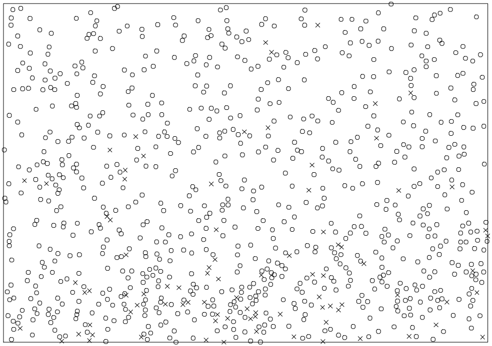
Statistical models and methods for spatial point processes

Rasmus Waagepetersen
Department of Mathematics
Aalborg University
Denmark

August 22, 2022

Mucous membrane cells

Centres of cells in mucous membrane:



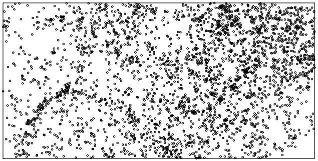
Repulsion due to physical extent of cells

Inhomogeneity - lower intensity in upper part

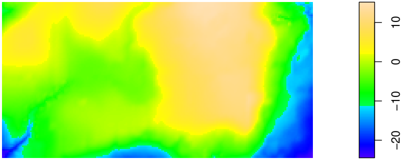
Bivariate - two types of cells

Same type of inhomogeneity for two types ?

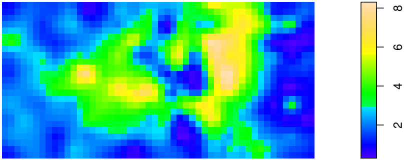
Data example: *Capparis Frondosa*



- *observation window W* $= 1000$ m $\times$ 500 m
- seed dispersal$\Rightarrow$ *clustering*
- environment $\Rightarrow$ *inhomogeneity*



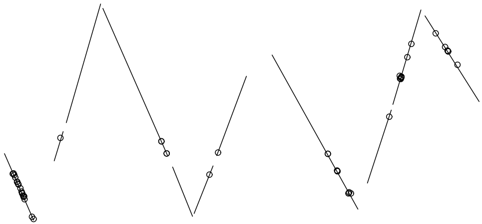
Elevation



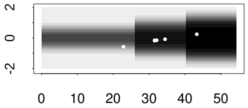
Potassium content in soil.

Objective: quantify dependence on environmental variables and clustering

Whale positions



Close up (last transect):



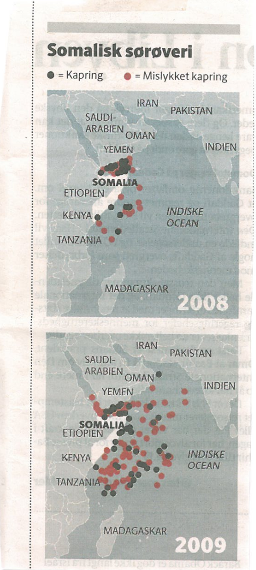
Aim: estimate whale intensity λ

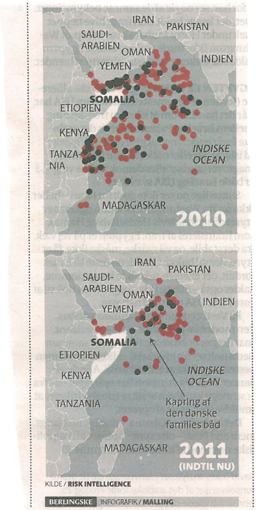
Observation window $W =$ narrow strips around transect lines

Varying detection probability: inhomogeneity (thinning)

Variation in prey intensity: clustering

Somalian pirates - two-type space-time





What is a spatial point process ?

Definitions:

1. a locally finite random subset $\mathbf{X}$ of $\mathbb{R}^2$ ($\#(\mathbf{X} \cap A)$ finite for all bounded subsets $A \subset \mathbb{R}^2$)
2. stochastic process of count variables $\{N(B)\}_{B \in \mathcal{B}_0}$ indexed by bounded Borel sets $\mathcal{B}_0$.
3. a random counting measure N on $\mathbb{R}^2$

Equivalent provided no multiple points: $(N(\{u\}) \in \{0, 1\})$

This course: appeal to 1. and skip measure-theoretic details.

In practice distribution specified by an explicit construction (second and third lecture) or in terms of a probability density (second and fifth lecture).

Moments of a spatial point process

Fundamental characteristics of point process: mean and covariance of counts $N(A) = \#(\mathbf{X} \cap A)$.

Intensity measure μ:

$$\mu(A) = \mathbb{E}N(A), \quad A \subseteq \mathbb{R}^2$$

In practice often given in terms of *intensity function*

$$\mu(A) = \int_A \rho(u)\mathrm{d}u$$

(if such a $\rho(\cdot)$ exists)

Infinitesimal interpretation: $N(A)$ binary variable (presence or absence of point in A) when A very small. Hence

$$\rho(u)\mathrm{d}A \approx \mathbb{E}N(A) \approx P(\mathbf{X} \text{ has a point in A})$$

Second-order moments

Second order factorial moment measure:

$$\alpha^{(2)}(A \times B) = \mathrm{E} \sum_{u,v \in \mathbf{X}}^{\neq} \mathbf{1}[u \in A,\ v \in B] \qquad A, B \subseteq \mathbb{R}^2$$
$$= \int_A \int_B \rho^{(2)}(u, v)\, \mathrm{d}u\, \mathrm{d}v$$

where $\rho^{(2)}(u, v)$ is the *second order product density* (provided it exists)

Infinitesimal interpretation of $\rho^{(2)}$:

$$\rho^{(2)}(u, v) dA dB \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)$$

$(u \in A\ , v \in B)$

Second moment vs. second factorial moment measure

Second moment measure

$$\mu^{(2)}(A \times B) = \mathbb{E}N(A)N(B) = \mathbb{E}\sum_{u,v\in\mathbf{X}} 1[u \in A, v \in B]$$

$$
\begin{aligned}
=&\mathbb{E}\sum_{u,v\in\mathbf{X}}^{\neq} 1[u \in A, v \in B] + \mathbb{E}\sum_{u\in\mathbf{X}} 1[u \in A \cap B] \\
=&\alpha^{(2)}(A \times B) + \mu(A \cap B)
\end{aligned}
$$

Hence $\mu^{(2)}$ not absolutely continuous due to "diagonal terms" in sum $N(A)N(B) = \sum_{u,v\in\mathbf{X}} 1[u \in A, v \in B]$.

Campbell formulae

By definition of intensity function and product density and the standard proof we obtain the useful Campbell formulae:

$$\mathbb{E} \sum_{u \in \mathbf{X}} h(u) = \int h(u)\rho(u)\mathrm{d}u$$

$$\mathbb{E} \sum_{u,v \in \mathbf{X}}^{\neq} h(u,v) = \iint h(u,v)\rho^{(2)}(u,v)\mathrm{d}u\mathrm{d}v$$

for non-negative functions h on $\mathbb{R}^2$ or $\mathbb{R}^2 \times \mathbb{R}^2$.

(i.e. replace indicator functions $1[u \in A]$ and $1[u \in A, v \in B]$ with general non-negative functions $h(u)$ and $h(u,v)$)

Pair correlation function

Pair correlation tendency to cluster/repel relative to case of independent points:

$$g(u,v) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)} = \frac{P(\mathbf{X} \text{ has a point at } u \text{ and at } v)}{P(\mathbf{X} \text{ has a point at } u)P(\mathbf{X} \text{ has a point at } v)}$$

$$= \frac{P(\mathbf{X} \text{ has a point at } u | \mathbf{X} \text{ has a point at } v)}{P(\mathbf{X} \text{ has a point at } u)}$$

$$= 1 \text{ if independence (Poisson process, next section)}$$

Let $\rho(u|v)$ denote intensity of $\mathbf{X}$ given $v \in \mathbf{X}$ ('Palm' intensity). Then

$$g(u,v) = \frac{\rho(u|v)}{\rho(u)}$$

Covariance and pair correlation function

$$\mathbb{C}\mathrm{ov}[N(A), N(B)] = \int_{A \cap B} \rho(u)\mathrm{d}u + \int_A \int_B \rho(u)\rho(v)(g(u,v) - 1)\mathrm{d}u\mathrm{d}v \tag{1}$$

$$\begin{aligned}\mathbb{V}\mathrm{ar} N(A) &= \int_A \rho(u)\mathrm{d}u + \int_A \int_A \rho(u)\rho(v)(g(u,v) - 1)\mathrm{d}u\mathrm{d}v \\ &= \text{Poisson variance} + \text{additional/less variance due to interaction}\end{aligned}$$

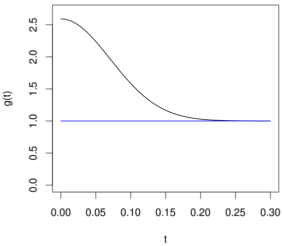
In some sense always a ‘nugget’ effect

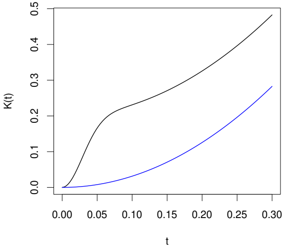
K-function

$$K(t) = \int_{\|h\| \le t} g(h) \mathrm{d}h$$

(provided $g(u, v) = g(u - v)$ i.e. **X** *second-order reweighted stationary*)

Examples of pair correlation and K-functions assuming g isotropic $g(u, v) = g_0(\|v - u\|)$:





Estimation and interpretation of $K(t)$

Unbiased estimate of K-function (W observation window):

$$\hat{K}(t) = \sum_{u,v \in \mathbf{X} \cap W}^{\neq} \frac{1[\|u - v\| \leq t]}{\rho(u)\rho(v)} e_{u,v}$$

($e_{u,v}$ edge correction factor, see exercises 3 and 4)

In the homogeneous case (constant intensity ρ) $K(t)$ has interpretation as conditional expectation:

$$\rho K(t) = \mathbb{E}[\text{number of further points within distance } t \text{ of } u | u \in \mathbf{X}]$$

Exercises

1. 1.1 Show that the covariance between counts $N(A)$ and $N(B)$ is

$$\mathbb{C}\mathrm{ov}[N(A), N(B)] = \mu(A \cap B) + \alpha^{(2)}(A \times B) - \mu(A)\mu(B)$$

   1.2 Verify covariance formula (1) (covariance in terms of pair correlation function).

2. Show that for any $B \subset \mathbb{R}^2$ with $0 < |B| < \infty$,

$$K(t) := \int_{\mathbb{R}^2} 1[\|u\| \leq t] g(u) \mathrm{d}u = \frac{1}{|B|} \mathbb{E} \sum_{\substack{u \in \mathbf{X} \cap B \\ v \in \mathbf{X}}}^{\neq} \frac{1[\|u - v\| \leq t]}{\rho(u)\rho(v)}$$

(Hint: use the Campbell formula)

3. Show that the following estimate is unbiased:

$$\hat{K}(t) = \sum_{u,v \in \mathbf{X} \cap W}^{\neq} \frac{1[\|u - v\| \leq t]}{\rho(u)\rho(v)|W \cap W_{u-v}|}$$

where W_{u-v} translated version of W (assume $|W \cap W_h| > 0$ for $\|h\| \leq t$).

4. Show that in the isotropic case $(g(u,v) = g_0(\|u - v\|))$, $K'(r) = 2\pi r g(r)$.

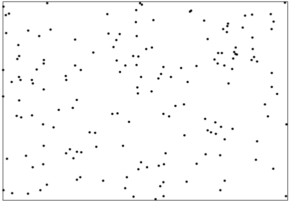
The Poisson process

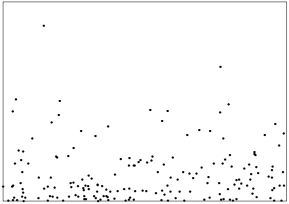
Assume μ locally finite measure on $S \subseteq \mathbb{R}^2$ with density ρ.

$\mathbf{X}$ is a Poisson process on S with intensity measure μ if for any region $B \subseteq S$ with $\mu(B) < \infty$:

1. $N(B) \sim \text{Poisson}(\mu(B))$
2. Given $N(B) = n$, $\mathbf{X} \cap B$ consists of n points that are *iid* with density $\rho(u)/\mu(B)$, $u \in B$[1]

Example: $B = [0,1] \times [0,0.7]$:



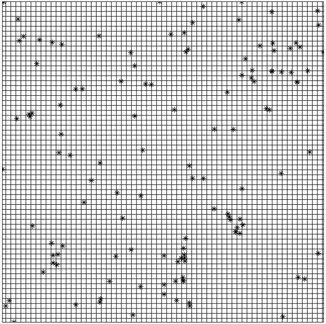


Homogeneous: $\rho = 150/0.7$ Inhomogeneous: $\rho(x,y) \propto \mathrm{e}^{-10.6y}$

[1]Provided $\mu(B) > 0$. If $\mu(B) = 0$ then $N(B) = 0$ almost surely

Homogeneous Poisson process as limit of Bernouilli trials

Consider disjoint subdivision $W = \cup_{i=1}^{n} C_i$ where $|C_i| = |W|/n$. With probability $\rho|C_i|$ a uniform point is placed in C_i.



Number of points in subset A is $b(n|A|/|W|, \rho|W|/n)$ which converges to a Poisson distribution with mean $\rho|A|$.

Hence, Poisson process default model when points occur independently of each other.

Characterization in terms of void probabilities

The distribution of any point process $\mathbf{X}$ is uniquely determined by the void probabilities $P(\mathbf{X} \cap B = \emptyset)$, for bounded subsets $B \subseteq \mathbb{R}^2$.

Intuition: consider very fine subdivision of observation window – then at most one point in each cell and (joint) probabilities of absence/presence determined by void probabilities.

Hence, a point process $\mathbf{X}$ with intensity measure μ is a Poisson process if and only if

$$P(\mathbf{X} \cap B = \emptyset) = \exp(-\mu(B))$$

for any bounded subset B.

Existence of Poisson process on $\mathbb{R}^2$: use definition on disjoint partitioning $\mathbb{R}^2 = \cup_{i=1}^{\infty} B_i$ of bounded sets B_i.

Check by assessing void probabilities, that constructed process is indeed a Poisson process.

Distribution and moments of Poisson process

X a Poisson process on S with $\mu(S) = \int_S \rho(u)\mathrm{d}u < \infty$ and F set of finite point configurations in S.

Examples of F: all point configurations with total number of points in a given interval, point configurations where all pairs of points separated by distance δ,...

By definition of a Poisson process and law of total probability

$$P(\mathbf{X} \in F) = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu(S)}}{n!} \int_{S^n} 1[\{x_1, x_2, \ldots, x_n\} \in F] \prod_{i=1}^{n} \rho(x_i)\mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (2)$$

Similarly,

$$\mathbb{E}h(\mathbf{X}) = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu(S)}}{n!} \int_{S^n} h(\{x_1, x_2, \ldots, x_n\}) \prod_{i=1}^{n} \rho(x_i)\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Independence properties:

- $\rho^{(2)}(u, v) = \rho(u)\rho(v)$ and $g(u, v) = 1$ (exercise)
- $\mathbb{C}\mathrm{ov}[N(A), N(B)] = \int_{A \cap B} \rho(u)\mathrm{d}u$
- $A, B \subseteq \mathbb{R}^2$ disjoint $\Rightarrow$ $\mathbf{X} \cap A$ and $\mathbf{X} \cap B$ independent

Superpositioning and thinning

If $\mathbf{X}_1, \mathbf{X}_2, \ldots$ are independent Poisson processes (ρ_i), then *superposition* $\mathbf{X} = \cup_{i=1}^{\infty} \mathbf{X}_i$ is a Poisson process with intensity function $\rho(u) = \sum_{i=1}^{\infty} \rho_i(u)$ (provided ρ integrable on bounded sets).

Conversely: *Independent π-thinning* of Poisson process $\mathbf{X}$: independent retain each point u in $\mathbf{X}$ with probability $\pi(u)$. Thinned process $\mathbf{X}_{\text{thin}}$ and $\mathbf{X} \setminus \mathbf{X}_{\text{thin}}$ are independent Poisson processes with intensity functions $\pi(u)\rho(u)$ and $(1 - \pi(u))\rho(u)$.

(Superpositioning and thinning results most easily verified using void probability characterization of Poisson process, see M & W, 2003)

For general point process $\mathbf{X}$: thinned process $\mathbf{X}_{\text{thin}}$ has intensity $\pi(u)\rho(u)$ and product density $\pi(u)\pi(v)\rho^{(2)}(u, v)$ - hence g and K invariant under independent thinning.

Primer regarding probability densities

We say that a random variable Y has density h if

$$P(Y \in C) = \int 1[y \in C] h(y) \mathrm{d}y.$$

More generally we can say that Y has density f with respect to the distribution of another random variable Z provided

$$P(Y \in C) = \mathbb{E}[1[Z \in C] f(Z)]$$

for all $C \subseteq \mathbb{R}$.

Example: suppose Z is $N(0,1)$ with density ϕ. Then

$$P(Y \in C) = \int 1[y \in C] \frac{h(y)}{\phi(y)} \phi(y) \mathrm{d}y = \mathbb{E}\left[1[Z \in C] \frac{h(Z)}{\phi(Z)}\right]$$

Hence, Y has density $f(z) = h(z)/\phi(z)$, $z \in \mathbb{R}$ with respect to the distribution of Z.

Density (likelihood) of a finite Poisson process

$\mathbf{X}_1$ and $\mathbf{X}_2$ Poisson processes on S with intensity functions ρ_1 and ρ_2 where $\int_S \rho_i(u)\mathrm{d}u < \infty$, $i = 1, 2$, and $\rho_2(u) = 0 \Rightarrow \rho_1(u) = 0$. Define $0/0 := 0$. Then

$$
\begin{aligned}
& P(\mathbf{X}_1 \in F) \\
=& \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu_1(S)}}{n!} \int_{S^n} 1[\mathbf{x} \in F] \prod_{i=1}^{n} \rho_1(x_i)\mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (\mathbf{x} = \{x_1, \ldots, x_n\}) \\
=& \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu_2(S)}}{n!} \int_{S^n} 1[\mathbf{x} \in F] \mathrm{e}^{\mu_2(S)-\mu_1(S)} \prod_{i=1}^{n} \frac{\rho_1(x_i)}{\rho_2(x_i)} \prod_{i=1}^{n} \rho_2(x_i)\mathrm{d}x_1 \ldots \mathrm{d}x_n \\
=& \mathbb{E}\big(1[\mathbf{X}_2 \in F] f(\mathbf{X}_2)\big)
\end{aligned}
$$

where

$$f(\mathbf{x}) = \mathrm{e}^{\mu_2(S)-\mu_1(S)} \prod_{i=1}^{n} \frac{\rho_1(x_i)}{\rho_2(x_i)}$$

Hence f is a density of $\mathbf{X}_1$ with respect to distribution of $\mathbf{X}_2$.

In particular (if S bounded): $\mathbf{X}_1$ has density

$$f(\mathbf{x}) = \mathrm{e}^{\int_S (1-\rho_1(u))\mathrm{d}u} \prod_{i=1}^{n} \rho_1(x_i)$$

with respect to unit rate Poisson process ($\rho_2 = 1$).

Exercises

1. What is $K(t)$ for a Poisson process ?
2. Check that the Poisson expansion (2) indeed follows from the definition of a Poisson process.
3. How can you simulate an inhomogeneous Poisson process on a bounded region B in case $\rho(u)/\mu(B)$ is not a standard probability density ?
4. Show that $\rho^{(2)}(u, v) = \rho(u)\rho(v)$ for a Poisson process $\mathbf{X}$.

   (Hint: a) use that counts on disjoint subsets uncorrelated or b) compute second order factorial measure using the Poisson expansion for $\mathbf{X} \cap (A \cup B)$ for bounded $A, B \subseteq \mathbb{R}^2$.)

5. Assume that $\mathbf{X}$ has second order product density $\rho^{(2)}$ and show that g (and hence K) is invariant under independent thinning (note that a heuristic argument follows easily from the infinitesimal interpretation of $\rho^{(2)}$).

   (Hint: introduce random field $\mathbf{R} = \{R(u) : u \in \mathbb{R}^2\}$, of independent uniform random variables on $[0, 1]$, and independent of $\mathbf{X}$, and compute second order factorial measure for thinned process $\mathbf{X}_{\text{thin}} = \{u \in \mathbf{X} | R(u) \leq p(u)\}$.)

(solutions to exercises 4 and 5 available in the end of the set of slides)

Cox processes

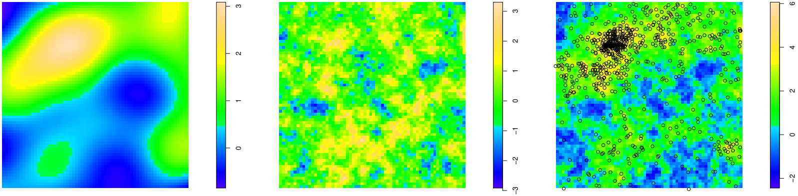
X is a *Cox process* driven by the random intensity function Λ if, conditional on $\Lambda = \lambda$, **X** is a Poisson process with intensity function λ.

Example: log Gaussian Cox process ("point process GLMM")

$$\log \Lambda(u) = \beta Z(u)^{\mathsf{T}} + Y(u)$$

where $\{Y(u)\}$ Gaussian random field.

$\beta Z(u)^{\mathsf{T}}$ $\qquad$ $Y(u)$ $\qquad$ $\beta Z(u)^{\mathsf{T}} + Y(u)$



Z: systematic variation Y: random clustering around peaks in Y

Wide range of covariance models available for Y: exponential, Gaussian, Matérn,...

Cox processes "bridge" between point processes and geostatistics.

Shot-noise Cox process

$$\Lambda(u) = \sum_{v \in \mathbf{C}} \gamma_v k(u - v)$$

where

- $\mathbf{C}$ homogeneous Poisson with intensity κ
- $k(\cdot)$ probability density.
- γ_v *iid* positive random variables independent of $\mathbf{C}$

NB: equivalent to cluster process

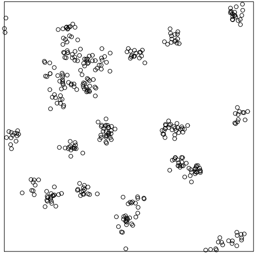
$$\mathbf{X} = \cup_{v \in \mathbf{C}} \mathbf{X}_v$$

where $\mathbf{X}_v$'s independent Poisson processes with intensity functions $\gamma_v k(\cdot - v)$.

Inhomogeneous shot-noise:

$$\Lambda(u) = \exp[\beta Z(u)^{\mathsf{T}}] \sum_{v \in \mathbf{C}} \gamma_v k(u - v)$$

Cox/Cluster process example: Inhomogeneous Thomas process



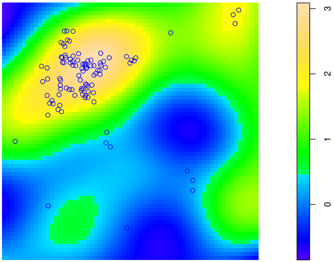
Parents stationary Poisson point process intensity κ

Poisson(α) number of offspring distributed around parents according to bivariate Gaussian density

Inhomogeneity: offspring survive according to probability

$$p(u) \propto \exp(Z(u)\beta^{\mathsf{T}})$$

depending on covariates (independent thinning).



Moments for Cox processes

Intensity function

$$\rho(u) = \mathbb{E}\Lambda(u)$$

Second-order product density and pair correlation

$$\rho^{(2)}(u,v) = \mathbb{E}\Lambda(u)\Lambda(v) = \mathbb{C}\mathrm{ov}[\Lambda(u),\Lambda(v)] + \rho(u)\rho(v)$$

$$g(u,v) = 1 + \frac{\mathbb{C}\mathrm{ov}[\Lambda(u),\Lambda(v)]}{\rho(u)\rho(v)}$$

$$\begin{aligned}
\mathbb{C}\mathrm{ov}[N(A),N(B)] &= \int_{A\cap B} \mathbb{E}\Lambda(u)\mathrm{d}u + \int_A\int_B \mathbb{C}\mathrm{ov}[\Lambda(u),\Lambda(v)]\mathrm{d}u\mathrm{d}v \\
&= \int_{A\cap B} \rho(u)\mathrm{d}u + \int_A\int_B \rho(u)\rho(v)[g(u,v)-1]\mathrm{d}u\mathrm{d}v \\
&= \text{ Poisson variance } + \text{ extra variance due to } \Lambda
\end{aligned}$$

(overdispersion relative to a Poisson process, see also exercise 1)

Common structure: log-linear model

Both log Gaussian and shot-noise Cox process of the form

$$\Lambda(u) = \Lambda_0(u) \exp[\beta Z(u)^\mathsf{T}]$$

where Λ_0 stationary non-negative reference process.

(interpretation: Cox process $\mathbf{X}$ independent inhomogeneous thinning of stationary $\mathbf{X}_0$ with random intensity function Λ_0).

Log-linear intensity (assume $\mathbb{E}\Lambda_0(u) = 1$)

$$\rho(u) = \mathbb{E}\Lambda(u) = \exp[\beta Z(u)^\mathsf{T}]$$

Pair correlation function ($\mathbb{E}\Lambda_0(u) = 1$):

$$g(h) = 1 + c_0(h) \quad c_0(h) = \mathbb{C}\mathrm{ov}[\Lambda_0(u), \Lambda_0(u+h)]$$

Specific models for $c_0(u-v) = \mathbb{C}\mathrm{ov}[\Lambda_0(u), \Lambda_0(v)]$

Log-Gaussian:

$$\Lambda_0(u) = \exp[Y(u)]$$

where Y Gaussian field.

Covariance (Laplace transform of normal distribution):

$$c_0(h) = \exp[\mathbb{C}\mathrm{ov}(Y(u), Y(u+h))] - 1$$

Shot-noise:

$$\Lambda_0(u) = \sum_{v \in C} \gamma_v k(u-v)$$

Covariance (convolution):

$$c_0(u-v) = \kappa\alpha^2 \int_{\mathbb{R}^2} k(u)k(u+h)\mathrm{d}u$$

$(\alpha = \mathbb{E}\gamma_v)$

Convolution

$$\int_{\mathbb{R}^2} k(u)k(u+h)\mathrm{d}u$$

is density of $Y_1 - Y_2$ where Y_1 and Y_2 independent each with density $k(\cdot)$.

We would like simple expression for convolution.

Ex: for Thomas Y_1 and Y_2 both bivariate Gaussian.

Normal-variance mixtures: suppose $Y = \sqrt{W}U$ where W positive random variable and $U \sim N_p(0, I)$.

Then easy to show that distribution of Y is closed under convolution if distribution of W is closed under convolution.

Ex: W Gamma $\Rightarrow$ Y variance-gamma distributed $\Rightarrow$ Matérn covariance

Ex: W inverse Gamma $\Rightarrow$ Y Cauchy $\Rightarrow$ Cauchy covariance

Density of a Cox process

- Restricted to a bounded region W, the density is

$$f(\mathbf{x}) = \mathbb{E}[f(\mathbf{x}|\Lambda)] = \mathbb{E}\left[\exp\left(|W| - \int_W \Lambda(u)\,\mathrm{d}u\right) \prod_{u\in\mathbf{X}} \Lambda(u)\right]$$

- Not on closed form
- likelihood-based inference: MCMC or Laplace approximation (INLA for log Gaussian Cox processes)
- estimating equations based on closed form expressions for intensity and pair correlation

Exercises

1. For a Cox process with random intensity function Λ, show that

   $$\mathbb{V}\mathrm{ar} N(A) \geq \mathbb{E}N(A), \quad \rho(u) = \mathbb{E}\Lambda(u), \quad \rho^{(2)}(u,v) = \mathbb{E}[\Lambda(u)\Lambda(v)]$$

   (hint: use conditioning on Λ)
2. Show that pair correlation for LCGP is
   $g(u,v) = \exp[\mathbb{C}\mathrm{ov}(Y(u), Y(v))]$

   (hint: use previous exercise and expression for Laplace transform of a Gaussian random variable)
3. Show that a cluster process with Poisson(α) number of iid offspring is a Cox process with random intensity function

   $$\Lambda(u) = \alpha \sum_{v \in \mathbf{C}} k(u-v)$$

   (using notation from previous slide on cluster processes. Hint: use void probability characterisation and superposition result for Poisson process. Note: **C** can be any point process)

Exercises

4. Compute the intensity and second-order product density for an inhomogeneous Thomas process. Deduce that the pair correlation function is

$$g(u, v) = 1 + (4\pi\omega^2)^{-d/2} \exp[-\{r/(2\omega)\}^2]/\kappa$$

(Hint: interpret the Thomas process as a Cox process and use the Campbell formulae)

1. Intro to point processes and moment measures
2. The Poisson process
3. Cox and cluster processes
4. Estimating functions
5. The conditional intensity and Markov point processes
6. References

Maximum likelihood estimation for Poisson

Log likelihood for Poisson process with intensity function ρ_θ:

$$l(\theta) = \sum_{u \in \mathbf{X}} \log \rho_\theta(u) - \int_W \rho_\theta(u) \mathrm{d}u$$

Score (first derivative):

$$s(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta} l(\theta) = \sum_{u \in \mathbf{X}} \frac{\rho'_\theta(u)}{\rho_\theta(u)} - \int_W \rho'_\theta(u) \mathrm{d}u$$

Find $\hat{\theta}$ by solving $s(\theta) = 0$. Unique solution if observed information

$$-\frac{\mathrm{d}^2}{\mathrm{d}\theta^\mathsf{T} \mathrm{d}\theta} l(\theta)$$

positive definite.

Information matrix:

$$i(\theta) = -\mathbb{E}\frac{\mathrm{d}^2}{\mathrm{d}\theta^\mathsf{T}\mathrm{d}\theta} l(\theta)$$

Under weak regularity conditions,

$$\hat{\theta} \approx N(\theta, i(\theta)^{-1})$$

If Poisson process not appropriate due to clustering we might consider Cox/cluster processes but likelihood function is then hard to compute.

To move on, estimating function perspective is useful.

Estimating function

Estimating function: $e(\theta)$ $[e(\theta, \mathbf{X})]$ function of θ and data $\mathbf{X}$.

Parameter estimate $\hat{\theta}$ solution of

$$e(\theta) = 0$$

First order Taylor:

$$e(\theta) \approx S(\hat{\theta} - \theta)$$

where sensitivity matrix:

$$S = -\mathbb{E}[\frac{\mathrm{d}}{\mathrm{d}\theta} e(\theta)]$$

minus expected derivative of $e(\theta)$

Using Taylor approximation: $\hat{\theta}$ approx. unbiased $\mathbb{E}\hat{\theta} = \theta$ if $e(\theta)$ unbiased $\mathbb{E}e(\theta) = 0$ (θ true value).

Moreover ('sandwich'-variance estimator):

$$\mathbb{V}\mathrm{ar}\hat{\theta} \approx S^{-1}\Sigma S^{-\mathsf{T}} \quad \Sigma = \mathbb{V}\mathrm{ar}e(\theta)$$

Note: in case of Poisson process and $e(\theta)$ equal to likelihood score, $S = \mathbb{V}\mathrm{ar}e(\theta) = i(\theta)$ whereby $\mathbb{V}\mathrm{ar}\hat{\theta} = i(\theta)^{-1}$.

How do we construct unbiased estimating functions involving **X** and θ ?

Composite likelihood

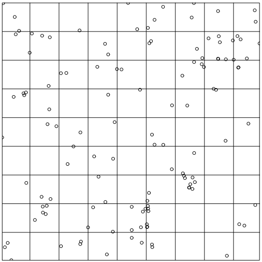
Disjoint subdivision $W = \cup_{i=1}^{m} C_i$ in 'cells' C_i.

$u_i \in C_i$ 'center' point.

Random indicator variables:

$$Y_i = 1[\mathbf{X} \text{ has a point in } C_i]$$

(presence/absence of points in C_i).



$$P(Y_i = 1) = |C_i| \rho_\theta(u_i)$$

Idea: form composite likelihoods based on Y_i, e.g.

$$\prod_i P(Y_i = 1)^{Y_i} (1 - P(Y_i = 1))^{1 - Y_i}$$

Consider limit when $|C_i| \to 0$.

Composite likelihood (in fact likelihood for Poisson):

$$\left[\prod_{u \in \mathbf{X}} \rho_\theta(u)\right] \exp\big[\int_W \rho_\theta(u)\mathrm{d}u\big]$$

Score:

$$e(\theta) = \sum_{u \in \mathbf{X}} \frac{\rho'_\theta(u)}{\rho_\theta(u)} - \int_W \rho'_\theta(u)\mathrm{d}u$$

unbiased estimating function by Campbell.

Sensitivity is equal to Information matrix for Poisson process.
Variance

$$\mathbb{V}\mathrm{ar}\, e(\theta) = \mathbb{V}\mathrm{ar} \sum_{u \in \mathbf{X}} \frac{\rho'_\theta(u)}{\rho_\theta(u)}$$

can be evaluated using second Campbell formula. Larger than $i(\theta)$ in case of Cox/cluster ($g_\theta(\cdot) > 1$).

Note: to evaluate sandwich estimator of variance

$$S^{-1}\mathbb{V}\mathrm{ar}e(\theta)S^{-\mathsf{T}}$$

of parameter estimates, we need estimate of pair correlation function (later).

Other issue:

- integral

  $$\int_W \rho'_\theta(u)\mathrm{d}u$$

  often not explicitly computable.

  Can be approximated fairly easy using numerical quadrature or Monte Carlo (later).

Estimation of pair correlation function

Suppose parametric model $g(\cdot;\psi)$ for pair correlation.

Some options:

1. minimum contrast estimation based on K-function.
2. second-order composite likelihood: composite likelihood based on indicators for joint occurrence of points in pairs of cells:

$$X_{ij} = 1[N_i > 0 \text{ and } N_j > 0]$$

Minimum contrast estimation for ψ

Computationally easy alternative if $\mathbf{X}$ second-order reweighted stationary so that K-function well-defined.

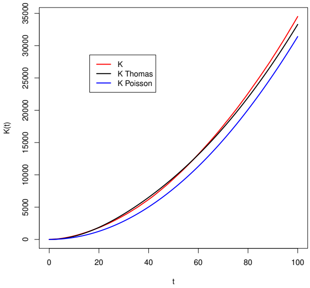
Estimate of K-function:

$$\hat{K}_\beta(t) = \sum_{u,v \in \mathbf{X} \cap W} \frac{1[0 < \|u-v\| \le t]}{\rho(u;\beta)\rho(v;\beta)} e_{u,v}$$

Unbiased if β 'true' regression parameter.

Minimum contrast estimation: minimize squared distance between theoretical K and $\hat{K}$:

$$\hat{\psi} = \operatorname*{argmin}_{\psi} \int_0^r \left(\hat{K}_{\hat{\beta}}(t) - K(t;\psi)\right)^2 \mathrm{d}t$$



Second-order composite likelihood

Consider indicators for *joint* occurrence of points in pairs of cells:

$$X_{ij} = 1[N_i > 0 \text{ and } N_j > 0]$$

with

$$\begin{aligned} P_{\beta,\psi}(X_{ij} = 1) &= \rho^{(2)}(u, v; \beta, \psi)|C_i||C_j| \\ &= \rho_\beta(u_i)\rho_\beta(v_j)g(u_i - u_j; \psi)|C_i||C_j| \end{aligned}$$

Second-order composite likelihood:

$$\mathrm{CL}_2(\beta, \psi) = \prod_{\substack{u,v \in \mathbf{X} \cap W \\ \|u-v\| \le R}}^{\neq} \rho^{(2)}(u, v; \beta, \psi) \times \exp\left[- \iint_{\|u-v\| \le R} \rho^{(2)}(u, v; \beta, \psi) \mathrm{d}u \mathrm{d}v \right]$$

NB: second-order reweighted stationarity (translation invariant pair correlation) not required.

In practice we plug in $\hat{\beta}$ from first order composite likelihood.

Two-step estimation

Obtain estimates $(\hat{\beta}, \hat{\psi})$ in two steps

1. obtain $\hat{\beta}$ using composite likelihood
2. obtain $\hat{\psi}$ using minimum contrast/second order composite likelihood (replacing β by $\hat{\beta}$ from first step)

Implementation spatstat

Two-step estimation implemented in `spatstat` procedure `kppm`

Options composite likelihood, quasi-likelihood, minimum contrast, second-order composite likelihood,...

Exercises

1. Compute information matrix and variance of log likelihood score in case of a Poisson process with intensity function $\rho_\theta(\cdot)$.
2. Obtain expression for $\mathbb{V}\mathrm{ar}e(\theta)$ in terms of pair correlation function g in case of first order composite likelihood.
3. Check that the derivative of minimum contrast criterion and the score of the second order composite likelihood function are unbiased estimating functions when β is equal to the true value.
4. How can you partition a Poisson-cluster process $\mathbf{X}$ into a union $\cup_{i=1}^{n}\mathbf{X}_i$ of *iid* Poisson-cluster processes ?
5. show that the approximate composite likelihood score (3) is of logistic regression score form when the intensity is log linear.
6. Derive the second-order product density of a stratified binomial point process with one point in each cell.

Approximation of integral in composite likelihood

Issue: integral

$$\int_W \rho'(u)\mathrm{d}u$$

in composite likelihood typically not available in closed form.

Deterministic numerical quadrature:

1. resulting estimating function not unbiased
2. difficult to quantify resulting bias of parameter estimates.

Monte Carlo approximation of integral in composite likelihood

Let **D** ‘quadrature/dummy’ point process of intensity κ and independent of **X**.

By Campbell

$$\int_W \rho'(u)\mathrm{d}u = \mathbb{E} \sum_{u \in \mathbf{X} \cup \mathbf{D}} \frac{\rho'(u)}{\rho(u)+\kappa} \approx \sum_{u \in \mathbf{X} \cup \mathbf{D}} \frac{\rho'(u)}{\rho(u)+\kappa}$$

Idea: replace integrals in pseudo- or composite likelihood with unbiased estimates using **D**.

Advantages:

1. unbiased approximation $\Rightarrow$ still unbiased estimating function !
2. CLT available for approximation $\Rightarrow$ CLT for parameter estimates.

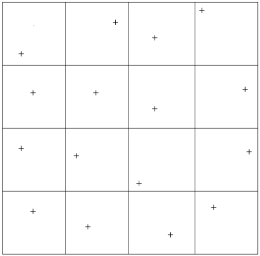
Dummy point process

Should be easy to simulate and mathematically tractable.

Possibilities:

1. Poisson process
2. binomial point process (fixed number of independent points)
3. stratified binomial point process

Stratified:



Approximate composite likelihood scores:

$$s(\theta) = \sum_{u \in \mathbf{X}} \frac{\rho_\theta'(u)}{\rho_\theta(u)} - \sum_{u \in (\mathbf{X} \cup \mathbf{D})} \frac{\rho_\theta'(u)}{\rho_\theta(u) + \kappa} \tag{3}$$

Note: of *logistic regression/case control* form with ‘probabilities’

$$p(u) = \frac{\rho_\theta(u)}{\rho_\theta(u) + \kappa}$$

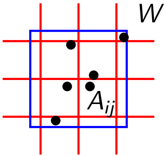
I.e. probabilities that $u \in \mathbf{X}$ given $u \in \mathbf{X} \cup \mathbf{D}$.

Hence computations straightforward with `glm()` software !

Monte Carlo and deterministic numerical quadrature implemented in spatstat procedure `ppm`

Asymptotic results - first order estimating function

Divide $\mathbb{R}^2$ into quadratic cells
$A_{ij} = [i, i+1[\times[j, j+1[$



Then

$$e_f(\beta) = \sum_{ij: A_{ij} \subseteq W} U_{ij}$$

where

$$U_{ij} = \sum_{u \in \mathbf{X} \cap A_{ij}} f_\beta(u) - \int_{A_{ij}} f_\beta(u) \rho_\beta(u) \mathrm{d}u$$

Assuming $\mathbf{X}$ is mixing, $\{U_{ij}\}_{ij}$ mixing random field and

$$|W|^{-1/2} e_f(\beta) \approx N(0, \Sigma_f)$$

(CLT for mixing random field $\{U_{ij}\}_{ij}$).

Asymptotic results cntd.

Estimate $\hat{\beta}$ solves

$$e_f(\beta) = 0$$

And (Taylor)

$$e_f(\beta) \approx |W| S_f(\hat{\beta} - \beta) \Leftrightarrow (\hat{\beta} - \beta) = |W|^{-1} S_f^{-1} e_f(\beta)$$

where

$$S_f = -\mathbb{E} \frac{\mathrm{d}}{\mathrm{d}\beta^{\mathsf{T}}} e_f(\beta) / |W|$$

It follows that

$$\hat{\beta} \approx N(\beta, V_f / |W|)$$

where

$$V_f = S_f^{-1} \Sigma_f S_f^{-\mathsf{T}}$$

Alternative: “infill” /increasing intensity-asymptotics

If $\mathbf{X}$ infinitely divisible (e.g. Poisson or Poisson-cluster) then

$$\mathbf{X} = \cup_{i=1}^{n} \mathbf{X}_i$$

where $\mathbf{X}_i$ *iid* and intensity of $\mathbf{X}$ is $\rho_\beta(u) = n\tilde{\rho}(u;\beta)$ where $\tilde{\rho}_\beta$ intensity of $\mathbf{X}_i$.

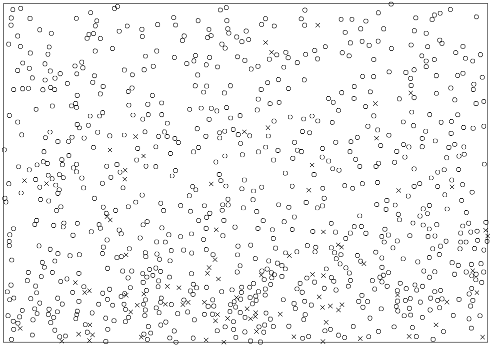
Thus

$$e_f(\beta) = \sum_{i=1}^{n} \Big[\sum_{u \in \mathbf{X}_i} f_\beta(u) - \int_W f_\beta(u)\tilde{\rho}(u;\beta)\mathrm{d}u \Big].$$

Ordinary CLT applies !

Mucous membrane cells

Centres of cells in mucous membrane:



Repulsion due to physical extent of cells

Inhomogeneity - lower intensity in upper part

Bivariate - two types of cells

Same type of inhomogeneity for two types ?

Density with respect to a Poisson process

$\mathbf{X}$ on bounded S has density f with respect to unit rate Poisson $\mathbf{Y}$ if

$$P(\mathbf{X} \in F) = \mathbb{E}\big(1[\mathbf{Y} \in F] f(\mathbf{Y})\big)$$
$$= \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-|S|}}{n!} \int_{S^n} 1[\mathbf{x} \in F] f(\mathbf{x}) \mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (\mathbf{x} = \{x_1, \ldots, x_n\})$$

Example: Strauss process

For a point configuration $\mathbf{x}$ on a bounded region S, let $n(\mathbf{x})$ and $s(\mathbf{x})$ denote the number of points and number of (unordered) pairs of R-close points ($R \geq 0$).

A *Strauss process* $\mathbf{X}$ on S has density

$$f(\mathbf{x}) = \frac{1}{c} \exp(\beta n(\mathbf{x}) + \psi s(\mathbf{x}))$$

with respect to a unit rate Poisson process $\mathbf{Y}$ on S and

$$c = \mathbb{E} \exp(\beta n(\mathbf{Y}) + \psi s(\mathbf{Y})) \tag{4}$$

is the normalizing constant (unknown).

Note: only well-defined ($c < \infty$) if $\psi \leq 0$.

Intensity and conditional intensity

Suppose $\mathbf{X}$ has *hereditary* density f with respect to Y: $f(\mathbf{x}) > 0 \Rightarrow f(\mathbf{y}) > 0, \mathbf{y} \subset \mathbf{x}$.

Intensity function $\rho(u) = \mathbb{E} f(\mathbf{Y} \cup \{u\})$ usually unknown (except for Poisson and Cox/Cluster).

Instead consider *conditional intensity*

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})}$$

(does not depend on normalizing constant !)

Note

$$\rho(u) = \mathbb{E} f(\mathbf{Y} \cup \{u\}) = \mathbb{E}\big[\lambda(u, \mathbf{Y}) f(\mathbf{Y})\big] = \mathbb{E}\lambda(u, \mathbf{X})$$

and

$\rho(u)\mathrm{d}A \approx P(\mathbf{X} \text{ has a point in } A) = \mathbb{E}P(\mathbf{X} \text{ has a point in } A | \mathbf{X} \setminus A), u \in A$

Hence, $\lambda(u, \mathbf{X})\mathrm{d}A$ probability that $\mathbf{X}$ has point in very small region A given $\mathbf{X}$ outside A.

Density and conditional intensity - factorization

One-to-one correspondence between density and conditional intensity (up to normalizing constant)

$$f(\{x_1, \ldots, x_n\}) \propto h(\{x_1, \ldots, x_n\}) = \prod_{i=1}^{n} \lambda(x_i, \{x_1, \ldots, x_{i-1})$$

Normalizing constant:

$$f(\mathbf{x}) = \frac{1}{c} h(\mathbf{x}) \quad c = \mathbb{E} h(\mathbf{Z})$$

Typically c is intractable so likelihood estimation is not straightforward.

Options: pseudo-likelihood (later in this section) or Monte Carlo approximation of c.

Markov point processes

Def: suppose that f hereditary and $\lambda(u, \mathbf{x})$ only depends on $\mathbf{x}$ through $\mathbf{x} \cap b(u, R)$ for some $R > 0$ (*local Markov property*). Then f is *Markov* with respect to the R-close neighbourhood relation.

Thm (Hammersley-Clifford) The following are equivalent.

1. f is Markov.
2. $$f(\mathbf{x}) = \prod_{\mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y})$$

   where $\phi(\mathbf{y}) = 1$ whenever $\|u - v\| \geq R$ for some $u, v \in \mathbf{y}$.

Pairwise interaction process: $\phi(\mathbf{y}) = 1$ whenever $n(\mathbf{y}) > 2$.

NB: in H-C, R-close neighbourhood relation can be replaced by an arbitrary symmetric relation between pairs of points.

Modelling the conditional intensity function

Suppose we specify a model for the conditional intensity. Two questions:

1. does there exist a density f with the specified conditional intensity ?
2. is f well-defined (integrable) ?

Solution:

1. find f by identifying interaction potentials (Hammersley-Clifford) or guess f.

2. sufficient condition (local stability): $\lambda(u, \mathbf{x}) \leq K$

NB some Markov point processes have interactions of any order in which case H-C theorem is less useful (e.g. area-interaction process).

Some examples

Strauss (pairwise interaction):

$$\lambda(u,\mathbf{x}) = \exp\left(\beta+\psi\sum_{v\in\mathbf{x}} 1[\|u-v\| \le R]\right), \quad f(\mathbf{x}) = \frac{1}{c}\exp\left(\beta n(\mathbf{x})+\psi s(\mathbf{x})\right)$$

Overlap process (pairwise interaction marked point process):

$$\lambda((u,m),\mathbf{x}) = \frac{1}{c}\exp\left(\beta+\psi\sum_{(u',m')\in\mathbf{x}} |b(u,m)\cap b(u',m')|\right) \quad (\psi \le 0)$$

where $\mathbf{x} = \{(u_1,m_1),\ldots,(u_n,m_n)\}$ and $(u_i,m_i) \in \mathbb{R}^2 \times [a,b]$.

Area-interaction process:

$$f(\mathbf{x}) = \frac{1}{c}\exp\left(\beta n(\mathbf{x})+\psi V(\mathbf{x})\right), \quad \lambda(u,\mathbf{x}) = \exp\left(\beta+\psi(V(\{u\}\cup\mathbf{x})-V(\mathbf{x})\right)$$

$V(\mathbf{x}) = |\cup_{u\in\mathbf{x}} b(u,R/2)|$ is area of union of balls $b(u,R/2)$, $u \in \mathbf{x}$.

NB: $\phi(\cdot)$ complicated for area-interaction process.

The Georgii-Nguyen-Zessin formula ('Law of total probability')

$$\mathbb{E}\sum_{u\in\mathbf{X}} k(u,\mathbf{X}\backslash\{u\}) = \int_S \mathbb{E}[\lambda(u,\mathbf{X})k(u,\mathbf{X})]\,\mathrm{d}u = \int_S \mathbb{E}^![k(u,\mathbf{X})\,|\,u]\rho(u)\,\mathrm{d}u$$

$\mathbb{E}^![\cdot\,|\,u]$: expectation with respect to the conditional distribution of $\mathbf{X}\setminus\{u\}$ given $u\in\mathbf{X}$ (*reduced Palm distribution*)

Density of reduced Palm distribution:

$$f(\mathbf{x}\,|\,u) = f(\mathbf{x}\cup\{u\})/\rho(u)$$

NB: GNZ formula holds in general setting for point process on $\mathbb{R}^d$.

Statistical inference based on pseudo-likelihood

$\mathbf{x}$ observed within bounded S. Parametric model $\lambda_\theta(u, \mathbf{x})$.

Let $N_i = 1[\mathbf{x} \cap C_i \neq \emptyset]$ where C_i disjoint partitioning of $S = \cup_i C_i$.

$P(N_i = 1 \,|\, \mathbf{X} \setminus C_i) \approx \lambda_\theta(u_i, \mathbf{X} \setminus C_i)\mathrm{d}C_i$ where $u_i \in C_i$.

Hence composite likelihood based on the N_i:

$$\prod_{i=1}^{n} (\lambda_\theta(u_i, \mathbf{x} \setminus C_i)\mathrm{d}C_i)^{N_i}(1 - \lambda_\theta(u_i, \mathbf{x} \setminus C_i)\mathrm{d}C_i)^{1-N_i} \equiv$$

$$\prod_{i=1}^{n} \lambda_\theta(u_i, \mathbf{x} \setminus C_i)^{N_i}(1 - \lambda_\theta(u_i, \mathbf{x} \setminus C_i)\mathrm{d}C_i)^{1-N_i}$$

which tends to *pseudo-likelihood* function

$$\prod_{u \in \mathbf{x}} \lambda_\theta(u, \mathbf{x} \setminus \{u\}) \exp\big(- \int_S \lambda_\theta(u, \mathbf{x})\mathrm{d}u \big)$$

Score of pseudo-likelihood: unbiased estimating function by GNZ.

Pseudo-likelihood estimates asymptotically normal but asymptotic variance is not straightforward.

Integral approximated by numerical quadrature or Monte Carlo

Flexible implementation for log linear conditional intensity (fixed R) in `spatstat`

Estimation of interaction range R: profile likelihood (?)

Exercises

1. Suppose that S contains a disc of radius $\epsilon \leq R/2$. Show that (4) is not finite, and hence the Strauss process not well-defined, when ψ is positive.

   (Hint: $\sum_{n=0}^{\infty} \frac{(\pi\epsilon^2)^n}{n!} \exp(n\beta + \psi n(n-1)/2) = \infty$ if $\psi > 0$.)
2. Show that local stability for a spatial point process density ensures integrability. Verify that the area-interaction process is locally stable.
3. what is the unnormalized density of a Strauss (β, ψ) with respect to a Poisson process of intensity $\exp(\beta)$?
4. Starting with the conditional intensity for a Strauss process, identify the potential function ϕ
5. (if time) Verify the Georgii-Nguyen-Zessin formula for a finite point process.

   (Hint: consider first the case of a finite Poisson-process $\mathbf{Y}$ in which case the identity is known as the Slivnyak-Mecke theorem, next apply $\mathbb{E}g(\mathbf{X}) = \mathbb{E}\big[g(\mathbf{Y})f(\mathbf{Y})\big]$.)

Monte Carlo approximation

Let **D** 'quadrature/dummy' point process of intensity $\rho(\cdot)$ and independent of **X**. $\mathbf{X} \cup \mathbf{D}$ has conditional intensity $\lambda(u, \mathbf{X}) + \rho(u)$

By GNZ

$$\mathbb{E} \int_W \lambda'(u, \mathbf{X}) \mathrm{d}u = \mathbb{E} \sum_{u \in \mathbf{X} \cup \mathbf{D}} \frac{\lambda'(u, \mathbf{X} \setminus \{u\})}{\lambda(u, \mathbf{X} \setminus \{u\}) + \rho(u)}$$

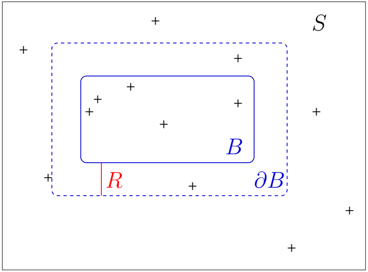
Idea: replace integral in pseudo-likelihood with unbiased estimates using **D**.

Resulting estimating function formally equivalent to logistic regression

The spatial Markov property and edge correction

Let $B \subset S$ and assume $\mathbf{X}$ Markov with interaction radius R.

Define: ∂B points in $S \setminus B$ of distance less than R from B



Factorization (Hammersley-Clifford):

$$f(\mathbf{x}) = \prod_{\substack{\mathbf{y} \subseteq \mathbf{x} \cap (B \cup \partial B): \\ \mathbf{y} \cap B \neq \emptyset}} \phi(\mathbf{y}) \prod_{\mathbf{y} \subseteq \mathbf{x} \setminus B} \phi(\mathbf{y})$$

Hence, conditional density of $\mathbf{X} \cap B$ given $\mathbf{X} \setminus B$

$$f_B(\mathbf{z}|\mathbf{y}) \propto f(\mathbf{z} \cup \mathbf{y})$$

depends on $\mathbf{y}$ only through $\partial B \cap \mathbf{y}$.

Edge correction using the border method

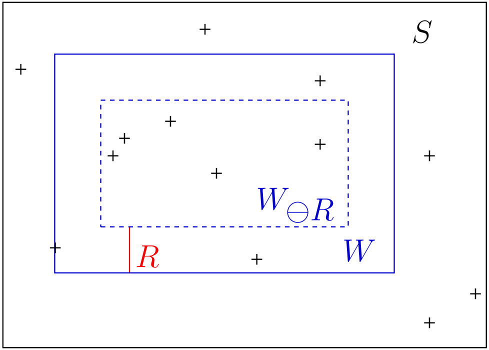
Suppose we observe $\mathbf{x}$ realization of $\mathbf{X} \cap W$ where $W \subset S$.

Problem: density (likelihood) $f_W(\mathbf{x}) = \mathbb{E} f(\mathbf{x} \cup Y_{S \setminus W})$ unknown.

Border method: base inference on

$$f_{W_{\ominus R}}(\mathbf{x} \cap W_{\ominus R} | \mathbf{x} \cap (W \setminus W_{\ominus R}))$$

i.e. conditional density of $\mathbf{X} \cap W_{\ominus R}$ given $\mathbf{X}$ outside $W_{\ominus R}$.



Solution: second order product density for Poisson

$$
\begin{aligned}
&\mathbb{E} \sum_{u,v \in \mathbf{X}}^{\neq} 1[u \in A, v \in B] \\
=& \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu(A \cup B)}}{n!} \int_{(A \cup B)^n} \sum_{u,v \in \{x_1,\dots,x_n\}}^{\neq} 1[u \in A, v \in B] \prod_{i=1}^{n} \rho(x_i) \mathrm{d}x_1 \dots \mathrm{d}x_n \\
=& \sum_{n=2}^{\infty} \frac{\mathrm{e}^{-\mu(A \cup B)}}{n!} 2 \binom{n}{2} \int_{(A \cup B)^2} \int_{(A \cup B)^{n-2}} 1[x_1 \in A, x_2 \in B] \prod_{i=1}^{n} \rho(x_i) \mathrm{d}x_1 \dots \mathrm{d}x
\\
=& \sum_{n=2}^{\infty} \frac{\mathrm{e}^{-\mu(A \cup B)}}{(n-2)!} \mu(A)\mu(B)\mu(A \cup B)^{n-2} \\
=& \mu(A)\mu(B) = \int_{A \times B} \rho(u)\rho(v) \mathrm{d}u \mathrm{d}v
\end{aligned}
$$

Solution: invariance of g (and K) under thinning

Since $\mathbf{X}_{\text{thin}} = \{u \in \mathbf{X} : R(u) \leq \pi(u)\}$,

$$
\begin{aligned}
&\mathbb{E} \sum_{u,v \in \mathbf{X}_{\text{thin}}}^{\neq} 1[u \in A, v \in B] \\
=&\mathbb{E} \sum_{u,v \in \mathbf{X}}^{\neq} 1[R(u) \leq \pi(u), R(v) \leq \pi(v), u \in A, v \in B] \\
=&\mathbb{E}\,\mathbb{E}\Big[\sum_{u,v \in \mathbf{X}}^{\neq} 1[R(u) \leq \pi(u), R(v) \leq \pi(v), u \in A, v \in B] \,\Big|\, \mathbf{X} \Big] \\
=&\mathbb{E} \sum_{u,v \in \mathbf{X}}^{\neq} \pi(u)\pi(v)1[u \in A, v \in B] \\
=&\int_A \int_B \pi(u)\pi(v)\rho^{(2)}(u,v)\mathrm{d}u\mathrm{d}v
\end{aligned}
$$

References

1. Baddeley, A.J., Coeurjolly, J.F., Rubak, E. and Waagepetersen R. (2014) A logistic regression estimating function for Gibbs point processes, *Biometrika*, **101**, 377-392.
2. Guan, Y., Jalilian, A. and Waagepetersen, R. (2015) Quasi-likelihood for spatial point processes, *Journal of the Royal Statistical Society, Series B*, **77**, 677-697.
3. Jalilian, A., Guan, Y. and Waagepetersen, R. (2013) Decomposition of variance for spatial Cox processes, *Scandinavian Journal of Statistics*, **40**, 119-137.
4. Møller, J. and Waagepetersen, R. (2003) Statistical inference and simulation for spatial point processes, Chapman and Hall/CRC Press.
5. Møller, J. and Waagepetersen, R. (2007) Modern statistics for spatial point processes, *Scandinavian Journal of Statistics*, **34**, 643-684.
6. Møller, J. and Waagepetersen, R. (2016) Some recent developments in statistics for spatial point processes, *Annual review of Statistics and its Applications*, to appear, arXiv: 1609.00908.
7. Waagepetersen, R. (2007) An estimating function approach to inference for inhomogeneous Neyman-Scott processes with Appendix, *Biometrics*, **63**, 252-258.
8. Waagepetersen, R. (2008) Estimating functions for inhomogeneous spatial point processes with incomplete covariate data, *Biometrika*, **95**, 351-363.
9. Waagepetersen, R. and Guan, Y. (2009) Two-step estimation for inhomogeneous spatial point processes, *Journal of the Royal Statistical Society, Series B*, **71**, 685-702.

(see the monograph M & W '03, and the two review papers, M & W '07, '16, for further references)