Spatial statistics and image analysis (TMS016/MSA301)

Marked point processes, point processes with noise

2022-05-02

Practical information

We had a (midterm) meeting with the students' representatives, Ivan Flensburg, Henrik Häggström, and Viktoria Löfgren. Some comments

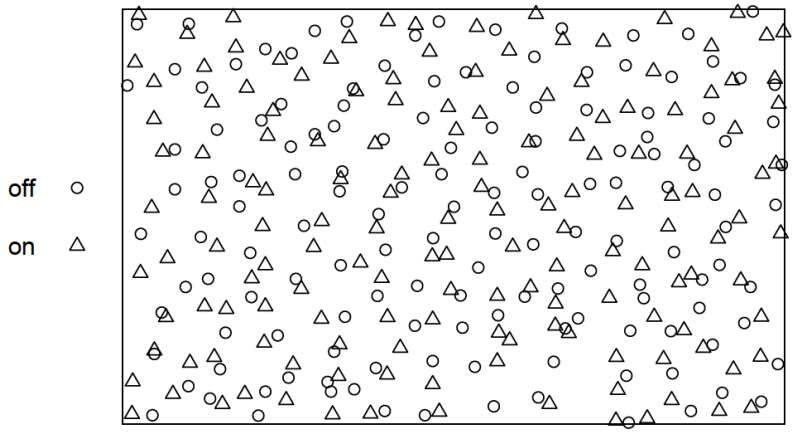
- Course literature could be structured better.
  $\rightarrow$ There is a folder called "Course literature" under "Files" in Canvas now. The list of books and the book chapters by Glasbye and Horgan are there.
- Hard to find information on the first two project parts.
  $\rightarrow$ We will keep this in mind for next year.
- Question: Is the schedule for the projects reasonable? Please, let us know if something should be changed.

Marked point process

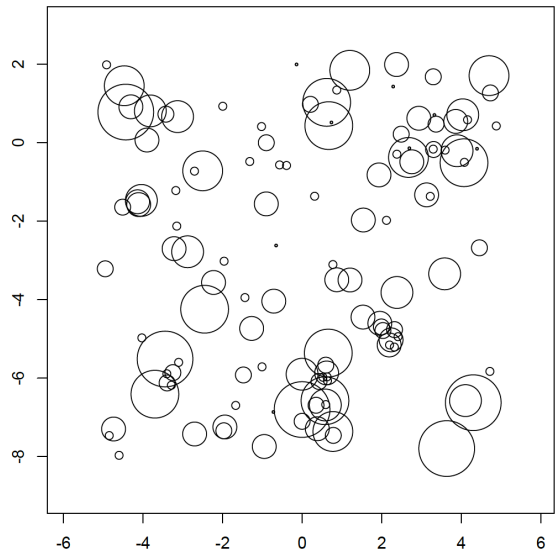
- Generalizations of (unmarked) point processes, where each event (point of the process) x_i is assigned a further quantity $m(x_i)$ (called a mark).
- Often, the marks are integers or real numbers but much more general marks can be considered.
- A marked point process is denoted by M

Marked point patterns

Amacrine cells: a point pattern with two types of points.



Finnish pines: a point pattern with continues marks (diameters).



Stationarity and isotropy

If $M = \{[x_1; m(x_1)], [x_2; m(x_2)], ...\}$ then its translated process is

$$M_x = \{[x_1 + x; m(x_1)], [x_2 + x; m(x_2)], ...\}$$

Note that in the translated process the marks stay the same and only the points are translated.

A marked point process M is stationary if and only if M and M_x have the same distribution.

The definition of isotropy is analogous.

First order characteristics

The intensity function $\lambda(x, m)$ gives the mean behaviour of the marked point process. The mean number of marked points of M located in the set B and having marks in C is given by

$$\mathbb{E}(M(B \times C)) = \int_B \int_C \lambda(x, m)\, dm\, dx,$$

where $M(B \times C)$ denotes the number of events in B that have marks in C.

In the stationary case we can use two types of first-order characteristics, one that concerns the points (intensity λ), and another that describes the marks (mark probabilities p_i for qualitative marks and density function for quantitative marks).

The intensity λ is the same intensity as in the unmarked case: the mean number of events per unit area (volume)

Second-order characteristics: qualitative marks

From now on, we assume stationarity and isotropy.

Let N_1 and N_2 be two sub point processes with mark 1 and intensity λ_1 and mark 2 and intensity λ_2, respectively.

Ripley's cross K function K_{12} is defined as

$$\lambda_2 K_{12}(r) = \mathbb{E}_{o1}(N_2(b(o,r))) \text{ for } r \geq 0,$$

where $\mathbb{E}_{o1}$ is the conditional expectation given that there is an event of type 1 at o.

Therefore, $\lambda_2 K_{12}(r)$ gives the mean number of events of type 2 in a disc of radius r centered at an arbitrary event of type 1.

Can be estimated in a similar way as the "usual" Ripley's K function (including edge corrections).

Remarks

- K_{11} and K_{22} are the "usual" K functions for the types 1 and 2, respectively.
- $K_{12}(r) = K_{21}(r)$ for $r \geq 0$. (Remark! Estimators may not coincide completely.)
- Transformation to L function can be done as in the univariate case.

Special situations

- If the two sub processes are independent (here, in $2D$), i.e. we have a superposition of two processes, then

$$K_{12}(r) = \pi r^2 \text{ for } i \neq j$$

(compare to the completely spatially random case). However, that the equation above holds does not guarantee independence between the two processes.

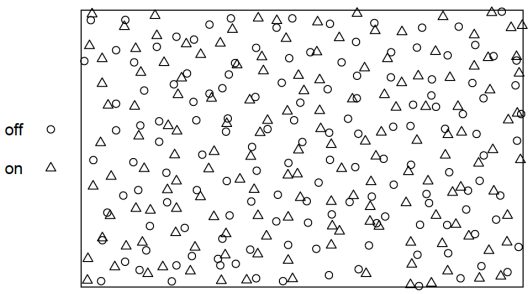
- If the marking can be regarded as random labelling (i.e. the sub processes are independent random thinnings of the entire process), then

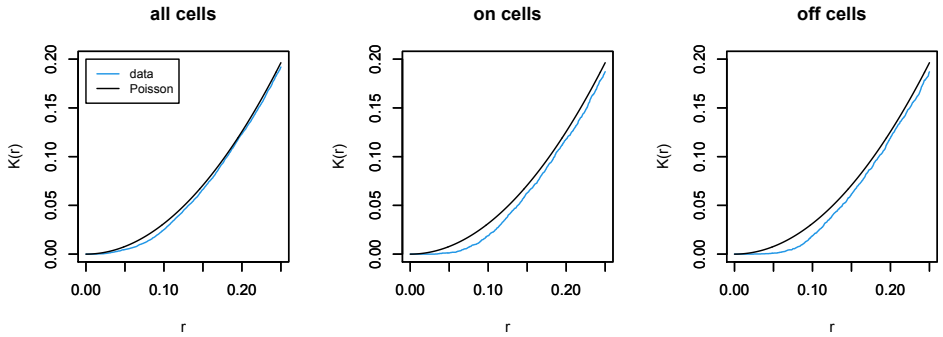
$$K_{12}(r) = K_{11}(r) = K_{22}(r) = K(r),$$

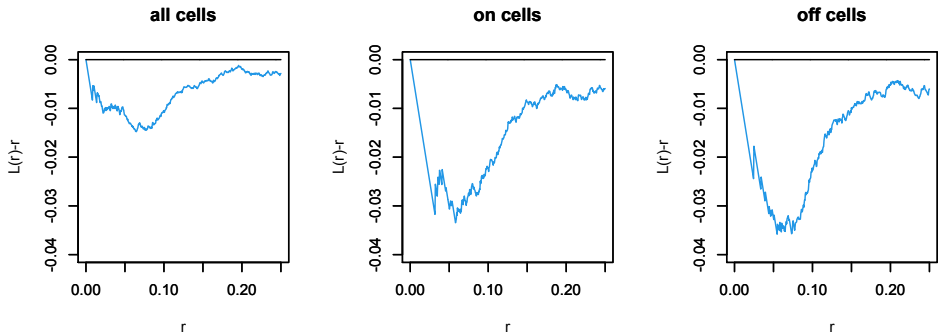
where $K(r)$ is the K function of the entire process N (M without marks).

Amacrine cells

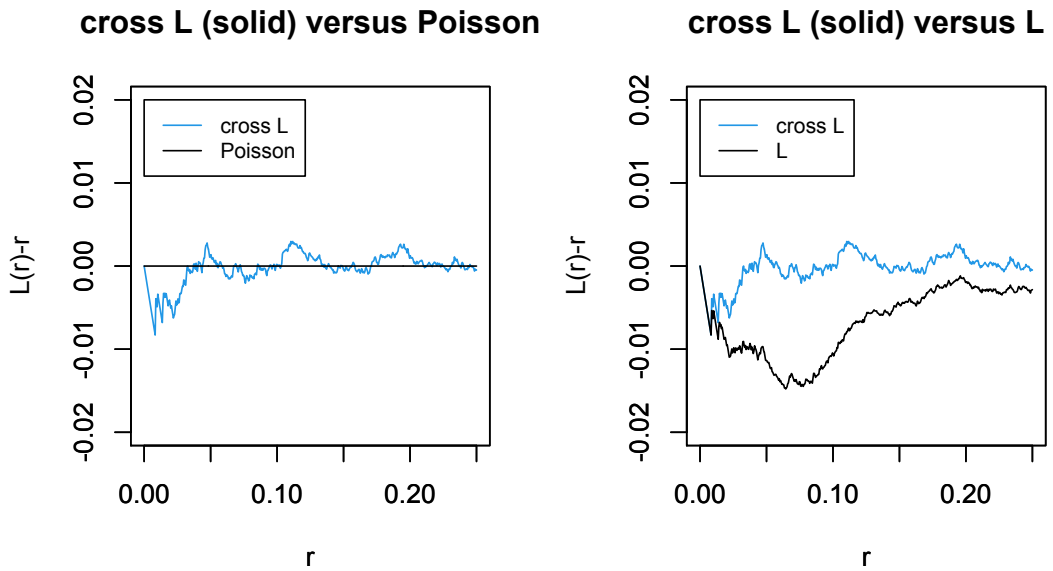
Estimated K (top) and $L(r) - r = \sqrt{K(r)/\pi} - r$ (bottom) functions for the whole pattern (left), the on cells (middle), and the off cells (right).







Amacrine cells: cross L function



Left: The cross $L_{12}(r) - r$ function is close to the Poisson line. Sub processes independent?

Right: The cross $L_{12}(r) - r$ seems to differ from the (unmarked) $L(r) - r$ function giving evidence against random labelling.

Remark: Formal tests would be needed.

Mark-correlation function (quantitative marks)

Let us define

$$c_{mm}(r) = \mathbb{E}_{or}(m(o)m(\mathrm{r})) \text{ for } r > 0,$$

the conditional expectation of the product of the marks of a pair of events in M given that one of the events is in the origin and the other distance r away.

For example, a value of $c_{mm}(r)$ that is larger than the squared mean mark μ^2 indicates increased marks at distance r.

The mark-correlation function is defined by

$$k_{mm}(r) = \frac{c_{mm}(r)}{\mu^2} \text{ for } r > 0.$$

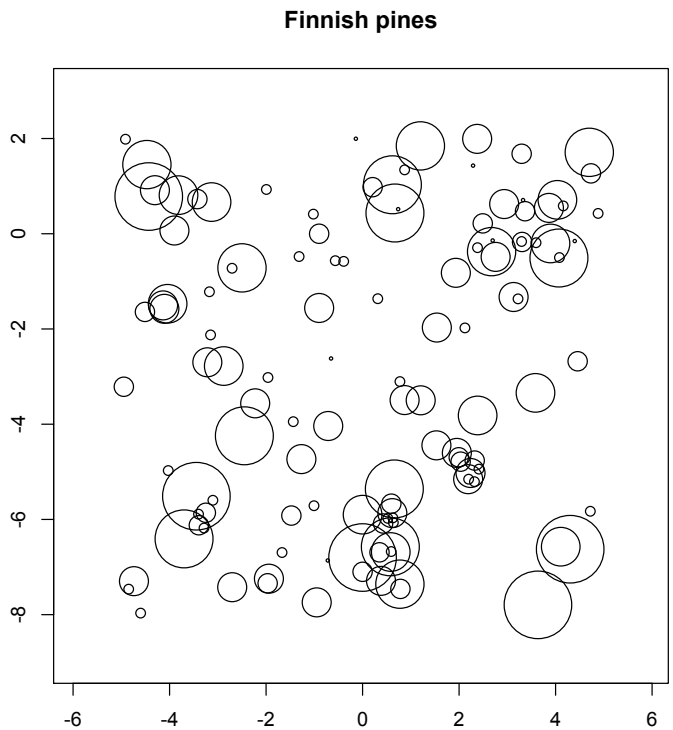
The mark correlation function has the value 1 if the marks are uncorrelated.

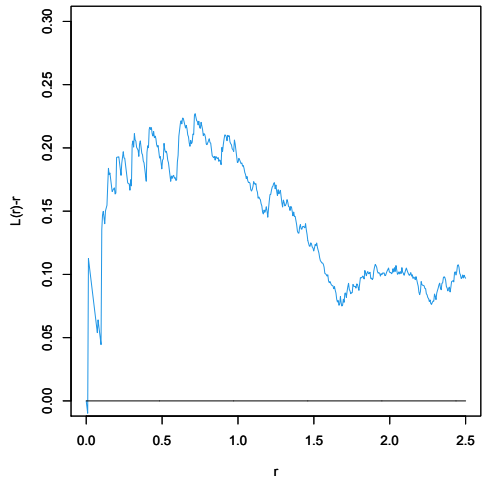
Remarks

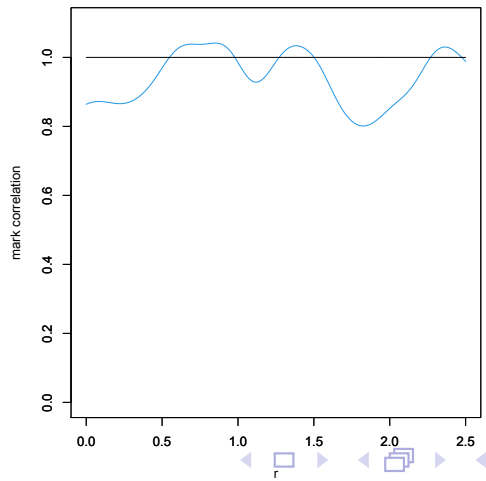
- $c_{mm}(r)$ in the mark correlation function can be replaced by some other function of marks.
- Mark-correlation function can be estimated by kernel estimation.

Finnish pines

Estimated $L(r) - r$ function (top right) and mark correlation function (bottom right) for the Finnish pines data (left).







Models for marked point patterns

How to generate a point pattern with varying sized discs, where the discs do not overlap?

Step 1: Generate an initial configuration by

- generating a realization of a Poisson point process with constant intensity λ in your observation window.
- generating an identically distributed radius for each point (event) from some predefined distribution function F_{pr}. The radii are mutually independent and independent of the point process.

Models for marked point patterns: non-overlapping discs

Step 2: Generate the final pattern by thinning the generated (Poisson) point pattern by letting all pairs of points whose associated discs intersect to compete with each other. A point is kept if it has a higher weight in all pairwise comparisons.

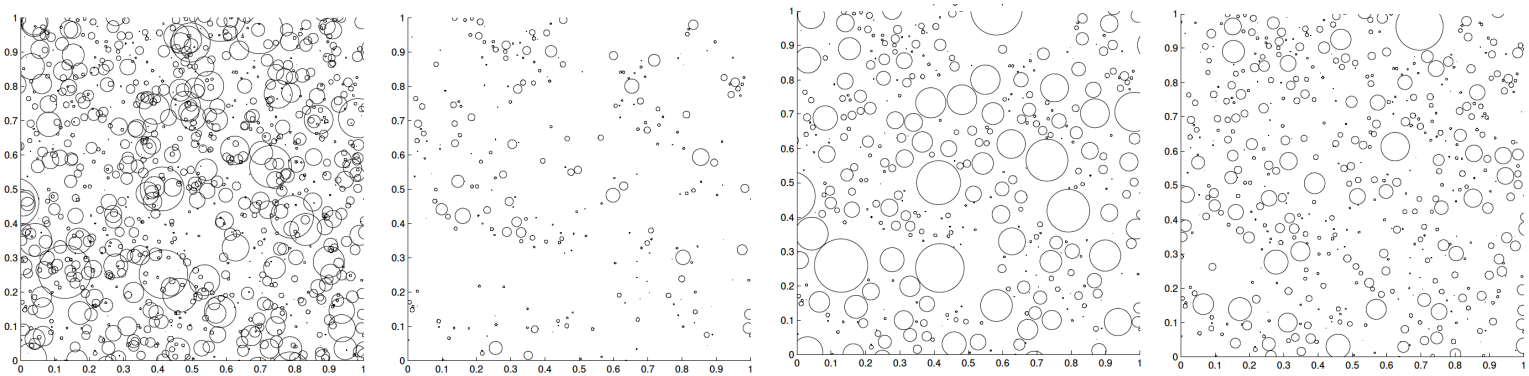
The weights can be assigned e.g. by

1. Pairwise assignment of weights: For each comparison, independent weights are assigned to the involved pair of points.
2. Global assignment of weights: Independent weights are assigned once and for all to all points.

The weight may depend on the associated radius. When the weights are deterministic functions of the radii, the two approaches coincide.

Models for marked point patterns: example patterns

Initial pattern: A realization of a Poisson process with intensity 1000 (left). Disc radii are exponentially distributed with mean 0.01.



From left to right, initial Poisson pattern, all intersecting discs removed, large discs kept, and a global thinning with uniformly distributed weights.

Strauss process

- Two points are neighbours if they are closer than distance R apart
- The density function for the point pattern x is

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

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 the interaction radius (range of interaction).
- α is a normalizing constant

Models for marked point patterns: Strauss process with two types of points

- Two type 1 points are neighbours if they are closer than distance R_1 apart, two type 2 points if they are closer than distance R_2 apart, and a type 1 and a type 2 points if they are closer than distance R_{12} apart.
- The density function for the point pattern x is

$$f(x) = \alpha \beta_1^{n(x_1)} \beta_2^{n(x_2)} \gamma_1^{s_1(x)} \gamma_2^{s_2(x)} \gamma_{12}^{s_{12}(x)},$$

where

  - $\beta_1 > 0$ and $\beta_2 > 0$ are the effects of a single type 1 and type 2 events, respectively.
  - $0 < \gamma_1 \leq 1$, $0 < \gamma_2 \leq 1$, and $0 < \gamma_{12} \leq 1$ are interaction parameters for $1-1$, $2-2$, and $1-2$ pairs, respectively.
  - $n_1(x)$ and $n_2(x)$ are the numbers of points in the sub patterns.
  - $s_1(x)$, $s_2(x)$, and $s_{12}(x)$ are the numbers of R_1, R_2, R_{12} close $1-1$, $2-2$, and $1-2$ pairs, respectively.
  - α is a normalizing constant

Models for marked point patterns: marked Strauss process, models for overlapping discs

- Two points (x_i, m_i) and (x_j, m_j) are neighbours if the discs $b(x_i; m_i)$ and $b(x_j; m_j)$ overlap.
- The density function for the point pattern $\{[x_i; m(x_i)]\}$is

$$f(x; m) = \alpha(\prod_{x_i \in x} \beta(m_i))\gamma^{s_m(x)},$$

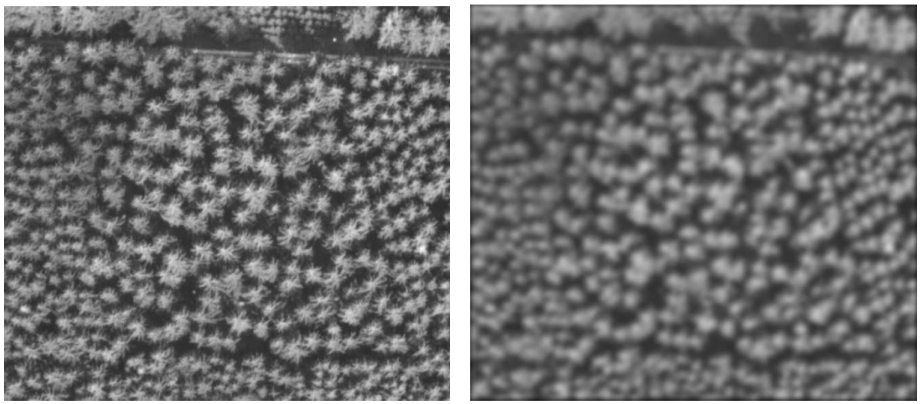
where

  - $\beta(m_i) > 0$ is the effect of a single event with mark m_i
  - $0 < \gamma \leq 1$ is an interaction parameter
  - $s_m(x)$ is the number of neighbours, i.e. the number of pairs of discs that overlap.
  - α is a normalizing constant

- Can be simulated by MCMC.

Point processes with noise

Example: How to estimate the tree top and/or tree base positions from aerial photos?



An original image on the left and a smoothed (filtered) image on the right.

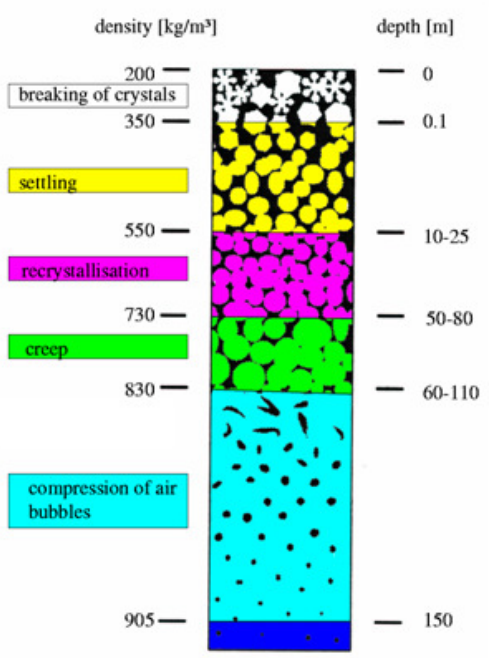
Point processes with noise: general model

Three sources of distortion:

1. Some points can be lost. In the tree example,
   - some trees (small trees shadowed by large ones) may not be identified.
   - model: each tree is given a probability to be identified as a tree, i.e. to give rise to a maximum.
2. Some points can be displaced, for example due to measurement error. In the tree example,
   - some trees are displaced as a consequence of image geometry and lighting conditions.
   - model: displacement of a tree is systematic but may even have a random (measurement error) component.
3. Some extra points that do not correspond to any real point are generated. In the tree example,
   - some ghost trees (large trees counted twice) may be added.
   - model: extra trees added according to a Poisson process with constant intensity

Point pattern with extra points: spatial pattern of air bubbles in polar ice

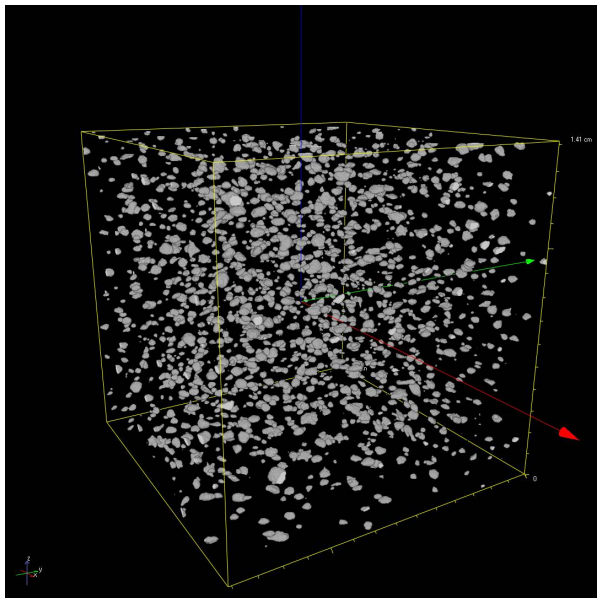
- Polar ice has information on the climate of the past
- To be able to interpret the ice core records, one has to know how old the ice is. How to determine the age?
- Polar ice is compacted snow. If we go deep enough, the air pores are isolated in the ice.
  $\rightarrow$ Study the anisotropy (deformation) of these air inclusions in the ice samples from a deep ice core at different depths.



Reference: Redenbach, C., Särkkä, A. and Sormani, M. Classification of Points in Superpositions of Strauss and Poisson Processes. *Spatial Statistics* **12**, (2015), 81–95.

Data

- Computer tomographic (CT) images of ice samples from Antarctica (imaged inside a cold room at -15° C (5° F)).
- In addition to the "real" air bubbles, the samples contain relaxation (extra) bubbles that do not give any information on the motion of the ice.



Problem/question

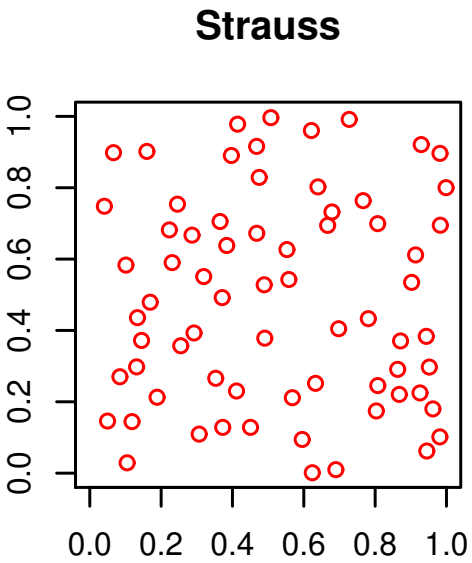
Since the extra bubbles do not carry any information on the motion of the ice and disturb the directional analysis, we would like to remove them before performing the analysis

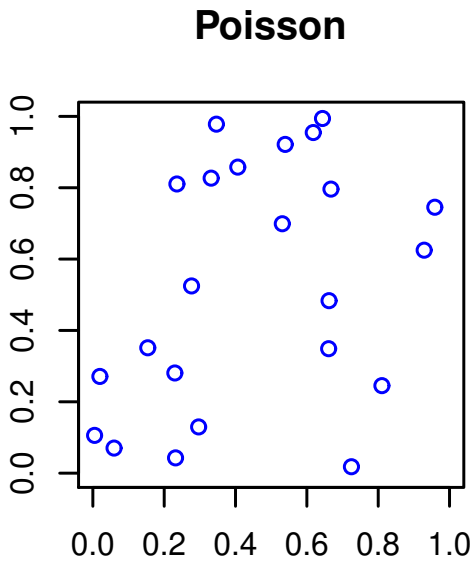
Question: How to classify each bubble either as "real pore" or as noise?

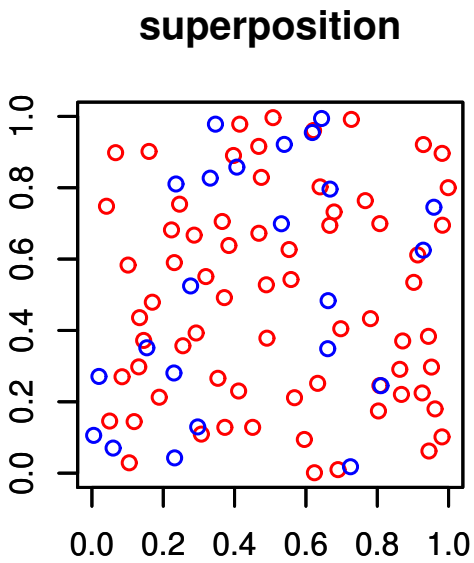
Model assumptions

- Real air pores modeled by a Strauss process (regular process).
- Noise bubbles are modeled as Poisson process with intensity λ_0.
- The complete point process N is a superposition of the Strauss process N_1 and the Poisson process N_0
- Parameter vector $\theta = (\lambda_0, \beta, \gamma, R)$

Strauss-Poisson superposition







Model assumptions

- The data consist of n bubbles, n_1 Strauss (real) bubbles and n_0 Poisson (noise) bubbles
- $Z \in \{0,1\}^n$ is a random vector such that

$$Z_i = \begin{cases} 1 & \text{if the } i\text{th bubble belongs to the Strauss pattern} \\ 0 & \text{if the } i\text{th bubble belongs to the Poisson pattern} \end{cases}$$

- We want to estimate
  - the vector Z (classification into real (Strauss) and noise (Poisson) bubbles)
  - the vector of parameters $\theta = (\lambda_0, \beta, \gamma, R)$

Bayesian approach

- To estimate the parameters, sample from the posterior distribution

$$\mathrm{P}(\theta, Z|y) \propto \mathrm{P}(y|\theta, Z)\Pi(\theta, Z),$$

where y is the point pattern we observe, $\mathrm{P}(y|\theta, Z)$ is a product of the Strauss density and the Poisson density, and $\Pi(\theta, Z)$ is the prior distribution for the parameters.

- Independent uniform priors for the parameters in $\theta = (\lambda_0, \beta, \gamma, R)$
- Prior for Z is

$$\Pi(Z|\theta) = \Pi(Z|n_0, \theta)\Pi(n_0|\theta) = \left(\frac{\lambda_0}{\lambda_0 + \lambda_1}\right)^{n_0} \left(\frac{\lambda_1}{\lambda_0 + \lambda_1}\right)^{n_1},$$

where λ_0 and λ_1 are the intensities of the Poisson and Strauss processes, respectively, and $\Pi(Z|n_0, \theta) = 1/\binom{n}{n_0}$.

MCMC algorithm

- Select initial parameters according to the prior distributions
- Iterate the following steps
  - Update θ: Pick one of the parameters in θ at random and propose to change it to $\theta_i' = \theta_i \exp(X)$, where $X \sim N(0, \tau^2)$.
  - Update Z: Pick one data point at random and propose to move it to the other process.

  Accept a move with the probability given by the Hastings ratio.
- Discard the first $K-1$ iterations. The remaining sample $(\theta^K, Z^K), ..., (\theta^N, Z^N)$ (or every kth value of it) is considered a sample from the posterior distribution of (θ, Z).

Results

- In a simulation study with fixed θ, we observed three classes of posterior probabilities for Z: high (classified as real), low (classified as noise) and intermediate:
  - The intermediate class consists of close pairs of bubbles: one of the points was randomly classified as real and the other one as noise.
  - The spatial pattern of real bubbles remained the same
- When estimating the model parameters simultaneously with the classification
  - isolated noise bubbles were typically classified as real bubbles
  - λ_0 was underestimated and β overestimated (too many real bubbles and too few noise bubbles)

Note on spatio-temporal point processes

- Classical spatial point processes are often only "snapshots" of a process that develops in time. A space-time description typically gives much more information and a deeper understanding of the underlying biological and physical processes.
- Spatio-temporal point processes are models for time-dependent, dynamic point patterns, where the points represent
  - events that take place at random times and at random locations (earthquakes), or
  - objects that move through space (animals, storm centers), or
  - objects that appear at random instants at random locations and remain there for a random length of time (trees in a forest)