IMPERIAL

MATH70139

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Spatial Statistics

Date: Friday, May 31, 2024

Time: 14:00 – 16:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer All Questions in 1 Answer Booklet

Candidates should start their solutions to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

1. (a) Define what it means for a random field $X = (X_t)_{t \in \mathbb{R}^d}$ to be:

(i) Strictly stationary. (2 marks)

(ii) Weakly stationary. (2 marks)

(iii) Intrinsically stationary. (2 marks)

(b) (i) Can a random field be strictly stationary but not weakly stationary? If so, provide an example. If not, explain your reasoning. (2 marks)

(ii) Can a random field be intrinsically stationary but not weakly stationary? If so, provide an example. If not, explain your reasoning. (2 marks)

(iii) Can a random field be weakly stationary but not intrinsically stationary? If so, provide an example. If not, explain your reasoning. (2 marks)

(c) Consider the function

$$\rho(s,t) = \langle s, t \rangle$$

where $\langle s, t \rangle = \sum_{i=1}^{d} s_i t_i$ is the inner product of $s, t \in \mathbb{R}^d$.

(i) Show that this function is a valid covariance function for a random field $X = (X_t)_{t \in \mathbb{R}^d}$. (4 marks)

(ii) Is a zero-mean random field specified by this covariance function a weakly or intrinsically stationary random field? Justify your answer. (2 marks)

(iii) Suppose we simulate realisations from a zero-mean random field in one dimension ($d = 1$) specified by this covariance function, where we simulate in the range $-1 < t < 1$. Describe two features of these realisations which will not be common to all zero-mean random fields. (2 marks)

(Total: 20 marks)

2. (a) Let the random field $X = (X_t)_{t\in\mathbb{R}^d}$ be intrinsically stationary. Define the semi-variogram $\gamma : \mathbb{R}^d \to \mathbb{R}$ of X.

(2 marks)

(b) Show that the semi-variogram follows a property called *conditionally non-positive definite*, that is, $\forall a \in \mathbb{R}^n$ such that $\sum_{i=1}^n a_i = 0$, and for any finite set $t_1, \dots, t_n \in \mathbb{R}^d$, $n \in \mathbb{N}$,

$$\sum_{i=1}^n \sum_{j=1}^n a_i \gamma(t_i - t_j) a_j \le 0$$

(6 marks)

(c) Consider the *circular* semi-variogram, with $\alpha, \beta, R > 0$

$$\gamma(t) = \begin{cases} 0 & t = 0 \\ \alpha + \beta \left[1 - \frac{2}{\pi}\cos^{-1}\left(\frac{||t||}{R}\right) + \frac{2||t||}{\pi R}\sqrt{1 - \frac{||t||^2}{R^2}}\right] & 0 < ||t|| < R \qquad t \in \mathbb{R}^2 \\ \alpha + \beta & ||t|| \ge R \end{cases}$$

(i) Write down the values of the nugget, sill, partial sill, and range.

(4 marks)

(ii) Is the semi-variogram isotropic? Explain your reasoning.

(2 marks)

(d) Let $Z = [X_{t_1}, \dots, X_{t_n}]'$, where $t_i \in \mathbb{R}^d$, $i = 1, \dots, n$, be a length-n column vector of samples from some random field $(X_t)_{t\in\mathbb{R}^d}$. Suppose we specify some mean function $m : \mathbb{R}^d \to \mathbb{R}$ and covariance function, $\rho_0 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Define:

* Σ as the $n \times n$ covariance matrix with i, jth entry $\Sigma_{i,j} = \rho_0(t_i, t_j)$, which is assumed to be non-singular,
* K as the length-n column vector with entries $K_i = \rho_0(t_i, t_0)$,
* M as the length-n column vector with entries $M_i = m(t_i)$.

Then the simple Kriging predictor of $X_{t_0}, t_0 \in \mathbb{R}^d$ is

$$\hat{X}_{t_0} = m(t_0) + K'\Sigma^{-1}(Z - M)$$

and the mean squared prediction error is given by

$$\mathrm{err}(\hat{X}_{t_0}) = \rho_0(t_0, t_0) - K'\Sigma^{-1}K$$

(i) Under what assumptions would $\hat{X}_{t_0}$ be the best linear unbiased predictor (BLUP) of X_{t_0} with prediction error given by $\mathrm{err}(\hat{X}_{t_0})$?

(2 marks)

(ii) Suppose we replace the covariance function with $\rho_1 = 2\rho_0$ in the above simple Kriging procedure (but keep m the same), such that we use a covariance function that is double the original covariance function. Given the same values of Z, will our predictor and prediction error of X_{t_0} change? Justify your reasoning, and if so, explain by how much these will change with respect to the original prediction and prediction error.

(4 marks)

(Total: 20 marks)

3. (a) Let X be an L-valued random field on $T = \{1, \ldots, N\}, N \in \mathbb{N}$, with probability mass or density function given by $\pi_X(x), x \in L^T$, where L could be finite, countably infinite or $L \subseteq \mathbb{R}$. Define the local characteristics of the random field X on T. (2 marks)

(b) Assume that $\pi_X(x) > 0$ for all $x \in L^T$. Besag's factorisation theorem states that for all $x, y \in L^T$,

$$\frac{\pi_X(x)}{\pi_X(y)} = \prod_{i=1}^{N} \frac{\pi_i(x_i|x_1, \ldots, x_{i-1}, y_{i+1}, \ldots, y_N)}{\pi_i(y_i|x_1, \ldots, x_{i-1}, y_{i+1}, \ldots, y_N)}$$

Using Besag's factorisation Theorem, or otherwise, show that when $\pi_X(x) > 0$ the local characteristics define the whole probability distribution, that is, if Y is a random field having the same local characteristics as X, necessarily $\pi_Y \equiv \pi_X$. (3 marks)

(c) Consider a special case of the Conditional autoregressive (CAR) model X on $T = \{1, \ldots, N\}, N \in \mathbb{N}$, that is distributed as a multivariate normal with mean zero and positive-definite covariance matrix $(I - B)^{-1}$, where B is a symmetric $N \times N$ matrix whose diagonal elements b_{ii} are zero and I is the $N \times N$ identity matrix.

(i) Using 3(b) or otherwise, show that the local characteristics of the CAR model are normally distributed with

$$\begin{cases} \mathbb{E}(X_i|X_j, j \neq i) = \sum_{j \neq i} b_{ij} X_j \\ \mathrm{Var}(X_i|X_j, j \neq i) = 1 \end{cases}$$

for all $i \in T$. (5 marks)

(ii) Set $B = \phi A$ where $|\phi| < 1/2$ is a scalar and A is the neighbourhood or adjacency matrix, which defines a symmetric relation $\sim$ on T, given by

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Draw the undirected graph corresponding to the adjacency matrix A where each vertex in the graph is a region in T. Find the family of all cliques represented by this graph. (4 marks)

(iii) Recall that X is a Gibbs state with interaction potentials $V = \{V_A : A \subseteq T\}, V_A : L^T \to \mathbb{R}$, if

$$\pi_X(x) = \frac{1}{Z} \exp\left[\sum_{A \subseteq T} V_A(x_A)\right], \quad x \in L^T$$

Using the model in part 3(c)(ii), write down all interaction potentials V for this CAR model. Is X a Markov random field with respect to the symmetric relation $\sim$? Explain your reasoning briefly. (6 marks)

(Total: 20 marks)

4. (a) Define an inhomogeneous Poisson process X on $\mathbb{R}^d$ in terms of its intensity function $\lambda(x)$, and $N_X(A)$, which is a random variable corresponding to the number of points of X that fall in A for a bounded Borel set $A \subset \mathbb{R}^d$. (4 marks)

(b) For a point process X on $\mathbb{R}^d$, and for Borel sets $A, B \subseteq \mathbb{R}^d$, recall that the first-order factorial moment measure, the second-order moment measure, and the second-order factorial moment measure are respectively given by

$$\begin{aligned}\alpha^{(1)}(A) &= \mathbb{E}N_X(A) \\ \mu^{(2)}(A \times B) &= \mathbb{E}\left[N_X(A)N_X(B)\right] \\ \alpha^{(2)}(A \times B) &= \mathbb{E}\left[\sum_{x\in X}\sum_{y \in X}{}^{\neq} 1\{x \in A; y \in B\}\right]\end{aligned}$$

where the notation $\sum^{\neq}$ is used to indicate that the sum is taken over all $(x, y) \in X^2$ for which $x \neq y$.

(i) Write down the relationship between $\alpha^{(2)}(A \times B)$ and $\mu^{(2)}(A \times B)$ in terms of the first-order factorial moment measure $\alpha^{(1)}$. Explain which measure, $\alpha^{(2)}(A \times B)$ or $\mu^{(2)}(A \times B)$, is greater than or equal to the other, and when will these two measures be equal for all point process models? (3 marks)

(ii) Find $\alpha^{(1)}(A)$, $\mu^{(2)}(A \times B)$, and $\alpha^{(2)}(A \times B)$ for the inhomogeneous Poisson process of 4(a) for bounded Borel sets $A, B \subset \mathbb{R}^d$. (5 marks)

(c) Consider the Log-Gaussian Cox process X on $\mathbb{R}^d$ where a Gaussian random field $(Y_t)_{t\in\mathbb{R}^d}$ is first specified, and then conditional on Y_t, an inhomogeneous Poisson process is generated with intensity function $\exp(Y_t)$, where it is assumed that $\exp(Y_t)$ is almost surely integrable on bounded Borel sets. Show that the log-Gaussian Cox process is overdispersed, that is

$$\mathrm{Var}(N_X(A)) \geq \mathbb{E}N_X(A)$$

for all bounded Borel sets $A \subset \mathbb{R}^d$, where $N_X(A)$ is a random variable corresponding to the number of points of X that fall in A. Comment on how this result differs from the relationship between $\mathrm{Var}(N_X(A))$ and $\mathbb{E}N_X(A)$ for the inhomogeneous Poisson process of 4(a).

(8 marks)

(Total: 20 marks)

5. Consider a zero-mean space-time random process $Z(\mathbf{s}, t)$, where $\mathbf{s} \in \mathbb{R}^d (d \geq 1)$ denotes a spatial location and $t \in \mathbb{R}$ denotes a time point. Assume that the second moments of $Z(\mathbf{s}, t)$ exist and are finite. Denote N space-time coordinates as $(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_N, t_N) \in \mathbb{R}^d \times \mathbb{R}$, and the covariance function of $Z(\mathbf{s}, t)$ by $C(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = \operatorname{cov}\{Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)\}$, where $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$ are space-time coordinates.

(a) Explain the necessary and sufficient conditions for C to be a valid covariance function.

(3 marks)

(b) Define what it means for C to be a stationary covariance function in space and time.

(2 marks)

(c) Define what it means for C to be a separable covariance function in space and time.

(2 marks)

(d) Show that if a covariance function C is stationary and separable it can be expressed as

$$C(\mathbf{h}, u) = \frac{C(\mathbf{h}, 0)C(\mathbf{0}, u)}{C(\mathbf{0}, 0)}$$

for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$.

(4 marks)

(e) Consider the Cressie-Huang model for the space-time covariance given by

$$C(\mathbf{h}, u) = \int_{\mathbb{R}^d} \exp(i\langle \omega, \mathbf{h} \rangle) k(\omega) \rho(\omega, u) d\omega \tag{1}$$

where $\langle \omega, \mathbf{h} \rangle = \sum_{i=1}^d \omega_i h_i$ is the inner product on $\mathbb{R}^d$, $k(\cdot)$ is a spectral density on $\mathbb{R}^d$ and where, for each $\omega \in \mathbb{R}^d$, $\rho(\omega, \cdot)$ is an autocorrelation function on $\mathbb{R}$.

(i) Find the form of $C(\mathbf{h}, u)$ (up to a constant of proportionality) when

$$\rho(\omega, u) = \exp\left(-\frac{\|\omega\|^2 u^2}{4}\right) \exp\left(-\delta u^2\right), \quad \delta > 0$$
$$k(\omega) = \exp\left(-\frac{c_0 \|\omega\|^2}{4}\right), \quad c_0 > 0$$

(5 marks)

(ii) Is the covariance function $C(\mathbf{h}, u)$ separable with the form of ρ and k given in (e)(i)?

(2 marks)

(iii) In general, under what conditions on ρ will the covariance function $C(\mathbf{h}, u)$ defined in (1) be separable? Explain your reasoning briefly.

(2 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2024

This paper is also taken for the relevant examination for the Associateship.

MATH60139/MATH70139

Spatial Statistics (Solutions)

Setter's signature	Checker's signature	Editor's signature
. . . .	. . . .	. . . .

1. (a) (i) A random field $X = (X_t)_{t \in \mathbb{R}^d}$ is **strictly stationary** if for all finite sets $t_1, \ldots, t_n \in \mathbb{R}^d, n \in \mathbb{N}$, all $k_1, \ldots, k_n \in \mathbb{R}$, and all $s \in \mathbb{R}^d$, seen ⇓

$$\mathbb{P}(X_{t_1+s} \leq k_1; \ldots; X_{t_n+s} \leq k_n) = \mathbb{P}(X_{t_1} \leq k_1; \ldots; X_{t_n} \leq k_n).$$

2, A

(ii) A random field $X = (X_t)_{t \in \mathbb{R}^d}$ is **weakly stationary** if

- $\mathbb{E}X_t^2 < \infty$ for all $t \in \mathbb{R}^d$;
- $\mathbb{E}X_t \equiv m$ is constant;
- $\mathrm{Cov}(X_{t_1}, X_{t_2}) = \rho(t_2 - t_1)$ for some $\rho : \mathbb{R}^d \to \mathbb{R}$. 2, A

(iii) A random field $X = (X_t)_{t \in \mathbb{R}^d}$ is **intrinsically stationary** if

- $\mathbb{E}X_t^2 < \infty$ for all $t \in \mathbb{R}^d$;
- $\mathbb{E}X_t \equiv m$ is constant;
- $\mathrm{Var}(X_{t_2} - X_{t_1}) = f(t_2 - t_1)$ for some $f : \mathbb{R}^d \to \mathbb{R}$. 2, A

unseen ⇓

(b) (i) Yes, a random field can be strictly stationary but not weakly stationary. An example is a random field where X_t at each location t is a realisation of an independent t distribution with 1 or 2 degrees of freedom such that the second moment is not finite as required for weak stationarity, but the random field is clearly strictly stationary as per the definition given in (a)(i). 2, B

sim. seen ⇓

(ii) Yes, a random field can be intrinsically stationary but not weakly stationary. An example is a Gaussian random field with constant mean and covariance function $\rho(s,t) = \frac{1}{2}(|s| + |t| - |t - s|)$ (a Brownian surface). 2, B

seen ⇓

(iii) No, all weakly stationary random fields are intrinsically stationary. This is true because

$$\mathrm{Var}(X_{t_2} - X_{t_1}) = \mathrm{Var}(X_{t_2}) + \mathrm{Var}(X_{t_1}) - 2\mathrm{Cov}(X_{t_2}, X_{t_1}) = 2\rho(0) - 2\rho(t_2 - t_1),$$

such that $\mathrm{Var}(X_{t_2} - X_{t_1})$ is function of $t_2 - t_1$ only. 2, A

unseen ⇓

(c) (i) We require that for any finite set $t_1, \ldots, t_n \in \mathbb{R}^d$, $n \in \mathbb{N}$, the matrix $(\rho(t_i, t_j))_{i,j=1}^n$ is non-negative definite. Therefore we require that $(\rho(t_i, t_j))_{i,j=1}^n$ be symmetric and satisfy the property that

$$\sum_{i=1}^n \sum_{j=1}^n a_i \rho(t_i, t_j) a_j \geq 0,$$

for any $a \in \mathbb{R}^n$. To show this is true then we clearly have that $(\rho(t_i,t_j))_{i,j=1}^n$ is symmetric as $\langle t_i, t_j\rangle = \langle t_j, t_i\rangle$ and furthermore,

$$\sum_{i=1}^n\sum_{j=1}^n a_i\langle t_i,t_j\rangle a_j = \sum_{i=1}^n\sum_{j=1}^n a_i\left(\sum_{k=1}^d t_{i,k}t_{j,k}\right)a_j$$
$$= \sum_{k=1}^d\sum_{i=1}^n\sum_{j=1}^n a_i t_{i,k}t_{j,k}a_j = \sum_{k=1}^d\left\{\left(\sum_{i=1}^n a_i t_{i,k}\right)^2\right\} \geq 0,$$

for any $a \in \mathbb{R}^n$, where $t_{i,k}$ denotes the value of t_i in dimension k. Hence the covariance function is non-negative definite. This could also be answered in vector-matrix form by showing $a'\Sigma a \geq 0$ for all column vectors $a \in \mathbb{R}^n$, where $\Sigma = (\rho(t_i,t_j))_{i,j=1}^n$, using similar reasoning. $\boxed{4,\text{ D}}$

(ii) The random field is not weakly stationary as clearly the covariance function $\rho(s,t)$ cannot be written as a function of $t-s$ only (as required for weak stationarity), e.g., $\rho(\mathbf{0},\mathbf{0}) = 0$ and $\rho(\mathbf{1},\mathbf{1}) = d$. However, the random field is intrinsically stationary as $\mathrm{Var}(X_s - X_t) = \mathrm{Var}(X_s) + \mathrm{Var}(X_t) - 2\mathrm{Cov}(X_s,X_t) = \sum_{i=1}^d s_i^2 + \sum_{i=1}^d t_i^2 - 2\sum_{i=1}^d s_i t_i = \sum_{i=1}^d (s_i - t_i)^2$ which is a function of $s-t$ only (as required for intrinsic stationarity). $\boxed{2,\text{ A}}$

(ii) Any 2 of:

- The variance at $t = 0$ is zero.
- The variance grows with $|t|$.
- The covariance $\rho(s,t)$ is positive if $s > 0, t > 0$ or $s < 0, t < 0$.
- The covariance $\rho(s,t)$ is negative if $s > 0, t < 0$ or $s < 0, t > 0$,

or any other sensible observation that doesn't apply to all valid $\rho(s,t)$. $\boxed{1,\text{ A}}$

$\boxed{1,\text{ C}}$

2. (a) The semi-variogram is defined by $\boxed{\text{seen} \Downarrow}$

$$\gamma(t) = \frac{1}{2}\text{Var}(X_t - X_0), \quad t \in \mathbb{R}^d$$

$\boxed{2, \text{A}}$

(b) This follows because $\gamma(t_i - t_j) = \frac{1}{2}\text{Var}(X_{t_i - t_j} - X_0) = \frac{1}{2}\text{Var}(X_{t_i} - X_{t_j})$ such that $\boxed{\text{unseen} \Downarrow}$

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \gamma(t_i - t_j) = \\
&\quad = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Var}(t_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Var}(t_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(t_i, t_j) \\
&\quad = \frac{1}{2}\left(\sum_{j=1}^{n} a_j\right)\sum_{i=1}^{n} a_i \text{Var}(t_i) + \frac{1}{2}\left(\sum_{i=1}^{n} a_i\right)\sum_{j=1}^{n} a_j \text{Var}(t_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(t_i, t_j) \\
&\quad = -\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(t_i, t_j) \leq 0,
\end{aligned}$$

where we use that $\sum_{i=1}^{n} a_i = 0$ and the fact that the covariance function is non-negative definite. $\boxed{6, \text{D}}$

$\boxed{\text{sim. seen} \Downarrow}$

(c) (i) The nugget is

$$\lim_{||t|| \to 0} \gamma(t) = \alpha,$$

the sill is

$$\lim_{||t|| \to \infty} \gamma(t) = \alpha + \beta,$$

such that the partial sill is β. The range is $||t|| = R$. $\boxed{4, \text{A}}$

(ii) Yes, the semi-variogram is isotropic as it is a function of distance $||t||$ only. $\boxed{2, \text{A}}$

$\boxed{\text{seen} \Downarrow}$

(d) (i) The predictor is the best linear unbiased predictor (BLUP), with mean squared prediction error $\text{err}(\hat{X}_{t_0})$, if the sampled data was drawn from a random field $(X_t)_{t \in \mathbb{R}^d}$ with the prescribed mean function m and covariance function ρ_0. $\boxed{2, \text{B}}$

$\boxed{\text{unseen} \Downarrow}$

(ii) The prediction will not change as $m(t_0)$ and $(Z - M)$ are unchanged, and so will the product $K'\Sigma^{-1}$ as the new K vector will be double the previous but Σ^{-1} will be half the previous so these effects will cancel.
In contrast, the prediction error will increase by a factor of 2, as the first term $\rho(t_0, t_0)$ will double, as will the second term $K'\Sigma^{-1}K$. $\boxed{4, \text{C}}$

3. (a) The local characteristics of a random field X on T with values in L are

$$\pi_i(x_i|x_{T\setminus i}), \quad i \in T, x \in L^T,$$

whenever well-defined.

seen ⇓

2, A

(b) Choose any element $a \in L$. By the use of Besag's factorisation Theorem, $\pi_X(x)/\pi_X(a,\dots,a)$ is a product of local characteristics only, and the distribution $\pi_X(x)$ is obtained by normalisation.

3, B

sim. seen ⇓

(c) (i) Using Besag's factorisation Theorem, consider the right hand side with a specific value of i, and with $(y_i)_{i=1,\dots,N} = 0$, then by the conditional expectation and variance given in in the question we have that

$$\frac{\pi_i(x_i|x_1,\dots,x_{i-1},y_{i+1},\dots,y_N)}{\pi_i(y_i|x_1,\dots,x_{i-1},y_{i+1},\dots,y_N)} = \frac{\exp\left[-\frac{1}{2}\left(x_i - \sum_{j<i} b_{ij}x_j\right)^2\right]}{\exp\left[-\frac{1}{2}\left(\sum_{j<i} b_{ij}x_j\right)^2\right]}$$
$$= \exp\left[-\frac{1}{2}\left(x_i^2 - 2x_i\sum_{j<i} b_{ij}x_j\right)\right].$$

Then by Besag's factorisation Theorem,

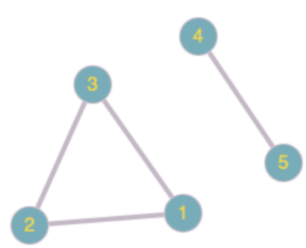
$$\frac{\pi_X(x)}{\pi_X(0)} = \exp\left[-\frac{1}{2}\sum_i x_i^2 + \sum_i\sum_{j<i} b_{ij}x_ix_j\right]$$
$$= \exp\left[-\frac{1}{2}x'x + \frac{1}{2}x'Bx\right] = \exp\left[-\frac{1}{2}x'(I-B)x\right],$$

which is proportional to the density of a multivariate normal with mean zero and covariance matrix $(I-B)^{-1}$, as given in the question.

5, B

unseen ⇓

(ii) The undirected graph looks like this:



The cliques are $\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{1,2\}, \{1,3\}, \{2,3\}, \{4,5\}, \{1,2,3\}$.

4, A

(iii) We have that

$$\pi_X(x) \propto \exp\left\{-\frac{1}{2}\sum_i\sum_j x_i(I-B)x_j\right\}.$$

Therefore:

$$V_{\{i\}}(x) = -x_i^2/2$$
$$V_{\{i,j\}}(x) = \phi A_{ij} x_i x_j$$

and $V_A(x) = 0$ for sets A of cardinality three or greater, and $V_\emptyset(\cdot) = 0$ by default. Note that $V_{\{i,j\}} = 0$ if $\{i,j\}$ is not in the family of cliques found in 3c(ii).

Finally, X is clearly a Markov random field with respect to $\sim$. This follows from the Hammersley-Clifford Theorem, which we can use here as $\pi_X(x) > 0$ for a multivariate normal distribution. The theorem states that, when $\pi_X(x) > 0$, X is a Markov random field with respect to $\sim$ where we can write $\pi_X(x)$ as a product of interaction functions over the cliques defined by $\sim$, as found already in 3(c)(ii).

Alternatively, we can say X is a Markov random field with respect to $\sim$ directly from the local characteristics given in 3(c)(i), as the local characteristics only depend on regions where i and j are neighbours ($i \sim j$ where $b_{ij} \neq 0$). $\boxed{1, \text{B}}$

$\boxed{5, \text{C}}$

4. (a) A point process X on $\mathbb{R}^d$ is a inhomogeneous Poisson process with intensity function $\lambda(x)$ if

seen ⇓

- $N_X(A)$ is Poisson distributed with mean $\mathbb{E}N_X(A) = \int_A \lambda(x)dx$ for some function $\lambda : \mathbb{R}^d \to \mathbb{R}^+$ that is integrable on all bounded Borel sets $A \subset \mathbb{R}^d$.
- for any k disjoint bounded Borel sets $A_1, \ldots, A_k, k \in \mathbb{N}$, the random variables $N_X(A_1), \ldots, N_X(A_k)$ are independent.

4, A

(b) (i) The relationship is given by

$$\mu^{(2)}(A \times B) = \alpha^{(2)}(A \times B) + \alpha^{(1)}(A \cap B),$$

such that $\mu^{(2)}(A \times B)$ is greater than or equal to $\alpha^{(2)}(A \times B)$ as all measures are positive. The measures become equal if the Borel sets A and B are disjoint.

3, A

(ii) By definition, $N_X(A)$ is Poisson distributed with expectation $\Lambda(A) = \int_A \lambda(x)dx < \infty$ for every bounded Borel set A. Hence, $\alpha^{(1)}(A) = \Lambda(A)$.
To compute the second order moment measure, we use the fact that the counts in disjoint sets are independent, such that for bounded Borel sets $A, B \subset \mathbb{R}^d$,

$$\begin{aligned}
\mu^{(2)}(A \times B) &= \mathbb{E}[N_X(A)\{N_X(A \cap B) + N_X(B \backslash A)\}] \\
&= \mathbb{E}[\{N_X(A \cap B) + N_X(A \backslash B)\}N_X(A \cap B)] + \Lambda(A)\Lambda(B \backslash A) \\
&= \Lambda(A \cap B) + \Lambda(A \cap B)^2 + \Lambda(A \backslash B)\Lambda(A \cap B) + \Lambda(A)\Lambda(B \backslash A) \\
&= \Lambda(A \cap B) + \Lambda(A \cap B)[\Lambda(A \cap B) + \Lambda(A \backslash B)] + \Lambda(A)\Lambda(B \backslash A) \\
&= \int_A \int_B \lambda(x)\lambda(y)dxdy + \int_{A \cap B} \lambda(x)dx.
\end{aligned}$$

Since

$$\mu^{(2)}(A \times B) = \alpha^{(2)}(A \times B) + \alpha^{(1)}(A \cap B)$$

the second order factorial moment measure is

$$\alpha^{(2)}(A \times B) = \int_A \int_B \lambda(x)\lambda(y)dxdy.$$

3, B

2, C

unseen ⇓

(c) For the log-Gaussian Cox process, the key step is realising that the process is a *conditional* inhomogeneous Poisson process, where the intensity function is a random variable itself. Denote, as in 4(b)(ii), the intensity for the inhomogeneous Poisson process in bounded Borel set A as $\Lambda(A) = \int_A \exp(Y_t)dt$. Then,

$$\begin{aligned}
\mathbb{E}(N_X(A)^2) &= \mathbb{E}\left[\mathbb{E}(N_X(A)^2|\Lambda(A))\right] \\
&= \mathbb{E}\left[\mathrm{Var}(Poisson(\Lambda(A))) + \{\mathbb{E}(Poisson(\Lambda(A)))\}^2\right] \\
&= \mathbb{E}\left[\Lambda(A) + \Lambda^2(A)\right]
\end{aligned}$$

and

$$\mathbb{E}N_X(A) = \mathbb{E}\left[\mathbb{E}(N_X(A)|\Lambda(A))\right] = \mathbb{E}[\Lambda(A)]$$

such that

$$\begin{aligned}\mathrm{Var}(N_X(A)) &= \mathbb{E}(N_X(A)^2) - \{\mathbb{E}(N_X(A))\}^2 \\ &= \mathbb{E}\left[\Lambda(A) + \Lambda^2(A)\right] - \{\mathbb{E}[\Lambda(A)]\}^2 \\ &= \mathbb{E}[\Lambda(A)] + \mathrm{Var}[\Lambda(A)] = \mathbb{E}(N_X(A)) + \mathrm{Var}[\Lambda(A)]\end{aligned}$$

such that $\mathrm{Var}(N_X(A)) \geq \mathbb{E}N_X(A)$ as required. $\boxed{6, \mathrm{D}}$

For the inhomogeneous Poisson process, $N_X(A)$ is Poisson distributed with mean $\mathbb{E}N_X(A) = \int_A \lambda(x)dx$, such that its variance is also $\mathrm{Var}(N_X(A)) = \int_A \lambda(x)dx$ from the properties of the Poisson distribution, such that $\mathbb{E}N_X(A) = \mathrm{Var}(N_X(A))$.

$\boxed{2, \mathrm{B}}$

5. (a) A necessary and sufficient condition for a function C to be a valid covariance function is that it is symmetric ($C(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) = C(\mathbf{s}_j, t_j, \mathbf{s}_i, t_i)$) and a non-negative definite function such that it satisfies

$$\sum_{i,j=1}^{N} c_i c_j C(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) \geq 0$$

for all finite N, all $(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_N, t_N) \in \mathbb{R}^d \times \mathbb{R}$, and all real $c_1, \ldots, c_N$.

seen ⇓

3, M

(b) A covariance function C is stationary in space and time if there exists a non-negative definite covariance function C defined on $\mathbb{R}^d \times \mathbb{R}$ such that $\mathrm{cov}\{Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)\} = C(\mathbf{h}, u)$ for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$, where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ and $u = t_1 - t_2$.

2, M

(c) A covariance function C is separable in space and time if it can be expressed as

$$C(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = C_{\mathrm{S}}(\mathbf{s}_1, \mathbf{s}_2) C_{\mathrm{T}}(t_1, t_2)$$

for all $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$, where C_{S} and C_{T} are purely spatial and purely temporal covariance functions, respectively.

2, M

unseen ⇓

(d) As the covariance is stationary and separable we have that

$$\begin{aligned} C(\mathbf{h}, u) = C_{\mathrm{S}}(\mathbf{h}) C_{\mathrm{T}}(u) &= \frac{C_{\mathrm{S}}(\mathbf{h}) C_{\mathrm{T}}(u) C(\mathbf{0}, 0)}{C(\mathbf{0}, 0)} \\ &= \frac{C_{\mathrm{S}}(\mathbf{h}) C_{\mathrm{T}}(0) C_{\mathrm{S}}(\mathbf{0}) C_{\mathrm{T}}(u)}{C(\mathbf{0}, 0)} = \frac{C(\mathbf{h}, 0) C(\mathbf{0}, u)}{C(\mathbf{0}, 0)} \end{aligned}$$

4, M

(e) (i) We have that

$$C(\mathbf{h}, u) = \exp\left(-\delta u^2\right) \int_{\mathbb{R}^d} \exp(i\langle\omega, \mathbf{h}\rangle) \exp(-\beta||\omega||^2) d\omega.$$

where $\beta = \frac{u^2 + c_0}{4}$. This integral factorises into d one-dimensional terms of the form

$$\begin{aligned} \int_{-\infty}^{\infty} \exp(i\omega h) \exp(-\beta\omega^2) d\omega &= \exp(-h^2/(4\beta)) \int_{-\infty}^{\infty} \exp(-\beta[\omega - ih/(2\beta)]^2) d\omega \\ &= \exp(-h^2/(4\beta)) \sqrt{\frac{\pi}{\beta}} \end{aligned}$$

for $h \in \mathbb{R}$. Collecting the d terms we find that

$$\begin{aligned} C(\mathbf{h}, u) &\propto \frac{1}{\beta^{d/2}} \exp\left(-\frac{\|h\|^2}{4\beta}\right) \exp\left(-\delta u^2\right) \\ &\propto \frac{1}{(u^2 + c_0)^{d/2}} \exp\left(-\frac{\|h\|^2}{u^2 + c_0}\right) \exp\left(-\delta u^2\right) \end{aligned}$$

5, M

(ii) This is not a separable covariance function as the $\mathbf{h}$ and u terms are combined in the exponential and clearly cannot be separated into space and time components for all $\mathbf{h}$ and u. 2, M

seen ⇓

(iii) The covariance function will clearly be separable if $\rho(\omega, u)$ is independent of ω. This is because k is a spectral density and by Bochner's Theorem therefore defines a valid covariance function in space only, and $\rho(\omega, u)$ is then by definition a valid covariance function in time, and both these covariances are separable by the model given in the question. 2, M

Review of mark distribution:
Total A marks: 32 of 32 marks
Total B marks: 20 of 20 marks
Total C marks: 12 of 12 marks
Total D marks: 16 of 16 marks
Total marks: 100 of 80 marks
Total Mastery marks: 20 of 20 marks

MATH60139 Spatial Statistics

Question	Marker's comment
1	Q1 answered very well overall, well done! I was pleased with your knowledge of key definitions and understanding of how they differ. Not all students managed to correctly answer 1(c)(i) in terms of proving non-negative definiteness for the inner product covariance function, please review the solutions for the key step.
2	Q2 also answered well with good understanding of semi-variograms and Kriging demonstrated, but students struggled most with 2(b) where the key step was to expand the semi-variogram as a combination of variances and covariances in t_i and t_j, and then use that the a_i sum to 0, please review the solutions!
3	Q3 also answered very well on what is perhaps the most abstract chapter - well done! Finding the interaction potentials (3(c)(iii)) was perhaps where most mistakes occurred.
4	Q4 also good attempts, with very good understanding of the inhomogeneous Poisson process demonstrated, but the final question (4(c)) challenged the most. There were several ways of answering this question, including the one given in the solutions, but perhaps the most elegant were those who used the law of total variance - well done if you did! Many of you tried to prove the statement using the moment generating function expansion - this was a possible route but ultimately unnecessary - where unfortunately a few got in a muddle in terms of combining the complex expansions that then occurred (marks for effort were still given if you got close!)

MATH70139 Spatial Statistics

Question	Marker's comment
1	Q1 answered very well overall, well done! I was pleased with your knowledge of key definitions and understanding of how they differ. Not all students managed to correctly answer 1(c)(i) in terms of proving non-negative definiteness for the inner product covariance function, please review the solutions for the key step.
2	Q2 also answered well with good understanding of semi-variograms and Kriging demonstrated, but students struggled most with 2(b) where the key step was to expand the semi-variogram as a combination of variances and covariances in t_i and t_j, and then use that the a_i sum to 0, please review the solutions!
3	Q3 also answered very well on what is perhaps the most abstract chapter - well done! Finding the interaction potentials (3(c)(iii)) was perhaps where most mistakes occurred.
4	Q4 also good attempts, with very good understanding of the inhomogeneous Poisson process demonstrated, but the final question (4(c)) challenged the most. There were several ways of answering this question, including the one given in the solutions, but perhaps the most elegant were those who used the law of total variance - well done if you did! Many of you tried to prove the statement using the moment generating function expansion - this was a possible route but ultimately unnecessary - where unfortunately a few got in a muddle in terms of combining the complex expansions that then occurred (marks for effort were still given if you got close!)
5	Q5 was answered well in terms of key definitions and understanding of space-time covariances - well done! 5(e)(i) challenged the most in terms of the calculus. A simpler example was given in lectures back in Chapter 2, but this was quite a tough one but I was impressed that a handful did get to the final solution, even finding the constant of proportionality! Working marks were given for those that got close.