Part 2: Point Processes

Last week:
- Examples
- Definitions + 3.4.1
- More examples and 3.4.2 Spatial statistical dependency

This week:
- Simulations + more models (3.4.3 + revisit ++).
- Estimation

Extra text book: Statistical Analysis and Modelling of Spatial Point Patterns by Janine Illian, Antti Penttinen, Helga Stoyan, Dietrich Stoyan (2008)

TODAY:

Clustering
- Neymann-Scott processes
- Cox processes
- Log-Gaussian Cox process

Regularization
- Gibbs processes

From Cressie and Wikle:

4.3 SPATIAL POINT PROCESSES

A spatial point process is a stochastic process governing the location of events (equivalently, points) $\{\mathbf{s}_i\}$ in some set $D_s \subset \mathbb{R}^d$, where the number of such events in D_s is also random (e.g., Diggle, 2003). In the simplest case, the

and we only consider:

Only *simple* spatial point processes in $\mathbb{R}^d$ (i.e., almost surely, either no event or a single event occurs at any point) will be considered. We characterize the

Data

Number of points: $Z(D_s) = m$

Locations: $\{s_1, s_2, \ldots, s_m\}$

from Illian et al 2008:

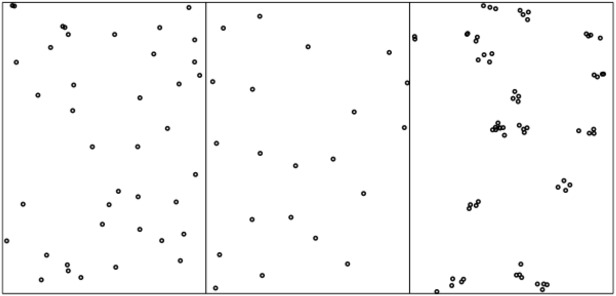


Figure 1.1 Three simulated point patterns: (left) random, (centre) regular, (right) clustered.

An important, but simple, example is the *Poisson point process* Z, for which

$$Z(A)|\lambda^o \sim Poi(\lambda^o|A|), \qquad A \subset D_s\,, \tag{4.159}$$

where $\lambda^o > 0$ is a parameter of the Poisson point process, and recall that $|A|$ is the d-dimensional volume of A. More details on this spatial point process are given in Section 4.3.1.

Sampling from Poisson point process

Trees in $1km \times 1km$ domain, with intensity $\lambda^0 = 15 trees/km^2$

- Sample $m \sim Pos(15)$
- for $i = 1 : m$
    - sample location randomly in domain $\{s_i\}$
- end

Clustering, how to sample/model?

- ...
- ...

Hierachical statistical models (HM), and Cox process

Data model: $[Z|Y,\theta]$

Process model: $[Y|\theta]$

Parameter model: $[\theta]$

Cox process

Data model: $[Z|\lambda] \sim Pos(\int_A \lambda(d)dx)$

Process model: $\lambda(s)$ (Cox process)

Overview clustering processes

from Note

Name of Process	Clusters	Parents
Independent cluster process	general	general
Poisson cluster process	general	Poisson
Cox cluster process	Poisson	general
Neyman-Scott process	Poisson	Poisson
Matérn cluster process	Poisson (uniform in ball)	Poisson (homog.)
Modified Thomas Process	Poisson (Gaussian)	Poisson (homog.)

Table 2: Nomenclature for independent cluster processes.

from Note

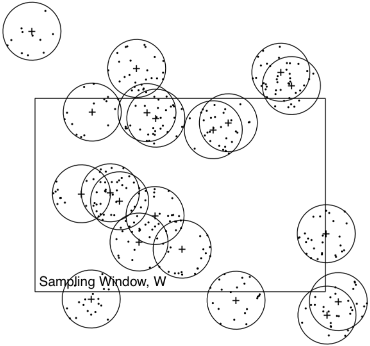


Figure 1: Edge effects for a Matérn process. Parent process is homogeneous Poisson with a constant rate of 20 in a 20×20 square field. The cluster processes are homogeneous Poisson centered at the parents (+) with a radius of 3 and a constant rate of 10.

First order intensity

ds denote a small region located at $\mathbf{s}$ with volume $|d\mathbf{s}|$. Then the *first-order intensity* function of the point process $Z(\cdot)$ is defined as

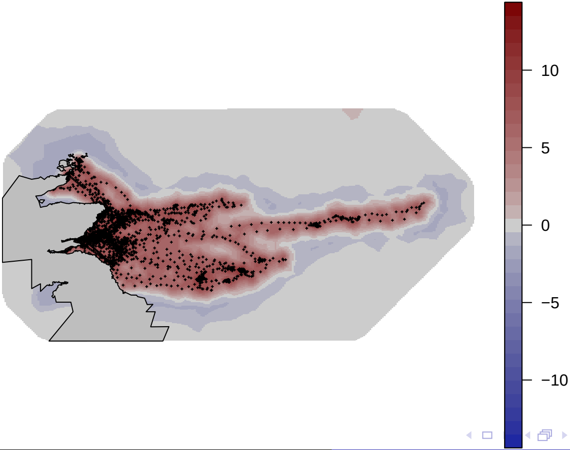
$$\lambda(\mathbf{s}) \equiv \lim_{|d\mathbf{s}| \to 0} E(Z(d\mathbf{s}))/|d\mathbf{s}|, \qquad \mathbf{s} \in D_s, \qquad (4.160)$$

provided the limit exists. Hence,

$$E(Z(A)) = \int_A \lambda(\mathbf{s})\, d\mathbf{s}, \qquad A \subset D_s. \qquad (4.161)$$

Estimation of intensity

We often want to estimate the $\lambda(s)$ from data (points).
Seals (from log-Gaussian Cox model):



Kernel estimation of intensity

SPATIAL POINT PROCESSES 207

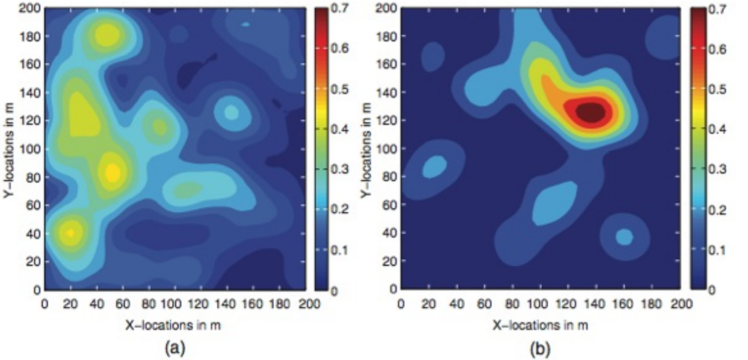


Figure 4.16 **(a)** Contour plot of estimated first-order intensity of $m_A = 271$ adult longleaf pine trees in the 4-ha study area in the Wade Tract; bandwidth $b = 30$ m. **(b)** Contour plot of estimated first-order intensity of $m_S = 159$ subadult longleaf pine trees in the same study area as **(a)**; bandwidth $b = 30$ m.

2nd order intensity and Pair-correlation function

2nd order intensity

$$\lambda_2(s,x) = \lim_{|ds|\to 0, |dx|\to 0} \frac{E(Z(ds)Z(dx))}{|ds||dx|}$$

Pair correlation function

$$g(s,x) = \frac{\lambda_2(x,s)}{\lambda(s)\lambda(x)}$$

For HPPP: $\lambda_2(s,x) = \lambda(s)\lambda(x)$ and $g(s,x) = 1$

Hierachical statistical models (HM), LGCP

Data model: $[Z|Y, \theta]$

Process model: $[Y|\theta]$

Parameter model: $[\theta]$

Log Gaussian Cox process (LGCP)

Data model: Conditional on $\lambda(\cdot)$, the point process $Z(\cdot)$ is an inhomogeneous Poisson point process with intensity $\lambda(\cdot)$.

Process model: Conditional on $\boldsymbol{\beta}$ and $C_Y(\cdot,\cdot)$, $Y(\cdot) \equiv \log \lambda(\cdot)$ is a Gaussian process with the following properties:

$$E(Y(\mathbf{s})) = \mathbf{x}(\mathbf{s})'\boldsymbol{\beta} \quad \text{and} \quad C_Y(\mathbf{s}, \mathbf{x}) \equiv \text{cov}(Y(\mathbf{s}), Y(\mathbf{x})) .$$

$\lambda_2(s, u)$ for Log Gaussian Cox Process (LGCP)

From page 214 Cressie & Wikle.

$$
\begin{aligned}
E(Z(d\mathbf{s})Z(d\mathbf{x})) &\overset{1}{=} E\{E(Z(d\mathbf{s})Z(d\mathbf{x})|\lambda(\cdot))\} \\
&= E\{\lambda(d\mathbf{s})\lambda(d\mathbf{x})\} && \text{2} \\
&= E(e^{Y(\mathbf{s})} \cdot e^{Y(\mathbf{x})})|d\mathbf{s}|\,|d\mathbf{x}|\,, && \text{3} \qquad (4.182)
\end{aligned}
$$

$$\lambda_2(\mathbf{s}, \mathbf{x}) = \exp\{\mu_Y(\mathbf{s}) + \mu_Y(\mathbf{x}) + C_Y(\mathbf{s}, \mathbf{s})/2 + C_Y(\mathbf{x}, \mathbf{x})/2 + C_Y(\mathbf{s}, \mathbf{x})\}\,. \qquad (4.183)$$

Intensity functions

1 order intensity:

$$\lambda_1(s) = \lim_{|ds|\to 0} \frac{E(Z(ds))}{|ds|}$$

2. order intensity:

$$\lambda_2(s_1, s_2) = \lim_{|ds_1|\to 0, |ds_2|\to 0} \frac{E(Z(ds_1)Z(ds_2))}{|ds_1||ds_2|}$$

...

mth order intersity

$$\lambda_m(s_1, s_2) = \lim_{\cdots\to 0} \frac{E(Z(ds_1)Z(ds_2)\ldots Z(ds_m))}{|ds_1||ds_2|\ldots|ds_m|}$$

Example Strauss

from Note. $\lambda^0 = \beta = 100$

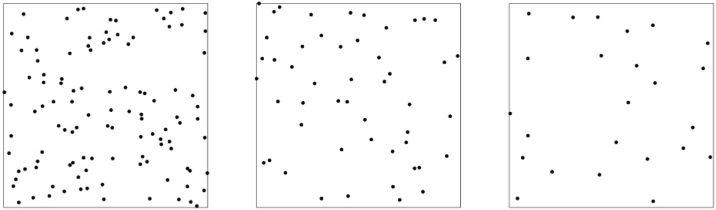


Table 3: Realizations of Strauss processes on $S = [0, 1]^2$. Here $\beta = 100$, $R = 0.1$ and $\gamma = 1.0, 0.5, 0.0$ from left to right. Processes generated from the function rStrauss in the R package spatstat by Adrian Baddeley.