Spatial point processes: introduction

Aila Särkkä

Mathematical sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden

1. Introduction
2. Repetitions, inhomogeneity, anisotropy, 2D→3D
3. Example 1: Analysis and modeling of epidermal nerve fiber patterns
   Repetitions and non-spatial covariates
4. Example 2: Air pore structure in polar ice
   Anisotropy (3D) and missing information

Illian *et al.* (2008)
R library spatstat used (Baddeley and Turner, 2005).

Definition

A point process N is a stochastic mechanism or rule to produce point patterns or realisations according to the distribution of the process.

A marked point process is a point process where each point x_i of the process is assigned a quantity $m(x_i)$, called a mark. Often, marks are integers or real numbers but more general marks can also be considered.

Two interpretations

- N is a counting measure. For a subset B of $\mathbb{R}^d$, $N(B)$ is the random number of points in B. It is assumed that $N(B) < \infty$ for all bounded sets B, i.e. that N is locally finate.
- N is a random set, i.e. the set of all points $x_1, x_2, \ldots$ in the process. In other words

$$N = \{x_i\} \text{ or } N = \{x_1, x_2, \ldots\}$$

Therefore, $x \in N$ means that the point x is in the set N. The set N can be finite or infinite. If it is finate the total number of points can be deterministic or random.

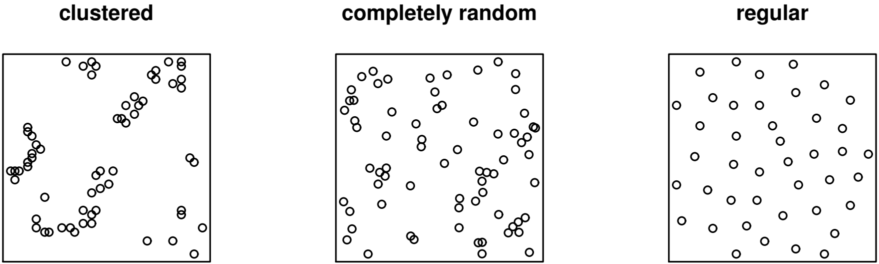
Remarks

Remark 1: We assume that all point processes are simple, i.e. that there are no multiple points ($x_i \neq x_j$ if $i \neq j$).

Remark 2: There is a large literature on processes $\{Z(t) : t \in T\}$, where T is a point process in time. There is an overlap of methods for point processes in space and in time but the temporal case is not only a special case of the spatial process with $d = 1$. Time is 1-directional.

Remark 3: To avoid confusion between points of the process and point of $\mathbb{R}^d$, the points of the process or point pattern (realization) are called events (or trees or cells).

Spatial point patterns

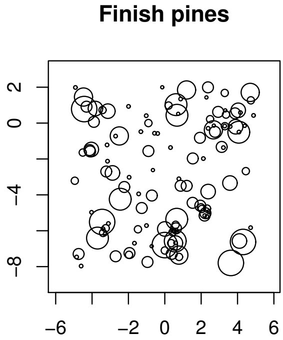


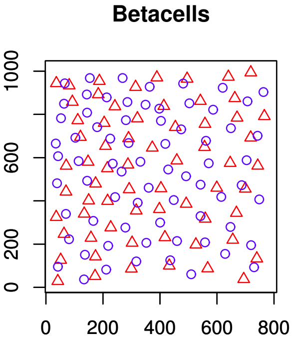
Examples

- Locations of betacells within a rectangular region in a cat's eye (regular)
- Locations of Finnish pine saplings (clustered)
- Locations of Spanish towns (regular)
- Locations of galaxes (clustered)

Remark: Very different scales, from microscopic to cosmic

Marked point patterns





Finnish pine saplings: locations and diameters

Beta-type retina cells in the retina of a cat: locations and type (red triangles "on", blue circles "off")

First-order properties (without marks)

The mean number of points of N in B is $\mathbb{E}(N(B))$ (depends on the set B). We use the notation

$$\Lambda(B) = \mathbb{E}(N(B))$$

and call Λ the intensity measure.

Under some continuity conditions, a density function λ, called the intensity function, exists, and

$$\Lambda(B) = \int_B \lambda(x)\,dx.$$

Some properties of point processes: stationarity and isotropy

A point process N is stationary (translation invariant) if N and the translated point process N_x have the same distribution for all translations x, i.e.

$$N = \{x_1, x_2, ...\} \text{ and } N_x = \{x_1 + x, x_2 + x, ...\}$$

have the same distribution for all $x \in \mathbb{R}^d$.

A point process is isotropic (rotation invariant) if its characteristics are invariant under rotations, i.e.

$$N = \{x_1, x_2, ...\} \text{ and } rN_x = \{rx_1, rx_2, ...\}$$

have the same distribution for any rotation r around the origin.
If a point process is both stationary and isotropic, it is called motion-invariant.

First-order properties

If N is stationary, then

$$\Lambda(B) = \lambda|B|,$$

where $0 < \lambda < \infty$ is called the intensity of N and $|B|$ is the volume of B.

λ is the mean number of points of N per unit area, i.e.

$$\lambda = \frac{\Lambda(B)}{|B|} = \frac{\mathbb{E}(N(B))}{|B|}.$$

Two distribution functions

1. Let D_1 denote the distance from an arbitrary event to the nearest other event. Then, the nearest neighbour distance function is

$$G(r) = P(D_1 \leq r)$$

   If the pattern is completely spatially random (CSR), $G(r) = 1 - \exp(-\lambda\pi r^2)$. For regular patterns $G(r)$ tends to lie below and for clustered patterns above the CSR curve.

2. Let D_2 denote the distance from an arbitrary point to the nearest event. Then,

$$F(r) = P(D_2 \leq r)$$

   If the pattern is completely spatially random, $F(r) = 1 - \exp(-\lambda\pi r^2)$. For regular patterns $F(r)$ tends to lie above and for clustered patterns below the CSR curve.

Combination of the two

Using G and F we can define the so-called J function as

$$J(r) = \frac{1 - G(r)}{1 - F(r)}$$

(whenever $F(r) > 0$)

If the pattern is completely spatially random, $J(r) \equiv 1$. For regular patterns $J(r) > 1$ and for clustered patterns $J(r) < 1$.

Second-order properties

The 2nd order properties of a stationary and isotropic point process can be characterized by Ripley's K function (Ripley, 1977)

$K(r) = \lambda^{-1}\mathbb{E}[\#$ further events within distance r of a typical event].

Often, (in 2D) a variance stabilizing and centered version of the K function (Besag, 1977) is used, namely

$$L(r) - r = \sqrt{K(r)/\pi} - r,$$

which equals 0 under CSR. Values less than zero indicate regularity and values larger than zero clustering.

Note on estimation of G, F, J and K

- Typically, a point pattern is observed in a (bounded) observation window and points outside the window are not observed.
- Estimators (except for $J(r)$) need to be edge-corrected
- Edge correction methods include plus sampling, minus sampling, Ripley's isotropic correction and translation (stationary) correction

Poisson process

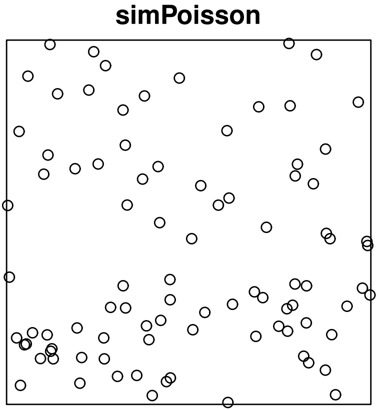
A point process is a homogeneous Poisson process (CSR) if

(P1) for some $\lambda > 0$ and any finite region B, $N(B)$ has a Poisson distribution with mean $\lambda|B|$

(P2) given $N(B) = n$, the events in B form an independent random sample from the uniform distribution on B

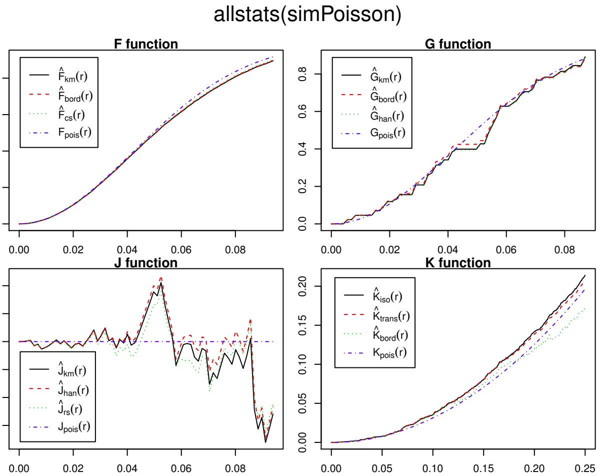
Inhomogeneous Poisson process: intensity λ (in homogeneous Poisson process) replaced by an intensity function $\lambda(x)$

Realization of a homogeneous Poisson process



Poisson process with intensity 100

Summary statistics



Matérn cluster process

Cluster processes are models for aggregated spatial point patterns

For Matérn cluster process

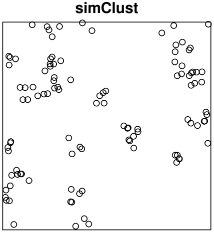
(MC1) parent events form a Poisson process with intensity λ

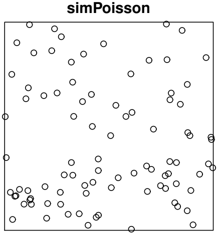
(MC2) each parent produces a random number S of daughters (offsprings), realized independently and identically for each parent according to some probability distribution

(MC3) the locations of the daughters in a cluster are independently and uniformly scattered in the disc of radius R centered at the parent point.

The cluster process consists only of the daughter points.

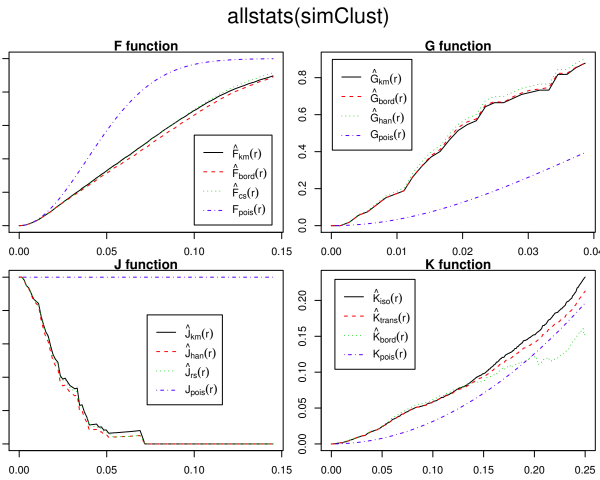
Realization of a Matérn cluster process





Left: Parent point intensity 20, cluster radius 0.05, average number of daughter points per cluster 5
Right: Poisson process with intensity 100

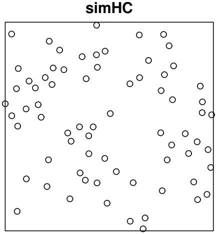
Summary statistics

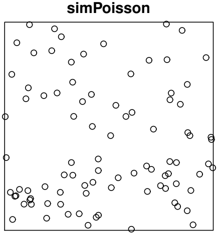


Matérn I hard-core process

- Hard-core processes are models for regular spatial point patterns
- There is a minimum allowed distance, called hard-core distance, between any two points
- Matérn I hard-core process: A Poisson process with intensity λ is thinned by delating all pairs of points that are at distance less than the hard-core radius apart.

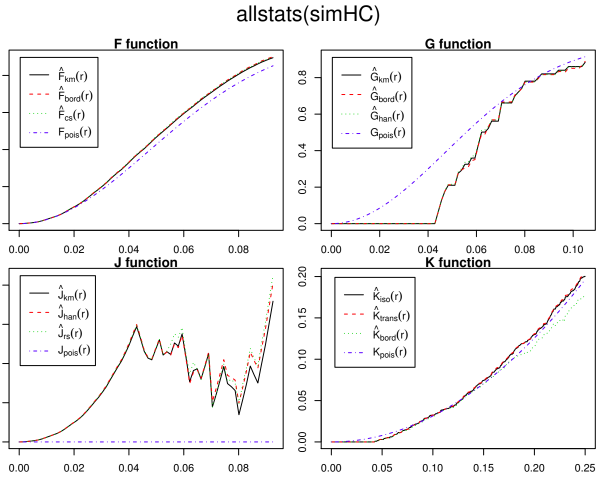
Realization of a Matérn I hard-core process





Left: Hard-core process with the initial Poisson intensity 300, hard-core radius 0.04
Right: Poisson process with intensity 100

Summary statistics



Pairwise interaction processes

- Pairwise interaction processes are a subclass of Markov point processes which are models for point patterns with interaction between the events
- There is interaction between the events if they are "neighbours", e.g. it they are close enough to each other
- Models for inhibition/regularity

Pairwise interaction procesees: Strauss process

- Two points are neighbours if they are closer than distance R apart
- The density function (with respect to a Poisson process with intensity 1) is

$$f(x) = \alpha\beta^{n(x)}\gamma^{s(x)}, \ \beta > 0, \ \gamma \geq 0,$$

  where
  - $\beta > 0$ is the effect of a single event (connected to the intensity of the process)
  - $0 < \gamma \leq 1$ is an interaction parameter
  - $n(x)$ is the number of points in the configuration
  - $s(x)$ is the number of R close pairs in the configuration, where $R > 0$ is an interaction radius (range of interaction)
  - α is a normalizing constant

References

- Baddeley, A.,Turner, R. Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* 12 (2005) 1-42.
- Besag, J.E., 1977. Comment on “Modelling spatial patterns” by B. D. Ripley. *Journal of the Royal Statistical Society* B (Methodological) 39 (1977) 193-195.
- Illian, J., Penttinen, A., Stoyan, H., Stoyan, D. *Statistical Analysis and Modelling of Spatial Point Patterns*. Chichester: Wiley (2008).
- Ripley, B.D. Modelling spatial patterns. *Journal of the Royal Statistical Society* B 39 (1977) 172-212.