theta}, a) = (\frac{a-\eta}{\sigma})^2, \underline{\theta} \in \Theta, a \in \mathbb{R} = \mathcal{A}$.

(a) Identify a group of transformations under which the above estimation problem is invariant;

(b) Find the MRIE of η. $\boxed{\text{4+8=12 Marks}}$

P.T.O.

4. Let $X_1, \ldots, X_n$ ($n \geq 2$) be a random sample from $N(\mu, \sigma^2)$, where $\underline{\theta} = (\mu, \sigma) \in \mathbb{R} \times (0, \infty) = \Theta$, say.

(a) Consider estimation of μ under the loss function $L(\underline{\theta}, a) = \frac{(a-\mu)^4}{\sigma^4}, \underline{\theta} \in \Theta, a \in \mathbb{R} = \mathcal{A}$. Find the MRIE of μ under the affine group of transformations.

(b) Consider estimation of σ under the loss function $L(\underline{\theta}, a) = \frac{(a-\sigma)^2}{\sigma^2}, \underline{\theta} \in \Theta, a \in \mathbb{R} = \mathcal{A}$. Find the MRIE of σ under the affine group of transformations. $\boxed{6+6=12 \text{ Marks}}$

Quiz 1: MTH441A

Question 1: Consider the following simple linear regression model without the intercept

$$y_i = \beta x_i + \epsilon_i \qquad \text{for} \quad i = 1, \ldots, n,$$

where $\{\epsilon_1, \ldots, \epsilon_n\}$ are independently distributed with mean 0 and variance σ^2. Find

(i) the least square estimate $\widehat{\beta}$ of β. (2 marks)

(ii) the expected value of $\widehat{\beta}$. (1 marks)

(iii) the variance of $\widehat{\beta}$. (2 marks)

Question 2: Consider the following multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \epsilon_i \qquad \text{for} \quad i = 1, \ldots, n,$$

where $\{\epsilon_1, \ldots, \epsilon_n\}$ are independently distributed with mean 0 and variance σ^2. Show that

$$\sum_{i=1}^{n} \text{Var}(\hat{y}_i) = (p+1)\sigma^2. \qquad \text{(5 marks)}$$

Question 3: The adjusted R^2 for a linear regression model is defined as

$$R^2_{\text{Adj}} = 1 - \frac{SS_{\text{Res}}/(n-p)}{SS_{\text{T}}/(n-1)}.$$

Find $\mathbb{E}[R^2_{\text{Adj}}]$. (5 marks) $e^{-\frac{x}{\theta}}$

$\chi^2_n = \Gamma(\frac{n}{2}, 2)$

Mid Sem: MTH441A **6 PM to 8 PM: 2 Hours**

Question 1: Consider the model

$$y_i = x_i\beta + \epsilon_i$$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$ are independently distributed with mean 0 and variance $\mathrm{Var}(\epsilon) = \mathrm{V} = \mathrm{Diag}(1/\omega_1, \ldots, 1/\omega_n)$. Find

(i) Find a general formula for the weighted least-squares estimator of β. **(2 marks)**

(ii) What is the variance of the weighted least-squares estimator? **(2 marks)**

(iii) Suppose that $\mathrm{Var}(y_i) = cx_i$. Using the results of parts (i) and (ii), find the weighted least- squares estimator of β and the variance of this estimator. **(3 marks)** $\frac{\bar{y}}{\bar{x}}$ $\frac{c}{n\bar{x}}$

(iv) Suppose that $\mathrm{Var}(y_i) = cx_i^2$. Using the results of parts (i) and (ii), find the weighted least-squares estimator of β and the variance of this estimator. **(3 marks)**

$\frac{1}{n}\sum \frac{y_i}{x_i}$ $\frac{c}{n}$

Question 2: Consider the following multiple linear regression model

$$y = X\beta + \epsilon$$

where $\mathbb{E}[\epsilon] = \mathbf{0}$ and $\mathrm{Var}(\epsilon) = \sigma^2\mathrm{V}$. Verify that

$$[y'\mathrm{V}^{-1}y - y'\mathrm{V}^{-1}X(X'\mathrm{V}^{-1}X)^{-1}X'\mathrm{V}^{-1}y]/(n-p)$$

is an unbiased or biased estimate of σ^2. **(5 marks)**

Hint: $\mathbb{E}[z'\mathrm{A}z] = \mathrm{trace}\{\mathrm{AV}\} + \mu'\mathrm{A}\mu$ where $\mathbb{E}[z] = \mu$ and $\mathrm{Var}(z) = \mathrm{V}$.

Question 3: If $\boldsymbol{Y} = (Y_1, Y_2)' \sim \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$, prove or disprove that

$$(\boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - \frac{Y_2^2}{\sigma_{22}}) \sim \chi_1^2.$$ **(5 marks)**

Question 4: Consider a multiple linear regression model with usual assumptions and n observations and $p-$ regressors. Derive the distribution of Residual Sum of Squares divided by σ^2. **(5 marks)**

Quiz 2: MTH441A **1 Hours**

NAME: **Roll Number:**

MCQ: Choose the correct options (more than one correct options might be possible) in questions 1 to 5. Each question carries 2 marks.

Question 1: Consider the following simple regression model with usual assumptions

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad i = 1, \ldots, n. \tag{1}$$

Let $\bar{y} = \sum_{i=1}^{n} y_i/n$, $\widehat{\beta}_1$ is the least square estimate of β_1 and $\mathrm{Var}(\epsilon_i) = \sigma^2$. The covariance between $\bar{y}$ and $\widehat{\beta}_1$ is

(a) $\sigma^2/\sum_{i=1}^{n}(x_i - \bar{x})^2$

(b) 0

(c) σ^2/n

(d) $\sum_{i=1}^{n}(x_i - \bar{x})$.

Answer:

Question 2: Consider the model (1). Define $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$. When σ^2 is known, an appropriate test statistics for testing the hypothesis $H_0 : \beta_1 = \beta_{10}$ versus $H_0 : \beta_1 \neq \beta_{10}$ is

(a) $Z_0 = (\widehat{\beta}_1 - \beta_{10})/\sqrt{\sigma^2/S_{xx}}$ and $Z_0 \sim \mathcal{N}(0,1)$.

(b) $t_0 = (\widehat{\beta}_1 - \beta_{10})/\sqrt{MS_{Res}/S_{xx}}$ and $t_0 \sim t_{n-2}$.

(c) $Z_0 = (\widehat{\beta}_1 - \beta_{10})/\sqrt{\sigma^2/S_{xx}}$ and $Z_0 \sim t_{n-2}$.

(d) $W_0 = (\widehat{\beta}_1 - \beta_{10})^2/(\sigma^2/S_{xx})$ and $W_0 \sim \chi^2_1$.

Answer:

Question 3: Consider the following regression model with n observations and p parameters in matrix form

$$y_{n\times 1} = X_{n\times p}\beta_{p\times 1} + \epsilon_{n\times 1}.$$

(a) Let $H = X(X'X)^{-1}X'$ and h_{ii} is the ith diagonal element of H. Then ith observation is a **leverage** point if $h_{ii} > 2p/n$.

(b) The ith observation is **influential**/outlier if the associated COOK's distance is greater than 1.

(c) The ith observation is **influential**/outlier if the associated $|DFFITS_i|$ is greater than $2\sqrt{p/n}$.

(d) The ith observation is **influential**/outlier if the associated $|DFBETAS_{j,i}|$ is greater than $2/\sqrt{n}$.

Answer:

Question 4:

Ans

(a) Subset selection criterion based on residual mean square is equivalent to the criterion based on adjusted R^2.

(b) For a subset model if Mallows's C_p Statistic is closed to p then that model is not recommended.

(c) Forward selection and Backward elimination methods lead to the same subset model.

(d) Forward selection and Step wise methods lead to the same subset model.

Question 5: Consider the model

(a) Rigde estimator is an unbiased estimator of the regression parameters.

(b) Principal–Component Regression can be used to address the issue of multicollinearity.

(c) VIF is a measure of multicollinearity.

(d) The data collection method can not be a problem of multicollinearity.

Answer:

Question 6: Consider the following regression model with usual assumptions

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i, \quad i = 1, \ldots, n.$$

(i) Write an appropriate test statistics for testing $H_0 : T\beta = 0$ versus $H_1 : T\beta \neq 0$, where

$$T = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \text{ (1 mark)}$$

(ii) What is the distribution of the test statistics you proposed? (1 mark)

Question 7: Let $Y \sim Gamma(\alpha, \beta)$ with α known. The density of Y is given by

$$f(y|\alpha, \beta) = \frac{\beta^\alpha y^{\alpha-1} e^{-\beta y}}{\Gamma(\alpha)}, \quad 0 < y < \infty.$$

(i) Is Y a member of exponential family distribution? (1 mark)

(ii) Find the mean and variance of Y? (2 mark)

α/β, α/β^2

End Sem: MTH441A **Duration: 2 Hours**

Question 1: Consider the following arrangement of complete randomized block design (CRBD) with blocks $i = a, b, c, d$ and treatments $j = 1, 2, 3, 4$.

Table 1: CRBD

Block	Treatment			
a	1	2	3	4
b	2	1	4	3
c	3	4	2	1
d	4	3	1	2

The following anova model is used

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij},\ i = a, b, c, d;\ j = 1, 2, 3, 4,$$

where μ is the overall mean parameter, α_i and β_j are parameters associated with ith block and jth treatment effects respectively. Following constraints are imposed for the identifiablity $\sum_{i=a}^{d} \alpha_i = 0$ and $\sum_{j=1}^{4} \beta_j = 0$.

Table 2: Data

18
9

Block	Treatment				
a	3	2	5	6	16
b	2	4	7	8	21
c	9	2	1	5	17
d	5	6	4	3	18

The CRBD given in Table 1 is used for the experiment and the responses are given in Table 2. Find the least square estimates of μ, α_i and β_j for $i = a, b, c, d$; and $j = 1, 2, 3, 4$. (7 Marks)

Question 2: Consider a Poisson regression model with the following linear predictor

$$\eta_i = \sum_{j=1}^{p} x_{ij}\beta_j;\quad i = 1, \ldots, n;\ j = 1, \ldots, p$$

and probability mass function

$$f(y_i; \mu_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$$

(a) Find the **canonical link** function relating mean to the linear predictor. (2 Marks)

(b) Find the mean and variance of y_i using the form of exponential family distribution. (3 Mark)

(c) Find the asymptotic variance covariance matrix of the least square estimator of $\beta = (\beta_1, \ldots, \beta_p)^{\mathrm{T}}$. (3 Marks)

Question 3:

(i) Consider the following regression model:

$$y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \cdots + \beta_{p-1}(x_{i,p-1} - \bar{x}_{p-1}) + \epsilon_i;\ i = 1, \ldots, n,$$

where $\bar{x}_j = \sum_i x_{ij}/n$. Find the LSE of β_0 and $\beta = (\beta_1, \ldots, \beta_{p-1})'$. Express your solution in a compact form. (4 Marks)

(ii) Consider a particular case of above regression model when $n = p = 3$, i.e.

$$y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i,2} - \bar{x}_2) + \epsilon_i,$$

$\mathbf{x}_1 = (1, 0, 0)'$, $\mathbf{x}_2 = (0, 1, 0)'$ and $\mathbf{y} = (1, 2, 3)'$. Find the least square estimators of $(\beta_0, \beta_1, \beta_2)$. (4 Marks)

Question 4: Define

$$X_A = \begin{pmatrix} X_{n\times p} \\ \sqrt{k}\mathrm{I}_p \end{pmatrix} \qquad y_A = \begin{pmatrix} y \\ \mathbf{0}_p \end{pmatrix}$$

where I_p is a square matrix of order $p \times p$ and $\mathbf{0}_p$ is a $p \times 1$ vector of zeros. Moreover, $X'X$ is a diagonal matrix with jth diagonal element as $j - k$ for $j = 1, \ldots, p$. Now consider the following model

$$y_A = X_A\beta + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2\mathrm{I}_n)$. Find $\sum_{j=1}^{p} \mathrm{Var}(\widehat{\beta}_j)$ where $\widehat{\beta}_j$ is the least square estimator of β. (7 Marks)

MTH516 : MidSem

Date: 22.09.2023

Time: 120 Minutes **Total Marks: 25**

Note: We follow usual notations from class lectures and assignments. Answer all parts of one question together. **State clearly any results that you may use.**

1

1. State Parzen's lemma and the Nadaraya-Watson (N-W) estimator.
   Use Parzen's lemma to prove consistency of the N-W estimator. State all necessary conditions. [1.5+3.5]

2. ~~(a)~~ Define a median of $X \sim F$. Let $f \equiv F'$. Define $\phi(a) = E_f|X - a| < \infty$ for $a \in \mathbb{R}$. Show that $\phi(a)$ is minimized by a median of X.

   ~~(b)~~ Let $\bar{X}_i$ and S_i^2 be the sample means and variances for an independent sample of size n_i for the i-th population with $1 \le i \le k$. Assume them to be realizations of a common cdf F which is symmetric about $\mu \in \mathbb{R}$. Define $S = \sum_{i=1}^{k} \frac{\bar{X}_i S_i}{\sum_{i=1}^{k} S_i}$.
   Prove that S is symmetric about μ. [(1+2)+2]

3. (a) Let X and Y be two non-negative random variables with infinite expectation. Define $Z_1 = \min(X, Y)$ and $Z_2 = \max(X, Y)$. Is it possible that
   (i) only Z_1 has finite expectation; (ii) only Z_2 has finite expectation; (iii) both Z_1 and Z_2 have finite expectations; (iv) one has finite expectation, but not the other one?
   In each case, justify your argument by give a proof, or an example.

   (b) Let $X_{(1)} \le \cdots \le X_{(n)}$ be order statistics based on a sample of size n from the standard Cauchy distribution. Fix $x \in \mathbb{R}$. Compute $\lim_{n\to\infty} P(n^{-1} X_{(n)} \le x)$. [3+2]

4. (a) Let $X_1, \ldots, X_n$ be i.i.d. from a continuous distribution function (df) F. Define the empirical df F_n of F. For each fixed $x \in \mathbb{R}$, prove that $F_n(x) \xrightarrow{P} F(x)$ as $n \to \infty$.

   (b) State Hellinger's statistic (say, H_n). For the k-category case, now prove that $H_n \xrightarrow{D} H \sim \chi^2_{(k-1)}$ as $n \to \infty$. [(1+1)+(1+3)]

5. State Wilcoxon's signed rank statistic W_n^+ and the associated testing problem.
   Under the null hypothesis, prove/compute the following: (i) W_n^+ is distribution-free, (ii) $W_n^+ \stackrel{D}{=} \sum_{j=1}^{n} jZ_j$ with $Z_j \stackrel{\text{i.i.d.}}{\sim} Ber(1/2)$ for $1 \le j \le n$, and (iii) $Var(W_n^+)$ by using part (ii).

   [1+(1∗3)]

$\frac{t_{mn}(nw)}{N'(N+1)}$ $(x_i - x_j)(y_i - y_j) > 0$

MTH 516A: End-semester

$K(x,\theta) = \frac{1}{n}\sum K$

Date: 21.11.2023

Time: 150 Minutes
$2\frac{1}{2}$ hr.

Total Marks: 25

We follow usual notations from class lectures. Answer ALL questions. Each question is of *five marks* only.

1. (a) Define the kernel density estimator (KDE). Derive an expression for the bias of the KDE. Give a choice of the kernel function K so that the bias is 0.
   (b) Let $\mathbf{X} \sim N_p(0, \Sigma)$. Find the distribution of $\|\mathbf{X}\|^2$.
   Let P be an orthogonal matrix. Is it true that $\mathbf{X} \stackrel{D}{=} P\mathbf{X}$? Give clear arguments. [2+(2+1)]

2. (a) Let $X_1, \ldots, X_4$ be an iid sample from a continuous df F, which is symmetric about 0. Use the equal in distribution technique to compute $P(X_1 + X_2 + X_3 + X_4 > 0)$.
   (b) State Kendall's sample correlation coefficient (say, K_n) for paired data (X, Y), and interpret it. Relate K_n with an appropriate U-statistic.
   If X and Y are independent, then prove that K_n is symmetric about 0. [2+(1+2)]

3. (a) Let $X_1, \ldots, X_n$ be an iid sample from a df F with $E_F(X_1^2) < \infty$. Derive the asymptotic distribution of $n^{-1}\sum\sum_{i \neq j} X_i X_j$ as $n \to \infty$.
   (b) For the usual two sample formulation, find an estimator of $\sigma_X^2 + \sigma_Y^2$. Establish its asymptotic normality under appropriate conditions. [2+(1+2)]

4. Let $X_1, \ldots, X_n$ be an iid sample from a continuous df F.
   (a) Define Walsh averages. Express Wilcoxon's signed rank statistic in terms of the Walsh averages.
   (b) Consider the median of all Walsh averages. Is this estimator better than the usual median? Justify your answer. [(1+2)+2]

5. (a) Let $(Z_1, \ldots, Z_n)$ be a vector of exchangeable rvs from a continuous df F. Give an example of such an F. Find the joint distribution of the vector of ranks (say, $(R_1, \ldots, R_n)$) using clear arguments.
   (b) Prove that the sign test is consistent. Compare it with the one sample t-test using the concept of ARE. [2.5*2]

$\frac{n}{2} + \sqrt{\frac{n}{2}}\, x$

$P_{1n} - \frac{1}{2}$

All the best!

$\frac{1}{n} I(x_i + x_j > 0)$

$P_{1n} > T_{1n}$

MTH442A: Time Series Analysis
Midsem Examination: Full Marks 60

[1] Let $X_t = (-1)^t A + \epsilon_t$; where $A \sim N(0,1)$ and $\{\epsilon_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. Further, A is independent of the sequence $\{\epsilon_t\}$. Prove or disprove the following statements:

(a) $\{X_t : t = 1,2,\ldots\}$ is covariance stationary with $Cov(X_{14}, X_{11}) = -1$.

(b) If $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, then $\{Y_n : n = 1,2,\ldots\}$ is covariance stationary.

(c) If $P_t = \epsilon_t + \epsilon_{4t+3}$, then $\{P_t : t = 0, \pm1, \pm2, \ldots\}$ is a white noise.

(d) If $Q_t = \epsilon_{2t} + \epsilon_{2t+1}$, then $\{Q_t : t = 0, \pm1, \pm2, \ldots\}$ is a white noise.

4+5+4+4

[2] (a) Consider the following real-valued function $f : \mathbb{Z} \to \mathbb{R}$, defined on the set of integers $\mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$:

$$f(h) = \begin{cases} 1, & h = 0 \\ 1, & h = \pm 1 \\ 0, & \text{otherwise.} \end{cases}$$

Verify whether $f(h)$ is ACVF or not.

(b) Let $Y_t = (a + b\,t)S_t + \epsilon_t$; a and b are deterministic fixed constants, S_t is seasonal component with periodicity 6 and $\{\epsilon_t\}$ is a sequence of i.i.d. $N(0,1)$ random variables. Let $\{Z_t\}$ denote the EWMA series obtained from $\{Y_t\}$ with $\alpha = \frac{2}{3}$; i.e. $Z_1 = Y_1$ and for all $t \geq 2$, $Z_t = \alpha\, Y_t + (1-\alpha)\, Z_{t-1}$.

(i) Find the joint distribution of the random vector $(Z_1, Z_3, Z_4)'$.

(ii) Is $\{Z_t\}$ strict stationary?

(iii) Prove or disprove: "A moving average filter, of window length 6, applied on $\{Y_t\}$ eliminates the seasonal component".

7+8

[3] Suppose A and B are independent $N(0,1)$ random variables and $\{\epsilon_t\}$ and $\{\delta_t\}$ are two mutually independent sequence of i.i.d. $N(0,1)$ random variables. Further, $A, B, \{\epsilon_t\}$ and $\{\delta_t\}$ are also mutually independent. Define

$Y_t = (\epsilon_t + i\,\delta_t)(A\cos(\omega t) + i\,B\sin(\omega t))$ and

$Z_t = (\epsilon_t + i\,\delta_t) + (A\cos(\omega t) + i\,B\sin(\omega t))$; where, $i = \sqrt{-1}$ and $\omega = \frac{\pi}{2}$.

Prove or disprove the following statements:

(a) The 2 complex-valued time series $\{Y_t\}$ and $\{Z_t\}$ are both covariance stationary.

(b) $\{Re(Y_t)\}$ is a real-valued covariance stationary time series; where, $Re(Y_t)$ is the real-valued time series part of $\{Y_t\}$.

(c) $Cov(Y_{t+h}, Y_t) = 0$ for all $h \neq 0$

(d) $Cov(Z_{t+h}, Z_t) = 0$ for all $|h| \geq 3$

6+4+2+2

[4] (a) Consider the following $ARMA(2,1)$ process

$$X_t = \frac{3}{2} X_{t-1} - \frac{1}{2} X_{t-2} + \epsilon_t + 2\,\epsilon_{t-1};$$

where, $\epsilon_t \sim WN(0, \sigma^2)$ and $Cov(\epsilon_t, X_{t-j}) = 0$ for all $j > 0$.

(i) Is $\{X_t\}$ covariance stationary?

(ii) Prove or disprove " ∇X_t is covariance stationary $ARMA(1,1)$".

(b) Consider the following covariance stationary $ARMA(1,1)$ process

$$Y_t = \frac{1}{3} Y_{t-1} + \epsilon_t + \epsilon_{t-1};$$

where, $\epsilon_t \sim WN(0, \sigma^2)$ and $Cov(\epsilon_t, Y_{t-j}) = 0$ for all $j > 0$.

Find $\gamma_Y(0), \gamma_Y(2), \gamma_Y(6)$ of the $\{Y_t\}$ process, as a function of σ^2 only.