Random fields on a lattice

1. Working with cliques:

a) How many cliques are there on the given graph?

b) Which neighbourhood relation on 13 vertices would result in the least possible number of cliques? What is the number of cliques in that case?

c) Which neighbourhood relation on 13 vertices would result in the maximum possible number of cliques? What is the number of cliques in that case?

d) Draw a neighbourhood relation that results in exactly 20 cliques.

e) Is there any other way of representing the neighbourhood relation, other than an undirected graph?

f) Is there any neighbourhood relation relevant for the regions of Czech Republic other than the one based on the common boundary?

2. Show that a Markov chain $\{Z_1, \ldots, Z_n\}$ is a Markov random field with respect to the relation $i \sim j \Leftrightarrow |i-j| \leq 1$. Prove that the converse implication holds as follows: if $\{Z_1, \ldots, Z_n\}$ is a Markov random field with a probability density function satisfying $p(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} = (z_1, \ldots, z_n)^T$ then it is a Markov chain.

3. Consider a Markov random field on a lattice L with respect to the relation $i \sim j$. If $i \in L$ has no neighbours, i.e. $\partial i = \emptyset$, does that imply that Z_i and Z_{-i} are independent?

4. Consider a Markov random field on a lattice L with respect to the relation $i \sim j$. If $i, j \in L$ are not neighbours, i.e. $i \nsim j$, does that imply that Z_i and Z_j are independent?

5. Let L be a lattice and $\sim$ be a relation on L given by the graph $G = (L, E)$. Assume that Z is a Markov random field on L with respect to $\sim$. Consider adding a new edge, obtaining the graph $G' = (L, E')$ with $E \subset E'$. Does the random field Z considered above have the Markov property with respect to the relation given by the graph G'?

6. Let L be a lattice and $\sim$ be the relation on L given by the complete graph $G = (L, 2^L)$. Let Z be a random field on L. Show that it is Markov with respect to $\sim$. What is the factorization given by the Hammersley-Clifford theorem?

7. Consider a Gaussian random field on a lattice L, i.e. the joint distribution of $\{Z_i, i \in L\}$ is n-dimensional Gaussian. Assume the covariance matrix $\boldsymbol{\Sigma}$ is regular and hence $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$ exists. The joint probability density function of $\{Z_i, i \in L\}$ is

$$p(\boldsymbol{z}) = \frac{\sqrt{\det \boldsymbol{Q}}}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2} \sum_{i,j \in L} q_{ij}(z_i - \mu_i)(z_j - \mu_j) \right\}, \quad \boldsymbol{z} \in \mathbb{R}^L.$$

If this random field is to be Markov, what should be the neighbourhood relation?

8. Local characteristics do not determine the joint distribution. Consider a lattice with two lattice points $L = \{i, j\}$ and assume that $Z_i \mid Z_j = z_j$ has an exponential distribution with rate z_j and $Z_j \mid Z_i = z_i$ has an exponential distribution with rate z_i. Show that these conditional distributions do not correspond to any probability distribution, i.e. a (proper) joint probability density function of the vector $(Z_i, Z_j)^T$ does not exist.

9. Let Z be a Markov random field on a lattice L and with respect to the relation $i \sim j$. Assume that the random variables $\{Z_i, i \in L\}$ are binary, i.e. $S = \{0, 1\}$, and that Z_i have the same expectation. We want to test the null hypothesis of independence of $\{Z_i, i \in L\}$, taking into account the neighbourhood relation $i \sim j$. Under the assumptions above, the null hypothesis in fact states that $\{Z_i, i \in L\}$ are i.i.d. random variables.

a) Propose an appropriate test statistics T;

b) discuss how to perform the test if we can simulate from the model under the null hypothesis;

c) discuss how to perform the test if we cannot simulate from the model under the null hypothesis;

d) if the point $i \in L$ has many neighbours and the point $j \in L$ has few neighbours, the impact of Z_i on the value of T can perhaps be much higher than the impact of Z_j – propose a way how to compensate for that.

Additional exercises

10. Show that any Gibbs random field satisfies

$$p(\boldsymbol{z}_A \mid \boldsymbol{z}_{-A}) = p(\boldsymbol{z}_A \mid \boldsymbol{z}_{\partial A})$$

for any $A \subseteq L$ and $\boldsymbol{z} \in S^L$. The symbol ∂A denotes the set of neighbours of the set A, i.e. $\partial A = (\cup_{i \in A} \partial i) \setminus A$.

11. Let $S = \mathbb{N}_0$ and L be a finite lattice in $\mathbb{R}^d$. Show that if $\beta_{ij} \geq 0$ for all $i, j \in L$ such that $i \sim j$, $i \neq j$, then the constant

$$\sum_{z \in S^L} \exp\left(-\sum_{i \in L} (\log z_i! + \beta_i z_i) - \sum_{\{i,j\} \in \mathcal{C}} \beta_{ij} z_i z_j \right)$$

is finite. On the other hand, it is infinite if $\beta_{ij} < 0$ for any $i, j \in L$: $i \sim j$, $i \neq j$.

Hint: In the first case consider the configurations with $\max z_i = k$ (there are $(k+1)^n - k^n$ of those). In the second case consider the configurations with $z_i = z_j = k$ and $z_l = 0$ for $l \in L \setminus \{i, j\}$.

12. Let the random variable Z_1 have a normal distribution $N(0, \frac{1}{1-\varphi^2})$ with $|\varphi| < 1$. Consider a first-order autoregressive sequence $\{Z_1, \ldots, Z_n\}$ defined by the formula

$$Z_t = \varphi Z_{t-1} + \varepsilon_t, \quad t = 2, \ldots, n,$$

where $\{\varepsilon_2, \ldots, \varepsilon_n\}$ is a sequence of independent identically distributed random variables with the $N(0, 1)$ distribution and the sequence is independent of Z_1. Determine the variance matrix $\boldsymbol{\Sigma}$ of the vector $(Z_1, \ldots, Z_n)^T$ and the matrix $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$. Show that $\{Z_1, \ldots, Z_n\}$ is a Gaussian Markov random field with respect to the relation $i \sim j \Leftrightarrow |i - j| \leq 1$.

13. Let $\{Z_i : i \in L\}$ be a Gaussian Markov random field with the precision matrix $\boldsymbol{Q}$. Show that

$$\operatorname{corr}(Z_i, Z_j \mid \boldsymbol{Z}_{-\{i,j\}}) = -\frac{q_{ij}}{\sqrt{q_{ii} q_{jj}}}, \quad i \neq j.$$

14. Consider random variables X_1, X_2 having only values 0 or 1. We specify the conditional distributions using the logistic regression models:

$$\operatorname{logit} \mathbb{P}(X_1 = 1 \mid X_2) = \alpha_0 + \alpha_1 X_2, \quad \operatorname{logit} \mathbb{P}(X_2 = 1 \mid X_1) = \beta_0 + \beta_1 X_1,$$

where $\operatorname{logit} p = \log \frac{p}{1-p}$ and $\alpha_0, \alpha_1, \beta_0, \beta_1$ are real-valued parameters. Determine the joint distribution of the random vector $(X_1, X_2)^T$ using the Brook lemma.

Random fields on a connected domain

1. Let $d > 1$ and let $C_d(h), h \in \mathbb{R}^d$, be an autocovariance function of a random field $\{Z(u), u \in \mathbb{R}^d\}$, i.e. C_d is a positive semidefinite function on $\mathbb{R}^d$. Assume that the random field is stationary and isotropic and hence $C_d(h) = f(\|h\|_d), h \in \mathbb{R}^d$, for some function $f : [0, \infty) \to \mathbb{R}$, where $\|h\|_d$ denotes the d-dimensional Euclidean norm. For $1 \leq k < d$ define $C_k(u) = f(\|u\|_k), u \in \mathbb{R}^k$. Prove that C_k is the autocovariance function of some random field $\{Y(u), u \in \mathbb{R}^k\}$.

2. Let $\{W^H(t) : t \in \mathbb{R}^d_+\}$ be a centered Gaussian random field with the covariances

$$\mathsf{E}\, W^H(t)W^H(s) = \frac{1}{2}(\|t\|^{2H} + \|s\|^{2H} - \|t-s\|^{2H}),\ t, s \in \mathbb{R}^d_+,$$

where $H \in (0, 1)$. Such a random field is called the *Lévy's fractional Brownian random field*. Show that it is an intrinsically stationary random field and determine its variogram.

3. Consider a spherical model for the autocovariance function of a stationary isotropic random field:

$$C(\|h\|) = \sigma^2 \frac{|b(o,\varrho) \cap b(h,\varrho)|}{|b(o,\varrho)|},\ h \in \mathbb{R}^d.$$

This model is valid in the dimension d and all the lower dimensions, see Exercise 1 above. However, it is not valid in higher dimensions. Express this autocovariance function for $d = 1$ and check that it is a positive semidefinite function. Show that this function considered in $\mathbb{R}^2$ (using $\|h\|, h \in \mathbb{R}^2$, as its argument) is not positive semidefinite.

Hint: Consider the points $x_{ij} = (i\sqrt{2}\varrho, j\sqrt{2}\varrho)$, $i, j = 1, \ldots, 8$ and the coefficients $\alpha_{ij} = (-1)^{i+j}$.

4. Express the autocovariance function from the previous Exercise for $d = 2$ using elementary functions.

5. Determine the spectral density of a weakly stationary random field with the autocovariance function

$$C(h) = \exp\{-\|h\|^2\},\ h \in \mathbb{R}^d.$$

6. Discuss how to estimate the semivariogram of an isotropic intrinsic stationary random field, based on the observations $Z(x_1), \ldots, Z(x_n)$.

7. Discuss how to test the independence of two stationary random fields defined on the same domain, based on the observations $(Z_1(x_1), Z_2(x_1)), \ldots, (Z_1(x_n), Z_2(x_n))$.

Random measures

1. Show that

a) $\mu \mapsto \mu(B)$ is a measurable mapping from $(\mathcal{M}, \mathfrak{M})$ to $([0,\infty], \mathcal{B}([0,\infty]))$ for every $B \in \mathcal{B}(E)$,

b) $\mu \mapsto \mu|_B$ is a measurable mapping from $(\mathcal{M}, \mathfrak{M})$ to $(\mathcal{M}, \mathfrak{M})$ for every $B \in \mathcal{B}(E)$.

c) $\mu \mapsto \int_E f(x)\,\mu(\mathrm{d}x)$ is a measurable mapping from $(\mathcal{M}, \mathfrak{M})$ to $([0,\infty], \mathcal{B}([0,\infty]))$ for every non-negative measurable function f on E.

2. Prove that Ψ is a random measure if and only if $\Psi(B)$ is a random variable for every $B \in \mathcal{B}$.

3. Consider intependent random variables U_1 a U_2 with uniform distribution on the interval $[0,a]$, $a > 0$, and the point process Φ on $\mathbb{R}^2$ defined as

$$\Phi = \sum_{m,n\in\mathbb{Z}} \delta_{(U_1+ma, U_2+na)}.$$

Determine the intensity measure of this process.

4. Let Ψ be a random measure. Check that the following formulas hold for $B, B_1, B_2 \in \mathcal{B}$:

a) $\operatorname{var} \Psi(B) = M^{(2)}(B \times B) - \Lambda(B)^2$,

b) $\operatorname{cov}(\Psi(B_1), \Psi(B_2)) = M^{(2)}(B_1 \times B_2) - \Lambda(B_1)\Lambda(B_2)$.

5. Why the measure $M^{(n)}$ cannot have a density w.r.t. the Lebesgue measure on $\mathbb{R}^{n\cdot d}$ but the measure $\alpha^{(n)}$ can? Consider $n = 2, d = 1$.

6. Let Φ be a simple point process. Check that the following formulas hold for $B, B_1, B_2, B_3 \in \mathcal{B}$:

a) $M^{(2)}(B_1 \times B_2) = \Lambda(B_1 \cap B_2) + \alpha^{(2)}(B_1 \times B_2)$,

b) $M^{(3)}(B_1 \times B_2 \times B_3) = \Lambda(B_1 \cap B_2 \cap B_3) + \alpha^{(2)}((B_1 \cap B_2) \times B_3) + \alpha^{(2)}((B_1 \cap B_3) \times B_2) + \alpha^{(2)}((B_2 \cap B_3) \times B_1) + \alpha^{(3)}(B_1 \times B_2 \times B_3)$,

c) $\alpha^{(n)}(B \times \cdots \times B) = \mathbb{E}[\Phi(B)(\Phi(B) - 1)\cdots(\Phi(B) - n + 1)]$.

Additional exercises

7. For $0 < a < b < c$ let us consider the sets $K_1 = \{0, a, a+b, a+b+c\}$ and $K_2 = \{0, a, a+c, a+b+c\}$. Let X_0 be a random variable with the uniform distribution on the interval $[0, a+b+c]$. We define simple point processes Φ_1 and Φ_2 on $\mathbb{R}$ such that $\operatorname{supp} \Phi_i = \{x \in \mathbb{R} : x = X_0 + y + z(a+b+c), y \in K_i, z \in \mathbb{Z}\}$, $i = 1, 2$. Show that $\mathbb{P}(\Phi_1(I) = 0) = \mathbb{P}(\Phi_2(I) = 0)$ for every interval $I \subseteq \mathbb{R}$ but the distributions of Φ_1 and Φ_2 are different.

8. The Prokhorov distance for finite measures μ, ν is defined as

$$\varrho_P(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(F) \le \nu(F^\varepsilon) + \varepsilon, \nu(F) \le \mu(F^\varepsilon) + \varepsilon) \text{ for every } F \in \mathcal{F}\},$$

where $F^\varepsilon = \{x \in E : \exists y \in F, d(x,y) < \varepsilon\}$ is an open ε-neighbourhood of a closed set F. Show that ϱ_P is a metric.

Binomial, Poisson and Cox point process

1. Show that the mixed binomial point process with the Poisson distribution (with parameter λ) of the number of points N is a Poisson process with the intensity measure $\lambda \frac{\nu(\cdot)}{\nu(B)}$.

2. Let Φ be a Poisson point process with the intensity measure Λ and $B \in \mathcal{B}$ be a given Borel set. Show that $\Phi|_B$ is a Poisson point process and determine its intensity measure.

3. Consider two independent Poisson point processes Φ_1 and Φ_2 with the intensity measures Λ_1 and Λ_2. Show that $\Phi = \Phi_1 + \Phi_2$ is a Poisson process and determine its intensity measure.

4. Consider the point pattern $\{x_1, \ldots, x_n\}$ observed in a compact observation window $W \subset \mathbb{R}^2$. Suggest a test of the null hypothesis that the point pattern is a realization of a Poisson process.

5. Let Φ be a Poisson point process with the intensity measure Λ. Determine the covariance $\mathrm{cov}(\Phi(B_1), \Phi(B_2))$ for $B_1, B_2 \in \mathcal{B}$.

6. Let Φ be a binomial point process with n points in B and the measure ν. Determine the covariance $\mathrm{cov}(\Phi(B_1), \Phi(B_2))$ for $B_1, B_2 \in \mathcal{B}$.

7. Determine the second-order factorial moment measure of a binomial point process.

8. Dispersion of a random variable $\Phi(B)$ is defined as

$$D(\Phi(B)) = \frac{\mathrm{var}\,\Phi(B)}{\mathsf{E}\,\Phi(B)}, \;\; B \in \mathcal{B}_0.$$

Show that

a) for a Poisson process $D(\Phi(B)) = 1$,

b) a binomial process is underdispersed, i.e. $D(\Phi(B)) \leq 1$,

c) a Cox process is overdispersed, i.e. $D(\Phi(B)) \geq 1$.

Additional exercises

9. Let Φ be a mixed Poisson point process with the driving measure $Y \cdot \Lambda$, where Y is a non-negative random variable and Λ is a locally finite diffuse measure. Determine the covariance $\mathrm{cov}(\Phi(B_1), \Phi(B_2))$ for $B_1, B_2 \in \mathcal{B}_0$ and show that it is non-negative.

10. Determine the Laplace transform of a binomial point process.

11. Let Y be a random variable with a gamma distribution. Show that the corresponding mixed Poisson process Φ is a negative binomial process, i.e. that $\Phi(B)$ has a negative binomial distribution for every $B \in \mathcal{B}_0$.

Stationary point process

1. Show that a homogeneous Poisson point process is stationary and isotropic. Is there any stationary non-isotropic Poisson point process?

2. Based on the interpretation of the Palm distribution determine the Palm distribution and the reduced Palm distribution of a binomial point process.

3. Consider independent random variables U_1 a U_2 with uniform distribution on the interval $[0, a]$, $a > 0$, and the point process Φ in $\mathbb{R}^2$defined as

$$\Phi = \sum_{m,n \in \mathbb{Z}} \delta_{(U_1 + ma, U_2 + na)}.$$

Determine the Palm distribution and the reduced second-order moment measure of the process. Express its contact distribution function and the nearest-neighbour distribution function.

4. Show that for a homogeneous Poisson process with the intensity λ it holds that $\mathrm{PI} = \mathrm{CE} = 1$, $F(r) = G(r) = 1 - \mathrm{e}^{-\lambda \omega_d r^d}$ and $J(r) = 1$.

5. Consider the point pattern $\{x_1, \dots, x_n\}$ observed in a compact observation window $W \subset \mathbb{R}^2$ and assume it is a realization of a stationary point process. How to estimate its intensity? How to estimate the values $F(r)$ and $G(r), r > 0$?

6. Let $Y = \{Y(x) : x \in \mathbb{R}^d\}$ be a weakly stationary Gaussian random field with the mean value μ and the autocovariance function $C(x, y) = \sigma^2 r(x - y)$, where σ^2 denotes the variance and r is the autocorrelation function of the random field Y. Consider the random measure

$$\Psi(B) = \int_B \mathrm{e}^{Y(x)} \, \mathrm{d}x, \quad B \in \mathcal{B}^d.$$

The Cox point process Φ with the driving measure Ψ is called a *log-Gaussian Cox process*. Show that the distribution of Φ is determined by its intensity and its pair-correlation function.

7. Determine the pair-correlation function of

a) the Thomas process,

b) the Matérn cluster process for $d = 2$.

Additional exercises

8. Determine the pair-correlation function of a binomial point process, provided it exists.

9. For a point process with the hard-core distance $r > 0$ and the intensity λ we define the *coverage density* as $\tau = \lambda |b(o, r/2)|$. It is in fact the mean volume fraction of the union of balls with the centers in the points of the process and the radii $r/2$. Determine the maximum possible value of τ for the following models:

a) Matérn hard-core process type I,

b) Matérn hard-core process type II.