IMPERIAL

MATH70139

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
Summer 2025

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

Spatial Statistics

Date: Thursday, May 22, 2025

Time: Start time 14:00 – End time 16:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer All Questions in 1 Answer Booklet

This is a closed book examination.

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.

Credit will be given for all questions attempted.

Allow margins for marking.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO DO SO

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) intrinsically stationary. (2 marks)

   (iii) a random field with isotropic covariance. (2 marks)

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

   (ii) Can a zero-mean random field have isotropic covariance but not be intrinsically stationary? If so, provide an example. If not, explain your reasoning. (2 marks)

(c) Consider the function

$$\rho(s,t) = \sum_{k=1}^{p} H(s)_k H(t)_k, \quad s,t \in \mathbb{R}^d$$

where $H : \mathbb{R}^d \to \mathbb{R}^p, p \in \mathbb{N}$ is a function, and $H(s)_k$ denotes the kth element of $H(s)$.

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

   (ii) Suppose we set the function H to be $H(s) = (||s||, 0.5)$, such that $p = 2$, where $||s|| = \sqrt{s_1^2 + \ldots + s_d^2}$ is the Euclidean norm of $s \in \mathbb{R}^d$. Write down the form of $\rho(s,t)$. Is a zero-mean random field specified by $\rho(s,t)$ a weakly or intrinsically stationary random field? Justify your answer. (4 marks)

   (iii) Suppose we simulate realisations from a zero-mean random field in one dimension ($d = 1$) specified by the covariance function in part (c)(ii), 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 $\gamma_X : \mathbb{R}^d \to \mathbb{R}$ of X.

(2 marks)

(b) Assume spatial observations are obtained from a random field $Y = (Y_t)_{t \in \mathbb{R}^d}$ which is related to X, at all locations $t_i \in \mathbb{R}^d$, by a linear model given by

$$Y_{t_i} = X_{t_i} + E_{t_i},$$

where E_{t_i} is a zero-mean i.i.d. (independent and identically distributed) noise term with variance $\sigma_E^2 > 0$ that is independent of the intrinsically stationary random field X.

(i) Find the form of the semi-variogram of Y in terms of the semi-variogram of X.

(2 marks)

(ii) Find the form of the covariance function of Y in terms of the covariance function of X.

(3 marks)

(iii) Describe the key special property of the semi-variogram and covariance function of Y, what this phenomenon is called, and what it represents in practice (one sentence for each).

(3 marks)

(iv) When will the random field Y be intrinsically or weakly stationary? Explain your reasoning in terms of the properties of X.

(4 marks)

(v) In simple Kriging, the best linear unbiased predictor of Y_{t_0} given n observations $Z = [Y_{t_1}, \ldots, Y_{t_n}]'$, where t_0 is not equal to any of $t_1, \ldots, t_n$, is given by

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

and the mean squared prediction error is given by

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

where $m_Y : \mathbb{R}^d \to \mathbb{R}$ and $\rho_Y : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ are respectively the mean and covariance functions of Y, and

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

Show that the mean squared prediction error, $\mathrm{err}(\hat{Y}_{t_0})$, satisfies the following inequality

$$\mathrm{err}(\hat{Y}_{t_0}) \geq \mathrm{err}(\hat{X}_{t_0}) + \sigma_E^2$$

where $\mathrm{err}(\hat{X}_{t_0})$ is the mean squared prediction error of predicting X_{t_0} from the noise-free observations $X_{t_1}, \ldots, X_{t_n}$ by simple Kriging. Comment on situations where this inequality approaches an equality or not. You may use without proof the Woodbury matrix identity

$$(A + B)^{-1} = A^{-1} - A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}$$

or any other well-known stated linear algebra result.

(6 marks)

(Total: 20 marks)

3. (a) Define a homogeneous Poisson process X on $\mathbb{R}^d$ in terms of its intensity λ, 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 $A \subset \mathbb{R}^d$. Is the process stationary in $\mathbb{R}^d$? (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

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

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 $\alpha^{(2)}(A \times A)$ and $\mu^{(2)}(A \times A)$, for a (single) Borel set $A \subseteq \mathbb{R}^d$, in terms of the first-order factorial moment measure $\alpha^{(1)}(A)$. Explain which measure, $\alpha^{(2)}(A \times A)$ or $\mu^{(2)}(A \times A)$, is greater than or equal to the other, and under what circumstances will these two measures be equal. (4 marks)

(ii) Find $\alpha^{(2)}(A \times A)$ and $\mu^{(2)}(A \times A)$ for the homogeneous Poisson process of 3(a) for a bounded Borel set $A \subset \mathbb{R}^d$. (4 marks)

(c) Let X be a finite point process on a bounded Borel set $W \subset \mathbb{R}^d$ whose distribution is defined by a probability density f with respect to the distribution of a unit rate Poisson process and let $\mathbf{x} \subset W$ be a finite point pattern. Then the *Papangelou conditional intensity* at $y \in W$ given $\mathbf{x}$ is defined as

$$\lambda(y \mid \mathbf{x}) = \frac{f(\mathbf{x} \cup \{y\})}{f(\mathbf{x})}$$

for $y \notin \mathbf{x}$ provided $f(\mathbf{x}) \neq 0$. Set $\lambda(y \mid \mathbf{x}) = 0$ otherwise. Also, recall that a Strauss process is defined by

$$f(\mathbf{x}) \propto \prod_{x\in\mathbf{x}} \beta(x) \prod_{\{u,v\}\subseteq\mathbf{x}} \gamma(u, v), \quad \mathbf{x} \subset W$$

where $\beta : W \to \mathbb{R}^+$ is some measurable function and

$$\gamma(u, v) = \begin{cases} \gamma_0 & \text{if } \|u - v\| \leq R \\ 1 & \text{if } \|u - v\| > R \end{cases} \quad R \in \mathbb{R}$$

Write down the Papangelou conditional intensity, $\lambda(y \mid \mathbf{x})$, of the Strauss process if:

(i) $\gamma_0 = 1$ (2 marks)

(ii) $0 < \gamma_0 < 1$ (3 marks)

(iii) $\gamma_0 = 0$ (3 marks)

In each case describe (in one sentence for each) the behaviour of the process.

(Total: 20 marks)

4. (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) Suppose X is an Ising model defined by a probability mass function given by

$$\pi_X(x) = \frac{1}{Z(\alpha, \beta)} \exp\left[\alpha \sum_{i \in T} x_i + \beta \sum_{\{i,j\}: i \sim j} x_i x_j\right], \quad x \in L^T$$

for constants $\alpha, \beta \in \mathbb{R}$, where $L = \{0, 1\}$, $Z(\alpha, \beta)$ is some normalising constant dependent on α and β, and $\sim$ denotes a symmetric relation.

(i) Find the local characteristics of the Ising model. (4 marks)

(ii) Consider the neighbourhood or adjacency matrix A, which defines the symmetric relation $\sim$ on T, where $N = |T| = 5$, given by

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

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

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

where Z is a normalising constant. Write down all interaction potentials V for the Ising model using the symmetric relation defined in 3(c)(ii). Is X a Markov random field with respect to the symmetric relation $\sim$? Explain your reasoning briefly. (5 marks)

(iv) Suppose we observe a realisation $\{x_1, \ldots, x_N\}$ from the Ising model and we wish to estimate the parameters of α and β. Maximum likelihood requires computation of $Z(\alpha, \beta)$ which becomes prohibitively expensive to compute when N is large. Instead consider using the log pseudo-likelihood function given by

$$PL(\theta; x) = \sum_{i=1}^{N} \log \pi_i(x_i | x_{T \setminus i}).$$

Show that the log pseudo-likelihood for the Ising model ($\theta = (\alpha, \beta)$) is given by

$$PL(\theta; x) = \sum_{i=1}^{N} x_i \left(\alpha + \beta \sum_{\{j\}: i \sim j} x_j\right) - \sum_{i=1}^{N} \log\left[1 + \exp\left(\alpha + \beta \sum_{\{j\}: i \sim j} x_j\right)\right]$$

which does not depend on $Z(\alpha, \beta)$ and can thus be optimised numerically over (α, β). (5 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) Define what it means for C to be fully symmetric.

(2 marks)

(e) Consider the following space-time covariance function:

$$C(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = \sigma^2 \exp\{-(||\mathbf{s}_2 - \mathbf{s}_1||/a + |t_2 - t_1|/b)\}$$

with $a, b, \sigma^2 > 0$. Is this covariance function separable and/or fully symmetric? Explain your reasoning. Interpret the parameters a and b in terms of their effect on the covariance (in max 2 sentences). (4 marks)

(f) Prove that if C is a continuous function on $\mathbb{R}^d \times \mathbb{R}$, then C is a stationary, fully symmetric space-time covariance function if and only if it is of the form

$$C(\boldsymbol{h}, u) = \iint \cos(\boldsymbol{h}'\boldsymbol{\omega}) \cos(u\tau) \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$$

where F is a finite, non-negative measure on $\mathbb{R}^d \times \mathbb{R}$. You may use without proof Bochner's Theorem below.

Bochner's Theorem: Suppose that C is a continuous and symmetric function on $\mathbb{R}^d \times \mathbb{R}$. Then C is a covariance function if and only if it is of the form

$$C(\boldsymbol{h}, u) = \iint e^{i(\boldsymbol{h}'\boldsymbol{\omega} + u\tau)} \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$$

where $i = \sqrt{-1}$, and F is a finite, non-negative and symmetric measure on $\mathbb{R}^d \times \mathbb{R}$.

(7 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2025

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 $s \in \mathbb{R}^d$, $\boxed{\text{seen} \Downarrow}$

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

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

(ii) 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 $f : \mathbb{R}^d \to \mathbb{R}$. $\boxed{2, \text{A}}$

(iii) A random field $X = (X_t)_{t\in\mathbb{R}^d}$ has an **isotropic covariance** if for each $t_1, t_2 \in \mathbb{R}^d$, $\mathrm{Cov}(X_{t_1}, X_{t_2})$ depends only on $||t_2 - t_1||$; specifically,

$$\exists \rho_0 : \mathbb{R}^+ \to \mathbb{R}\ s.t. : \forall t_1, t_2 \in \mathbb{R}^d,\ \rho(t_1, t_2) = \rho_0(||t_2 - t_1||) = \rho(t_2 - t_1).$$

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

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

(b) (i) Yes, a random field can be strictly stationary but not intrinsically 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 intrinsically stationarity, but the random field is clearly strictly stationary as per the definition given in (a)(i). $\boxed{2, \text{A}}$

(ii) No, all zero-mean random fields with isotropic covariance are intrinsically stationary. This is true because all isotropic covariances are stationary covariances (as can be seen from a(iii)), such that the random field is weakly stationary, and all weakly stationary random fields are intrinsically stationary.

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

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

(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 positive semi-definite. Therefore we require that $(\rho(t_i, t_j))_{i,j=1}^n$ be symmetric and satisfy the property that

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

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 $\sum_{k=1}^p H(s)_k H(t)_k = \sum_{k=1}^p H(t)_k H(s)_k$ and furthermore,

$$\sum_{i=1}^n \sum_{j=1}^n a_i \left(\sum_{k=1}^p H(t_i)_k H(t_j)_k \right) a_j = \sum_{k=1}^p \sum_{i=1}^n \sum_{j=1}^n a_i H(t_i)_k H(t_j)_k a_j$$
$$= \sum_{k=1}^p \left\{ \left(\sum_{i=1}^n a_i H(t_i)_k \right)^2 \right\} \ge 0,$$

for any $a \in \mathbb{R}^n$. Hence the covariance function is positive semi-definite. $\boxed{4, \text{D}}$

(ii) The covariance function is $\rho(s,t) = ||s||.||t|| + 0.25$. 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., $\rho(\mathbf{0},\mathbf{0}) = 0.25$ and $\rho(\mathbf{1},\mathbf{1}) = d + 0.25$.
Similarly, the random field is not intrinsically stationary as

$$\begin{aligned}\mathrm{Var}(X_t - X_s) &= \mathrm{Var}(X_s) + \mathrm{Var}(X_t) - 2\mathrm{Cov}(X_s, X_t)\\ &= ||s||^2 + 0.25 + ||t||^2 + 0.25 - 2(||s||.||t|| + 0.25)\\ &= ||s||^2 + ||t||^2 - 2||s||.||t|| = (||t|| - ||s||)^2\end{aligned}$$

cannot be written as a function of $t - s$ only, e.g., $\mathrm{Var}(X_{\mathbf{1}} - X_{-\mathbf{1}}) = 0$ but $\mathrm{Var}(X_{\mathbf{2}} - X_{\mathbf{0}}) = 4d$. $\boxed{4, \mathrm{B}}$

(iii) Any 2 of:

- The variance at $t = 0$ is 0.25.
- The variance grows with $|t|$.
- The covariance $\rho(s,t)$ is always positive.
- The correlation is 1 between t and $-t$ such that the random field is symmetric,

or any other sensible observation that doesn't apply to all valid $\rho(s,t)$. $\boxed{2, \mathrm{A}}$

2. (a) The semi-variogram is defined by

seen ⇓

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

2, A

(b) (i)

$$\begin{aligned}\frac{1}{2}\mathrm{Var}(Y_t - Y_0) &= \frac{1}{2}\mathrm{Var}(X_t - X_0) + \frac{1}{2}\mathrm{Var}(E_t - E_0)\\ &= \frac{1}{2}\mathrm{Var}(X_t - X_0) + \sigma_E^2 - \mathrm{Cov}(E_t, E_0) = \frac{1}{2}\mathrm{Var}(X_t - X_0) + \sigma_E^2 1\{t \neq 0\},\end{aligned}$$

such that

$$\gamma_Y(t) = \begin{cases}\gamma_X(t) + \sigma_E^2 & t \neq 0\\ \gamma_X(t) & t = 0\end{cases}$$

2, A

(ii)

unseen ⇓

$$\begin{aligned}\mathrm{Cov}(Y_s, Y_t) &= \mathrm{Cov}(X_s + E_s, X_t + E_t) = \mathrm{Cov}(X_s, X_t) + \mathrm{Cov}(E_s, E_t)\\ &= \rho_X(s,t) + \sigma_E^2 1\{s = t\},\end{aligned}$$

such that

$$\rho_Y(s,t) = \begin{cases}\rho_X(s,t) + \sigma_E^2 & s = t\\ \rho_X(s,t) & s \neq t\end{cases}$$

3, A

(iii) The semi-variogram and covariance functions exhibit discontinuities at $t = 0$ and $s = t$ respectively. This phenomenon is known as the nugget effect and occurs when there is observation noise or variability in values of an underlying random field collected at the same location.

seen ⇓

1, A

(iv) The random field Y is intrinsically stationary because the random field X is intrinsically stationary. We can show this by observing the three properties of intrinsic stationarity are still satisfied: the mean function is constant (it is the same with X and Y), the second moment is finite (it's just the second moment of X with σ_E^2 added), and $\mathrm{Var}(Y_{t+s} - Y_t)$ is a function of s only, as can be observed from (b)(i).
On the other hand Y is weakly stationary if and only if X is weakly stationary. This is because we require $\rho_Y(s,t)$ to be a function of $t - s$ only, which from b(ii) will only occur if $\rho_X(s,t)$ is a function of $t - s$ only.

2, B

unseen ⇓

4, C

(v) We have that

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

where from b(ii) we have that K is the same for X and Y and $\Sigma = \Sigma_X + \sigma_E^2 I$ (where I is the identity matrix) such that

$$\mathrm{err}(\hat{Y}_{t_0}) = \rho_X(t_0, t_0) + \sigma_E^2 - K'(\Sigma_X + \sigma_E^2 I)^{-1}K$$

such that

$$\mathrm{err}(\hat{Y}_{t_0}) - \mathrm{err}(\hat{X}_{t_0}) = \sigma_E^2 + (K'\Sigma_X^{-1}K - K'(\Sigma_X + \sigma_E^2 I)^{-1}K).$$

We therefore need to show that $K'\Sigma_X^{-1}K \geq K'(\Sigma_X + \sigma^2 I)^{-1}K$ and we are done. Using the Woodbury matrix identity with $A = \Sigma_X$ and $B = \sigma_E^2 I$ the right hand side is

$$K'\left[\Sigma_X^{-1} - \Sigma_X^{-1}(\Sigma_X^{-1} + I/\sigma_E^2)^{-1}\Sigma_X^{-1}\right]K = K'\Sigma_X^{-1}K - K'\Sigma_X^{-1}(\Sigma_X^{-1} + I/\sigma_E^2)^{-1}\Sigma_X^{-1}K.$$

Now, as Σ_X is positive definite then so is Σ_X^{-1} and so is $\Sigma_X^{-1} + I/\sigma_E^2$ (as the eigenvalues have just been made more positive), and finally $(\Sigma_X^{-1} + I/\sigma_E^2)^{-1}$ is positive definite, such that $K'\Sigma_X^{-1}(\Sigma_X^{-1} + I/\sigma_E^2)^{-1}\Sigma_X^{-1}K \geq 0$ (where the inequality is non-strict in case K is a zero vector) thus completing the proof.

The errors therefore increase by at least as much as the nugget (noise) variance as expected, but will increase even more where there are lots of nearby observations contaminated by noise (where the covariances in the K vector are large). In contrast, if the K vector approaches or becomes zero, which occurs far from observation locations or if the covariance decays quickly in space, then the inequality approaches an equality. $\boxed{6, \text{D}}$

3. (a) A point process X on $\mathbb{R}^d$ is a homogeneous Poisson process with intensity $\lambda > 0$ if seen $\Downarrow$

* $N_X(A)$ is Poisson distributed with mean $\mathbb{E}N_X(A) = \lambda|A|$ for every bounded Borel set $A \subset \mathbb{R}^d$.
* 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.

The homogeneous Poisson process is stationary in $\mathbb{R}^d$ as the distribution of $N_X(A)$ is translation invariant and determined by the volume of A only. 4, A

(b) (i) The relationship is given by sim. seen $\Downarrow$

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

such that $\mu^{(2)}(A \times A)$ is greater than or equal to $\alpha^{(2)}(A \times A)$ as all measures are positive. The measures become equal only in the degenerate case where $\alpha^{(1)}(A) = \mathbb{E}N_X(A) = 0$ such that there are always no points in A and by extension $\mu^{(2)}(A \times A) = \alpha^{(2)}(A \times A) = 0$. 4, A

unseen $\Downarrow$

(ii) By definition, $N_X(A)$ is Poisson distributed with expectation $\lambda|A|$ for every bounded Borel set A. Hence, $\alpha^{(1)}(A) = \lambda|A|$.
To compute $\mu^{(2)}(A \times A)$, we directly compute

$$\begin{aligned}\mu^{(2)}(A \times A) &= \mathbb{E}[N_X(A)^2] \\ &= \mathrm{Var}[N_X(A)] + \{\mathbb{E}[N_X(A)]\}^2 \\ &= \lambda|A| + \lambda^2|A|^2\end{aligned}$$

and hence

$$\alpha^{(2)}(A \times A) = \lambda|A| + \lambda^2|A|^2 - \lambda|A| = \lambda^2|A|^2.$$

4, B

(c) (i) $\lambda(y \mid \mathbf{x}) = \beta(y)$. This is in fact an inhomogeneous Poisson process with intensity $\beta(y)$. 2, C

(ii) $\lambda(y \mid \mathbf{x}) = \beta(y)\gamma_0^{S_R(y|\mathbf{x})}$, $y \notin \mathbf{x}$, where $S_R(y \mid \{x_1, \ldots, x_n\}) = \sum_{i=1}^n 1\{||x_i - y\| \leq R\}$ and thus depends only on points in $\mathbf{x}$ that are at most a distance R removed from y. This process makes points within distance R of $\mathbf{x}$ less likely to occur than under an inhomogeneous Poisson with the same $\beta(y)$, creating a repelling effect. 3, D

(iii) If there is a pair of points in $\mathbf{x}$ that are within distance R then $f(\mathbf{x}) = 0$ and $\lambda(y \mid \mathbf{x}) = 0$ by definition. Otherwise, if y is not within distance R of any points in $\mathbf{x}$ then $\lambda(y \mid \mathbf{x}) = \beta(y)$, otherwise $\lambda(y \mid \mathbf{x}) = 0$. This is a hardcore process where no points are allowed within distance R. 3, D

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

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

whenever well-defined. 2, A

(b) (i) First we note that

$$\pi_i(x_i|x_{T\setminus i}) \propto \exp\left[\alpha x_i + \beta \sum_{\{j\}:i\sim j} x_i x_j\right], \quad x_i \in \{0,1\},$$

such that

$$\pi_i(0|x_{T\setminus i}) \propto 1, \quad \pi_i(1|x_{T\setminus i}) \propto \exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right],$$

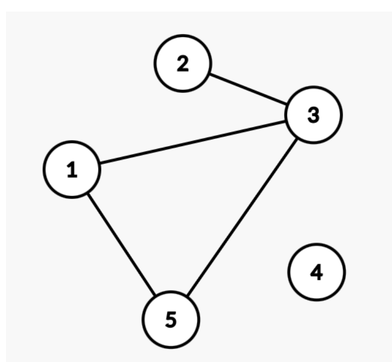
with the same constant of proportionality in both cases. Finally we can normalise to provide exact probabilities

$$\pi_i(0|x_{T\setminus i}) = \frac{1}{1+\exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]},$$
$$\pi_i(1|x_{T\setminus i}) = \frac{\exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]}{1+\exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]}.$$

3, B

1, C

(ii) The undirected graph looks like this: sim. seen ⇓



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

4, A

(iii) The Ising model is a Gibbs state with interaction potentials seen ⇓

$$V_{\{i\}}(x) = \alpha x_i$$
$$V_{\{i,j\}}(x) = \begin{cases} \beta x_i x_j & \text{if } i \sim j \\ 0 & \text{otherwise,} \end{cases}$$

and $V_A(x) = 0$ for sets A of cardinality three or greater. 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 $\pi_X(x) > 0$.

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 ($i \sim j$). $\boxed{\text{5, B}}$

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

(iv) We can observe from c(i) that as x_i takes values 0 or 1, then we can express $PL(\theta; x)$ as follows, which then simplifies to the form given in the question:

$$\begin{aligned}
PL(\theta; x) &= \sum_{i=1}^{N} \log \pi_i(x_i | x_{T\setminus i}) \\
&= \sum_{i=1}^{N} \log \left\{ \frac{\left(\exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]\right)^{x_i}}{1 + \exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]} \right\} \\
&= \sum_{i=1}^{N} \log \left\{ \left(\exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right]\right)^{x_i} \right\} - \sum_{i=1}^{N} \log \left\{ 1 + \exp\left[\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right] \right\} \\
&= \sum_{i=1}^{N} x_i \left(\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right) - \sum_{i=1}^{N} \log\left[1 + \exp\left(\alpha + \beta \sum_{\{j\}:i\sim j} x_j\right)\right].
\end{aligned}$$

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

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 positive semi-definite function such that it satisfies

seen ⇓

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

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 $c_1, \ldots, c_N$. 3, M

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

(d) A covariance function C is fully symmetric if

$$C(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = C(\mathbf{s}_1, t_2, \mathbf{s}_2, t_1)$$

for all $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$. 2, M

unseen ⇓

(e) The covariance function is separable as we can write it as in (c) above where $C_{\mathrm{S}}(\mathbf{s}_1, \mathbf{s}_2) \propto \exp(-(||\mathbf{s}_2 - \mathbf{s}_1||/a)$ and $C_{\mathrm{T}}(t_1, t_2) \propto \exp(-|t_2 - t_1|/b)$. The covariance function is fully symmetric as all separable functions are fully symmetric, or we directly observe that

$$\begin{aligned} C(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) &= \sigma^2 \exp\{-(||\mathbf{s}_2 - \mathbf{s}_1||/a + |t_2 - t_1|/b)\} \\ &= \sigma^2 \exp\{-(||\mathbf{s}_2 - \mathbf{s}_1||/a + |t_1 - t_2|/b)\} = C(\mathbf{s}_1, t_2, \mathbf{s}_2, t_1), \end{aligned}$$

as in part (d). The parameter a defines the rate at which the spatial covariances decay with distance, and the parameter b defines the rate at which the temporal covariances decay with time. 4, M

(f) Suppose that C is a stationary, fully symmetric covariance function. From Bochner's Theorem we can expand the exponential term in the integral as the product of $\cos(\boldsymbol{h}'\boldsymbol{\omega}) + i\sin(\boldsymbol{h}'\boldsymbol{\omega})$ and $\cos(u\tau) + i\sin(u\tau)$. Since C is real-valued, we can write Bochner's Theorem as

$$C(\boldsymbol{h}, u) = \iint \left(\cos(\boldsymbol{h}'\boldsymbol{\omega})\cos(u\tau) - \sin(\boldsymbol{h}'\boldsymbol{\omega})\sin(u\tau)\right) \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}.$$

Similarly,

$$C(\boldsymbol{h}, -u) = \iint \left(\cos(\boldsymbol{h}'\boldsymbol{\omega})\cos(u\tau) + \sin(\boldsymbol{h}'\boldsymbol{\omega})\sin(u\tau)\right) \mathrm{d}F(\boldsymbol{\omega}, \tau) \quad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$$

Full symmetry implies that $C(\boldsymbol{h}, u) = C(\boldsymbol{h}, -u)$ and therefore equating the two integrals above yields the representation in the question as desired. Conversely, any function C of the form in the question is fully symmetric and admits Bochner's representation, thus completing the proof. $\boxed{7, \mathrm{M}}$

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

MATH70139 Spatial Statistics Markers Comments

Question 1	Well done on this question overall, in particular the accuracy of definitions and working in higher-dimensions with a covariance constructed from the inner product of a p-dimensional vector with itself.
Question 2	This question had slightly lower scores where some students missed the discontinuity in the semi-variogram and covariance at t=0. The inequality in Kriging was a tough question but many got, the key thing was to establish what changed from adding noise, and what remained the same, and then run the algebra! More students could have commented on the properties of the differences between the errors in terms of the placement of points and the covariance function used (the inequality becomes an equality if the covariance vector K between the predicted and known points drops for example).
Question 3	This question was generally well answered, with good knowledge of the Poisson process shown. Many handled the (new) concept of the Papengelou conditional intensity very nicely, showing good understanding of the Strauss process in so doing. One common mistake in 3(b)(i) was not noticing the inequality becomes an equality for non-empty Borel sets if there are simply no points ever placed in it (the Borel set does not have to be empty for the equality to be satisfied).
Question 4	This was another question that was very well answered overall. I'm pleased that strong marks were achieved on the final chapter of the course rather than dropping off!
Question 5	Well done for a strong showing on the mastery material - thanks for engaging with it. The definitions you mostly mastered, as well as understanding on the exponential covariance in 5(e). The proof in 5(f) was the toughest but most got some partial marks at least for a good effort. The key was using the fully symmetric property to carefully cancel various terms, and then arguing the opposite direction admits a Bochner representation through its structure even though the measure F is not necessarily symmetric (for example, it can be made symmetric as the integrand is symmetric).