A brief introduction to spatial point processes

Jean-François Coeurjolly
http://www-ljk.imag.fr/membres/Jean-Francois.Coeurjolly/

Laboratoire Jean Kuntzmann (LJK), Grenoble University

Notes

Preliminary

Files which can de downloaded
http://www-ljk.imag.fr/membres/Jean-Francois.Coeurjolly/documents/Lille/
or more simply on the workshop webpage, program page
http://math.univ-lille1.fr/ heinrich/geostoch2014/

- introductionSPP_cours.pdf : pdf file of the slides. Beamer version.
- introductionSPP_print.pdf : pdf file of the printed version.
- Short R code used to illustrate the talks.
- The code is using the **excellent** R package spatstat which can be downloaded from the R CRAN website.

Notes

Notes

Spatial data . . .

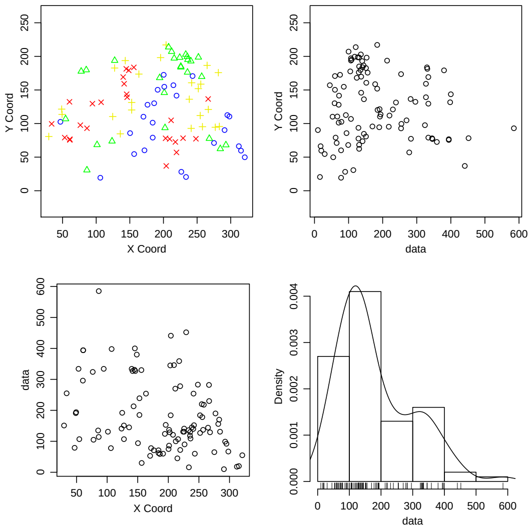
. . . can be roughly and mainly classified into three categories :

1. Geostatistical data.
2. Lattice data.
3. Spatial point pattern

Notes

Geostatistical data

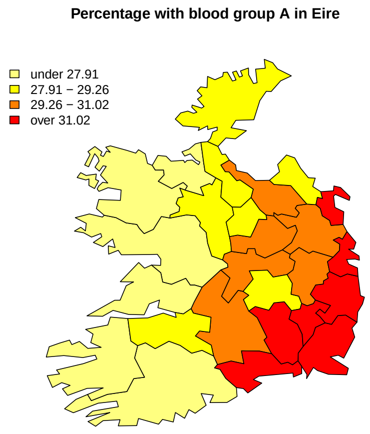
- sic.100 dataset (R package geoR)
- Cumulative rainfall in Switzerlan the 8th May.
- The observation consists in the **discretization** of a random field, $X = (X_u, u \in \mathbb{R}^2)$



Notes

Lattice data (1)

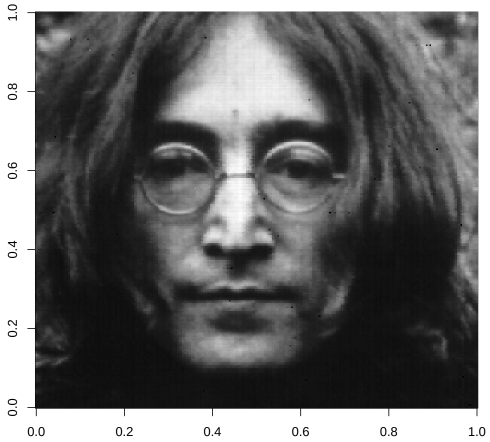
- Eire dataset (R package spdep)
- % of people with group A in eire, observed in 26 regions.
- The data are aggregated on the region ⇒ **random field on a network**.



Notes

Lattice data (2)

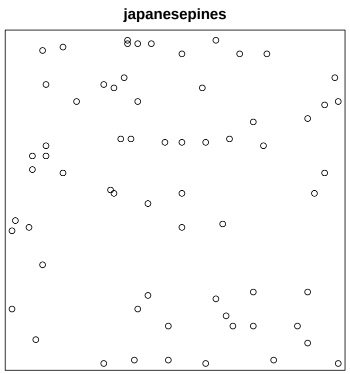
- Lennon dataset (`R` package `fields`)
- Real-valued random field (gray scale image with values in $[0, 1]$).
- Defined on the network $\{1, \dots, 256\}^2$.



Notes

Spatial point pattern (1)

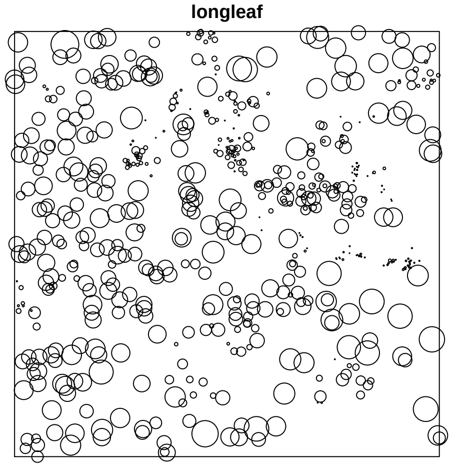
- Japanesepines dataset (`R` package `spatstat`)
- Locations of 65 trees on a bounded domain.
- $S = \mathbb{R}^2$ (equipped with $\|\cdot\|$).



Notes

Spatial point pattern (2)

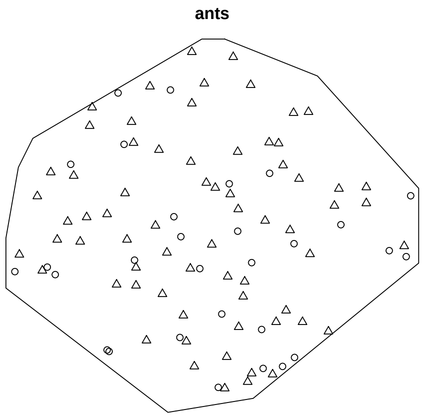
- Longleaf dataset (R package `spatstat`)
- Locations of 584 trees observed with their diameter at breast height.
- $\mathcal{S} = \mathbb{R}^2 \times \mathbb{R}^+$ (equipped with $\max(\|\cdot\|, |\cdot|)$).



Notes

Spatial point pattern (3)

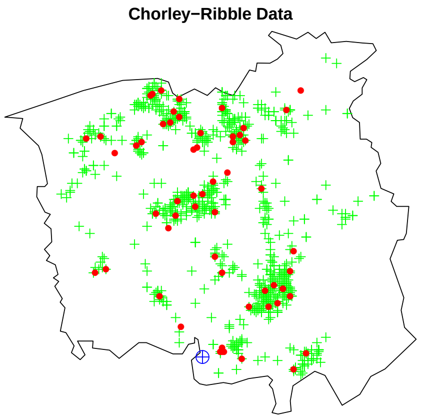
- Ants dataset (R package `spatstat`)
- Locations of 97 ants categorised into two species.
- $S = \mathbb{R}^2 \times \{0, 1\}$ (equipped with the metric $\max(\|\cdot\|, d_M)$ for any distance d_M on the mark space).



Notes

Spatial point pattern (3)

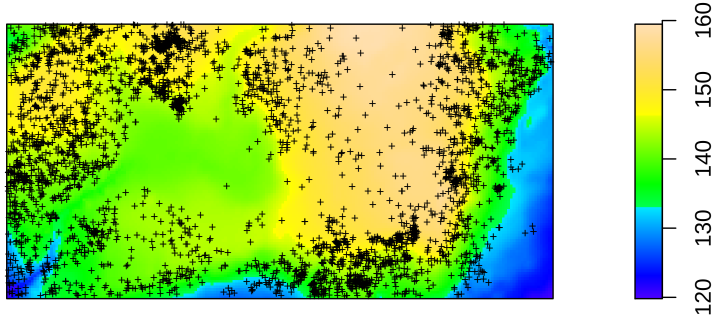
- chorley dataset (R package spatstat)
- Cases of larynx and lung cancers and position of an industrial incinerator.
- $S = \mathbb{R}^2 \times \{0,1\}$ (equipped with the metric $\max(\|\cdot\|, d_M)$ for any distance d_M on the mark space).



Notes

Spatial point pattern (4)

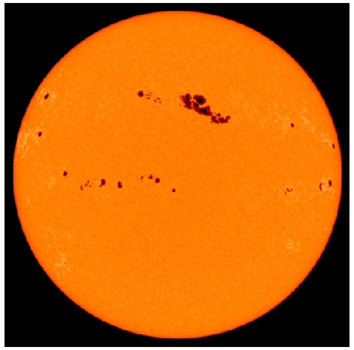
- Beischmedia dataset (R package spatstat)
- 3604 locations of trees observed with spatial covariates (here the elevation field).
- $S = \mathbb{R}^2$ (equipped with the metric $\|\cdot\|$), $z(\cdot) \in \mathbb{R}^2$.

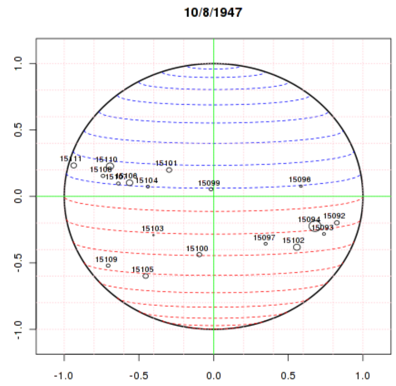


Notes

Spatial point pattern (5)

- Spatio-temporal point process on a complex space
- Daily observation of sunspots at the surface of the sun.
- can be viewed as the realization of a marked spatio-temporal point process on the sphere.
- $S = S_2 \times \mathbb{R}^+ \times \mathbb{R}^+$ (state, time, and mark)

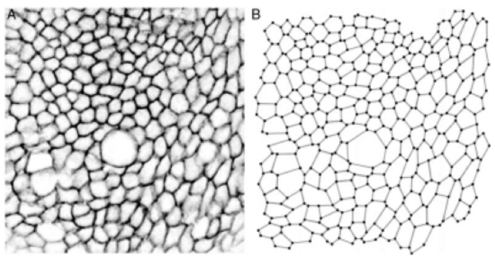


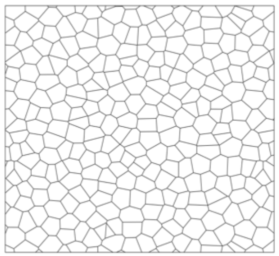


Notes

Spatial point pattern (6)

- Towards stochastic geometry ...
- Planar section of the pseudo-stratified epithelium of a drosophila wing marked with antibodies to highlight cell borders.
- The centers form of the tessellation form a point process.





Notes

References

A. Baddeley and R. Turner.
Spatstat : an R package for analyzing spatial point patterns.
Journal of Statistical Software, 12 :1–42, 2005.

N. Cressie.
Statistics for spatial data.
John Wiley and Sons, Inc, 1993.

P. J. Diggle.
Statistical Analysis of Spatial Point Patterns.
Arnold, London, second edition, 2003.

X. Guyon.
Random Fields on a Network.
Springer-Verlag, New York, 1991.

J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan.
Statistical Analysis and Modelling of Spatial Point Patterns.
Statistics in Practice. Wiley, Chichester, 2008.

J. Møller and R. P. Waagepetersen.
Statistical Inference and Simulation for Spatial Point Processes.
Chapman and Hall/CRC, Boca Raton, 2004.

Notes

Mathematical definition of a spatial point process ?

- S : Polish state space of the point process (equipped with the σ-algebra of Borel sets $\mathcal{B}$).
- A configuration of points is denoted $x = \{x_1, \ldots, x_n, \ldots\}$. For $B \subseteq S$: $x_B = x \cap B$.
- N_{lf} : space of **locally finite configurations**, i.e.

$$\{x, n(x_B) = |x_B| < \infty, \forall B \text{ bounded } \subseteq S\}$$

equipped with $\mathcal{N}_{lf} = \sigma\left(\{x \in N_{lf}, n(x_B) = m\}, B \in \mathcal{B}, B \text{ bounded}, m \geq 1\right)$.

Definition

A point process X defined on S is a measurable application defined on some probability space $(\Omega, \mathcal{F}, P)$ with values on N_{lf}.

Measurability of $X \Leftrightarrow N(B) = |X_B|$ is a r.v. for any bounded $B \in \mathcal{B}$.

Notes

Theoretical characterization of the distribution of X

Proposition

The distribution of a point process X

1. is determined by the finite dimensional distributions of its counting function, i.e. the joint distribution of $N(B_1), \dots, N(B_m)$ for any bounded $B_1, \dots, B_m \in \mathcal{B}$ and any $m \geq 1$.
2. is uniquely determined by its void probabilities, i.e. by

$$P(N(B) = 0), \qquad \text{for bounded } B \in \mathcal{B}.$$

- From now on, we assume that $S = \mathbb{R}^d$ (and even $d = 2$) or a bounded domain of $\mathbb{R}^2$.
- Everything can de extended to marked spatial point processes and/or to more complex domains.

Notes

Moment measures

- Moments play an important role in the modelling of classical inference.
- For point processes = moments of counting variables.

Definition : for $n \geq 1$ we define

- the n-th order moment measure (defined on S^n) by

$$\mu^{(n)} = \mathrm{E} \sum_{u_1,\dots,u_n} \mathbf{1}(\{u_1, \dots, u_n\} \in D), \quad D \subseteq S^n.$$

- the n-th order reduced moment measure (defined on S^n) by

$$\alpha^{(n)}(D) = \mathrm{E} \sum_{u_1,\dots,u_n}^{\neq} \mathbf{1}(\{u_1, \dots, u_n\} \in D), \quad D \subseteq S^n.$$

where the $\neq$ sign means that the n points are pairwise distinct.

Notes

Intensity functions

Assume $\mu^{(1)}$ and $\alpha^{(2)}$ are absolutely continuous w.r.t. Lebesgue measure, and denote by ρ and $\rho^{(2)}$ the densities.

Campbell Theorems

1. For any measurable function $h : S \to \mathbb{R}$

$$\mathrm{E} \sum_{u \in X} h(u) = \int_S h(u)\rho(u)\mathrm{d}u.$$

2. For any measurable function $h : S \times S \to \mathbb{R}$

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

$\rho(u)\mathrm{d}u \simeq$ Probability of the occurence of u in $B(u, \mathrm{d}u)$

$\rho^{(2)}(u,v) \simeq$ Probability of the occurence of u in $B(u, \mathrm{d}u)$ and v in $B(v, \mathrm{d}v)$.

Notes

Poisson point processes

Classical definition : $X \sim$Poisson(S, ρ)

- $\forall m \geq 1$, $\forall$ bounded and disjoint $B_1, \ldots, B_m \subset S$, the r.v. $X_{B_1}, \ldots, X_{B_m}$ are **independent**.
- $N(B) \sim \mathcal{P}\left(\int_B \rho(u)\mathrm{d}u\right)$ for any bounded $A \subset S$.
- $\forall B \subset S$, $\forall F \in N_{lf}$

$$P(X_B \in F) = \sum_{n \geq 0} \frac{e^{-\int_B \rho(u)\mathrm{d}u}}{n!} \int_B \cdots \int_B \mathbf{1}(\{x_1, \ldots, x_n\} \in F) \prod_{i=1}^{n} \rho(x_i)\mathrm{d}x_i.$$

- If $\rho(\cdot) = \rho$, X is said to be homogeneous which implies

$$\mathrm{E}N(B) = \rho|B|, \quad \mathrm{Var}N(B) = \rho|B|.$$

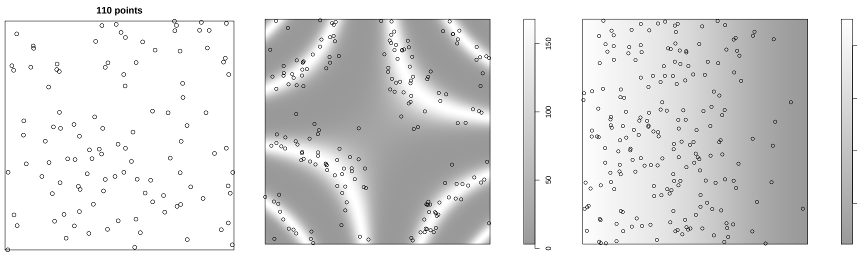
- and if $S = \mathbb{R}^d$, X is stationary and isotropic.

Notes

A few realizations on $S=[-1,1]^2$

- $\rho(u)=\beta e^{-u_1-u_1^2-.5u_1^3}$.
- $\rho=200$.
- $\rho(u)=\beta e^{2\sin(4\pi u_1 u_2)}$.

(β is adjusted s.t. the mean number of points in S, $\int_S \rho(u)\mathrm{d}u=200$.)



Notes

A few properties of Poisson point processes

Proposition : if $X\sim$Poisson(S,ρ)

- Void probabilities : $v(B)=P(N(B)=0)=e^{-\int_B(\rho(u)\mathrm{d}u)}$.
- For any $u,v\in S$, $\rho^{(2)}(u,v)=\rho(u)\rho(v)$ (also valid for $\rho^{(k)}$, $k\geq 1$)
- and if $|S|<\infty$, X admits a density w.r.t. Poisson$(S,1)$ given by

$$f(x)=e^{|S|-\int_S\rho(u)\mathrm{d}u}\prod_{u\in x}\rho(u).$$

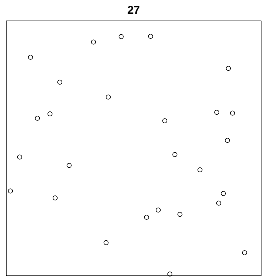
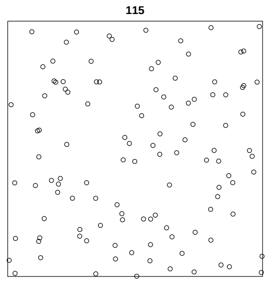
- Slivnyak-Mecke Theorem : for any non-negative function $h:S\times N_{lf}\to\mathbb{R}^+$, then

$$\mathrm{E}\sum_{u\in X}h(u,X\setminus u)=\int_S \mathrm{E}h(u,X)\rho(u)\mathrm{d}u.$$

<u>Example</u> : if $\rho(\cdot)=\rho$, $\mathrm{E}\sum_{u\in X\cap[0,1]^2}\mathbf{1}(d(u,X\setminus u)\leq R)=\rho\left(1-e^{-\rho\pi R^2}\right)$

Notes

Simulation

	
$\rho = 20,\ R = 0.1$	$\rho = 100,\ R = 0.05$
$\sum_{u \in x} \mathbf{1}(d(u, x \setminus u) \le R) = 9$	$\sum_{u \in x} \mathbf{1}(d(u, x \setminus u) \le R) = 60$
$\rho(1 - \exp(-\rho \pi R^2)) \simeq 9.33$	$\rho(1 - \exp(-\rho \pi R^2)) \simeq 54.41$

Notes

Statistical inference for a Poisson point process

- Simulation :
  - homogeneous case : very simple
  - inhomogeneous case : a **thinning** procedure can be efficiently done if $\rho(u) \le c$: simulate Poisson(c,W) and delete a point u with prob. $1 - \rho(u)/c$.
- Inference :
  - consists in estimating ρ, $\rho(\cdot;\theta)$ or $\rho(u)$ depending on the context.
  - All these estimates can be used even if the spatial point process is not Poisson (wait for a few slides)
  - Asymptotic properties very simple to derive under the Poisson assumption.
- Goodness-of-fit tests : tests based on quadrats counting, based on the void probability,. . .

Notes

Homogeneous case

- We consider here the problem of estimating the parameter ρ of a homogeneous Poisson point process defined on S and observed on a window $W \subseteq S$.
- Since $N(W) \sim \mathcal{P}(\rho|W|)$, the natural estimator of ρ is

$$\widehat{\rho} = N(W)/|W|$$

Properties

- (i) $\widehat{\rho}$ corresponds to the maximum likelihood estimate.
- (ii) $\widehat{\rho}$ is unbiased.
- (iii)$\mathrm{Var}\,\widehat{\rho} = \frac{\rho}{|W|}$.

Proof : (i) follows from the definition of the density (ii-iii) can be checked using the Campbell formulae.

Notes

Homogeneous case (2)

Asymptotic results

- For large $N(W)$, $\widehat{\rho}|W| \simeq \mathcal{N}(\rho|W|, \rho|W|)$ and so

$$|W|^{1/2}(\widehat{\rho} - \rho) \simeq \mathcal{N}(0, \rho).$$

  (the approximation is actually a convergence as $W \to \mathbb{R}^d$)
- Variance stabilizing transform :

$$2|W|^{1/2}(\sqrt{\widehat{\rho}} - \sqrt{\rho}) \simeq \mathcal{N}(0,1)$$

- We deduce a $1-\alpha$ ($\alpha \in (0,1)$) confidence interval for ρ

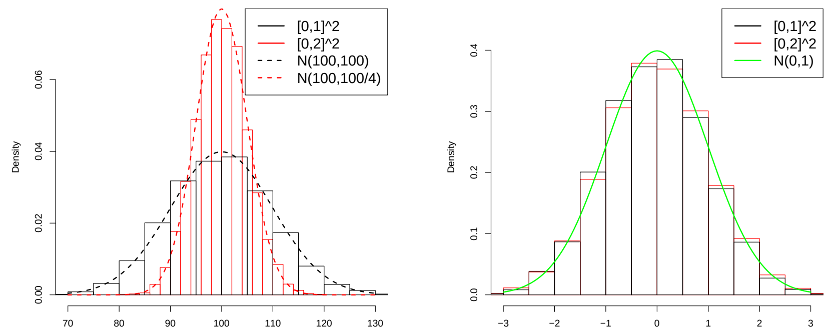
$$\mathrm{IC}_{1-\alpha}(\rho) = \left(\sqrt{\widehat{\rho}} \pm \frac{z_{\alpha/2}}{2|W|^{1/2}} \right)^2.$$

Notes

A simulation example

We generated $m = 10000$ replications of homogeneous Poisson point processes with intensity $\rho = 100$ on $[0,1]^2$ (blcak plots) and on $[0,2]^2$ (red plots).

Histograms of $\widehat{\rho}$	Histograms of $\widehat{2}|W|^{1/2}(\sqrt{\widehat{\rho}} - \sqrt{\rho})$



Notes

Examples | Definitions, Poisson | Summary statistics | Modelling and inference

A simulation example

We generated $m = 10000$ replications of homogeneous Poisson point processes with intensity $\rho = 100$ on $[0,1]^2$ (black plots) and on $[0,2]^2$ (red plots).

	$W = [0,1]^2$	$W = [0,2]^2$
Emp. Mean of $\widehat{\rho}$	100.17	100.07
Emp. Var. of $\widehat{\rho}$	98.57	25.69
Emp. Coverage rate of 95% confidence intervals	95.31%	94.78%

Notes

Application : pines datasets

- We consider three unmarked datasets : japanesepines, swedishpines, finpines.
- Plot the data, estimate the intensity parameter.
- Construct a confidence interval for each of them. Which one is significantly more abundant ?
- Judge the assumption of the Poisson model using a GoF test based on quadrats.

Notes

Inhomogeneous case : parametric estimation

- Assume that ρ is parametrized by a vector $\theta \in \mathbb{R}^p$ ($p \geq 1$). The most well-known model is the log-linear one :
$$\rho(u) = \rho(u;\theta) = \exp(\theta^\top z(u))$$
where $z(u) = (z_1(u), z_2(u), \ldots, z_p(u))$ correspond to known spatial functions or spatial covariates.
- θ can be estimated by maximizing the log-likelihood on W
$$l_W(X,\theta) = \sum_{u \in X_W} \log \rho(u;\theta) + \int_W (1-\rho(u;\theta))\mathrm{d}u$$
$$= |W| + \underbrace{\sum_{u \in X_W} \theta^\top z(u) - \int_W \exp(\theta^\top z(u))\mathrm{d}u}_{:=\ell_W(X,\theta)}.$$

In other words
$$\widehat{\theta} = \mathrm{Argmax}_\theta \;\; \ell_W(X,\theta).$$

Notes

Inhomogeneous case : parametric estimation (2)

- Why would $\widehat{\theta}$ be a good estimate ?
  Compute the score function

$$s_W(X,\theta) = \nabla \ell_W(X,\theta) = \sum_{u \in X_W} z(u) - \int_W z(u) \underbrace{\exp(\theta^\top z(u))}_{:=\rho(u)} \mathrm{d}u.$$

  The true parameter θ_0 (i.e. $X \sim P_{\theta_0}$) minimizes the expectation of the score function. Indeed from Campbell formula

$$\mathrm{E} s_W(X,\theta) = \int_W z(u) \left(\exp(\theta_0^\top z(u)) - \exp(\theta^\top z(u)) \right) \mathrm{d}u = 0$$

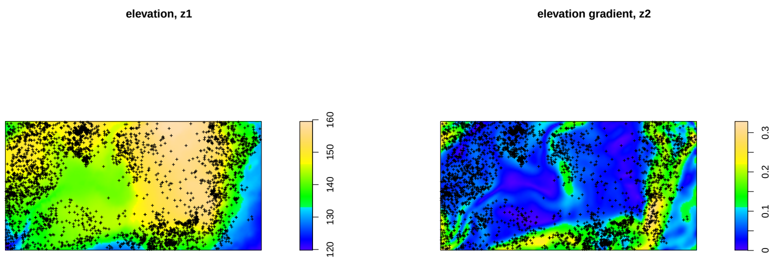
  when $\theta = \theta_0$.

- Rathbun and Cressie (1994) showed the **strong consistency** and the **asymptotic normality** of $\widehat{\theta}$ as $W \to \mathbb{R}^d$.

Notes

Data example : dataset `bei`

A point pattern giving the locations of 3605 trees in a tropical rain forest. Accompanied by covariate data giving the elevation (altitude) (z_1) and slope of elevation (z_2) in the study region.



Assume an inhomogeneous Poisson point process (which is not true, see the next chapter) with intensity

$$\log \rho(u) = \beta + \theta_1 z_1(u) + \theta_2 z_2(u).$$

<u>Question</u> : how can we prove that each covariate has a significant influence ?

Notes

Inhomogeneous case : nonparametric estimation

(Diggle 2003)

- Idea is to mimic the kernel density estimation to define a nonparametric estimator of the spatial function ρ.
- Let $k : \mathbb{R}^d \to \mathbb{R}^+$ a symmetric kernel with intensity one. Examples of kernels
  - Gaussian kernel : $(2\pi)^{-d/2}\exp(-\|y\|^2/2)$.
  - Cylindric kernel : $\frac{1}{\pi}\mathbf{1}(\|y\| \le 1)$.
  - Epanecnikov kernel : $\frac{3}{4}\mathbf{1}(|y| < 1)(1-|y|^2)$.
- Let h be a positive real number (which will play the role of a bandwidth window), then the nonparametric estimate (with border correction) at the location v is defined as

$$\widehat{\rho}_h(v) = K_h(v)^{-1} \sum_{u \in X_W} \frac{1}{h^d} k\left(\frac{\|v-u\|}{h}\right)$$

Notes

Intuitively, this works . . .

Indeed, using the Campbell formula and a change of variables we can obtain

$$\begin{aligned}
\mathrm{E}\,\widehat{\rho}_h(v) &= K_h(v)^{-1}\mathrm{E}\sum_{u \in X_W} \frac{1}{h^d} k\left(\frac{\|v-u\|}{h}\right) \\
&= K_h(v)^{-1}\int_W \frac{1}{h^d} k\left(\frac{\|v-u\|}{h}\right)\rho(u)\mathrm{d}u \\
&= K_h(v)^{-1}\int_{\frac{W-v}{h}} k\left(\|\omega\|\right)\rho(\omega h + v)\mathrm{d}\omega \\
&\overset{\text{h small}}{\simeq} K_h(v)^{-1}\int_{\frac{W-v}{h}} k\left(\|\omega\|\right)\rho(v)\mathrm{d}\omega \\
&\simeq \rho(v).
\end{aligned}$$

More theoretical justifications and properties and a discussion on the bandwidth parameter and edge corrections can be found in Diggle (2003).

Notes

Objective and classification

Objective :

- Define some descriptive statistics for s.p.p. (independently on any model so).
- Measure the abundance of points, the clustering or the repulsiveness of a spatial point pattern w.r.t. the Poisson point process.

Classification :

- First-order type based on the intensity function.
- Second-order type statistics : pair correlation function, Ripley's K function.
- Statistics based on distances : empy space function F, nearest-neigbour G, J function.

(We assume that ρ and $\rho^{(2)}$ exist in the rest of the talk)

Notes

Summary statistics based on the intensity function

Thanks to **Campbell formulae**, the estimates of the intensity for a Poisson point process can be used to estimate the intensity of a general spatial point process X. In particular

1. if X is stationary $\widehat{\rho} = N(W)/|W|$ is an estimate of ρ.
2. Non-stationary, parametric estimation of the intensity : if $\rho(u) = \rho(u;\theta)$ can be used using the "Poisson likelihood", i.e.

$$l_W(X,\theta) = \sum_{u \in X_W} \log \rho(u;\theta) - \int_W \rho(u;\theta)\mathrm{d}u.$$

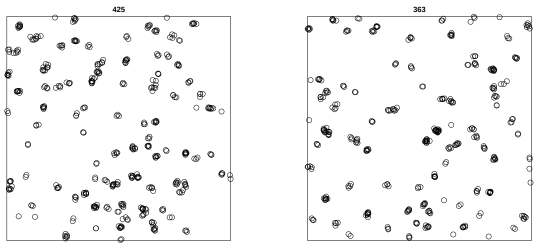
3. Non stationary, non-parametric estimation of the intensity (see previous chapter for notation) :

$$\widehat{\rho}_h(u) = K_h(u)^{-1} \sum_{v \in X_W} \frac{1}{h^d} k\left(\frac{\|v-u\|}{h}\right).$$

Notes

A simulation example in the stationary case

We generated $m = 10000$ replications of a stationary log-Gaussian Cox processes (Thomas process, $\kappa = 50$, $\sigma = .005$) with intensity $\rho = 400$.



	$W = [0,1]^2$	$W = [0,2]^2$
Emp. Mean of $\widehat{\rho}$	400.4	399.5
Emp. Var. of $\widehat{\rho}$	1741.4	507.4

- A survey of the estimation of the asymptotic variance of $\widehat{\rho}$ can be found in Prokesova and Heinrich (2010) and references therein.

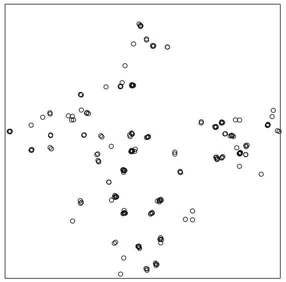
Notes

Parametric intensity estimation for non Poisson models

We generated $B = 1000$ replications of Thomas process with parameters $\kappa = 50$, $\sigma = .005$ and with intensity function

$$\rho(u) = \exp(\beta - \theta u_1^2 u_2^2)$$

with $\theta = -2$ and β adjusted s.t. $\mathrm{E}N(W) = 200$ for $W = [0,1]^2$ and 800 for $W = [0,2]^2$.



Then for each replication, θ is estimated using the "Poisson likelihood"

	$W = [0,1]^2$	$W = [0,2]^2$
Emp. Mean of $\widehat{\theta}$	-2.03	-2.01
Emp. Var. of $\widehat{\theta}$	0.13	0.03

- Asymptotic results are more awkward to derive and depend on mixing coefficients of the spatial point process X.
- See Guan (2006), Guan and Loh (2008), Waagepetersen, Guan and Jalilian (2012) and Coeurjolly and Møller (2012) for details and refinements.

Notes

Ripley's K function

We assume (for simplicity) the stationarity and isotropy of X.

Definition

The Ripley's K function is literally defined for $r \geq 0$ by

$$K(r) = \frac{1}{\rho}\,\mathrm{E}\Big(\text{number of extra events within distance r of a randomly chosen event}\Big)$$
$$= \frac{1}{\rho}\,\mathrm{E}\Big(N(B(0,r)\setminus 0)\mid 0\in X\Big)$$

We define the L function as $L(r) = (K(r)/\pi)^{1/2}$.

Properties :

- Under the Poisson case, $K(r) = \pi r^2$; $L(r) = r$.
- If $K(r) > \pi r^2$ or $L(r) > r$ (resp. $K(r) < \pi r^2$ or $L(r) < r$) we suspect clustering (regularity) at distances lower than r.

Notes

Examples Definitions, Poisson Summary statistics Modelling and inference

Pair correlation function

Definition

If ρ and $\rho^{(2)}$ exist, then the pair correlation function is defined by

$$g(u,v) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)}$$

where we set for convention $a/0 = 0$ for $a \geq 0$.

$$g(u,v)\begin{cases} = 1 & \text{if } X \sim \text{Poisson}(S,\rho). \\ > 1 & \text{for attractive point pattern.} \\ < 1 & \text{for repulsive point pattern.}\end{cases}$$

If $S = \mathbb{R}^d$ and X is stationary and isotropic, then

$$g(u,v) = \frac{\rho^{(2)}(\|v-u\|)}{\rho^2} = \overline{g}(\|v-u\|.$$

Notes

Particular case for stationary and isotropic processes

Theorem

For stationary and isotropic processes in $S = \mathbb{R}^d$

$$g(r) = \frac{K'(r)}{\sigma_d r^{d-1}}$$

where $\sigma_d = d\omega_d$ is the surface area of unit sphere $\mathbb{S}^{d-1}$ in $\mathbb{R}^d$.

Proof : Using polar decomposition we obtain

$$K(r) = \int_{B(0,r)} g(\|u\|)\mathrm{d}u = \int_0^r \int_{S^{d-1}} t^{d-1} g(t)\mathrm{d}t = \sigma_d \int_0^r t^{d-1} g(t)\mathrm{d}t.$$

Notes

Edge corrected estimation of the K function

Definition

We define

- the border-corrected estimate as

$$\widehat{K}_{BC}(r) = \frac{1}{\widehat{\rho}} \sum_{u \in X_{W_{\ominus r}}, v \in X_W}^{\neq} \frac{\mathbf{1}(v \in B(u,R))}{N(W_{\ominus r})}$$

  where $W_{\ominus r} = \{u \in W : B(u,r) \subseteq W\}$ is the erosion of W by r.

- the translation-corrected estimate as

$$\widehat{K}_{TC}(r) = \frac{1}{\widehat{\rho^2}} \sum_{u,v \in X_W}^{\neq} \frac{\mathbf{1}(v - u \in B(0,r))}{|W \cap W_{v-u}|}$$

  where $W_u = W + u = \{u + v : v \in W\}$.

Remark : everything extends to 2nd-order reweighted stationary point processes ; asymptotic properties depend on mixing conditions, . . .

Notes

Estimation of the pair correlation function

For convenience, we consider only stationary and isotropic point processes.

- Then, the pair correlation function $g(u,v) = g(\|u - v\|)$ can be estimated using the following edge corrected kernel estimate

$$\widehat{g}(r) = \frac{1}{\widehat{\rho^2}} \sum_{u,v \in X_W}^{\neq} \frac{k_h(\|v-u\| - r)}{\sigma_d r^{d-1} |W \cap W_{v-u}|}$$

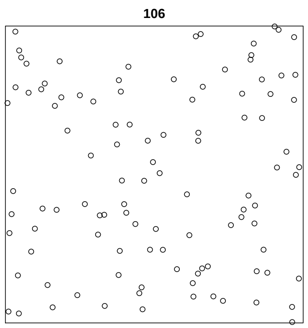
where $k_h(t) = h^{-d} k(t/h)$.

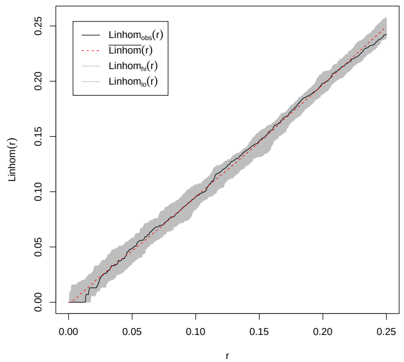
- Alternatively, we can estimate estimate the derivative of the K function (after smooting using e.g. spline techniques) and define

$$\widehat{g}(r) = \frac{\widehat{K'}(r)}{\sigma_d r^{d-1}}.$$

Notes

Example of L function for a Poisson point pattern

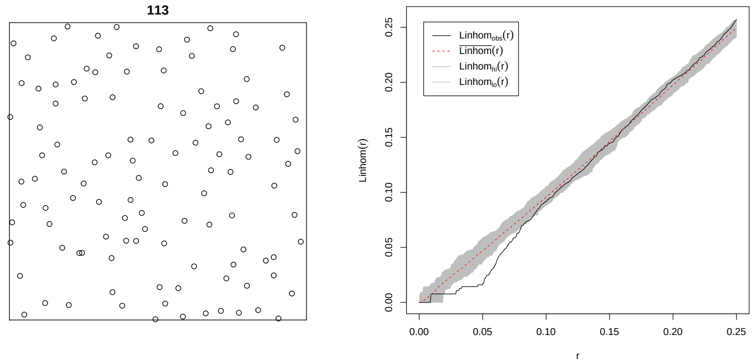




- The enveloppes are constructed using a Monte-Carlo approach under the Poisson assumption.
- $\Rightarrow$ we don't reject the Poisson assumption.

Notes

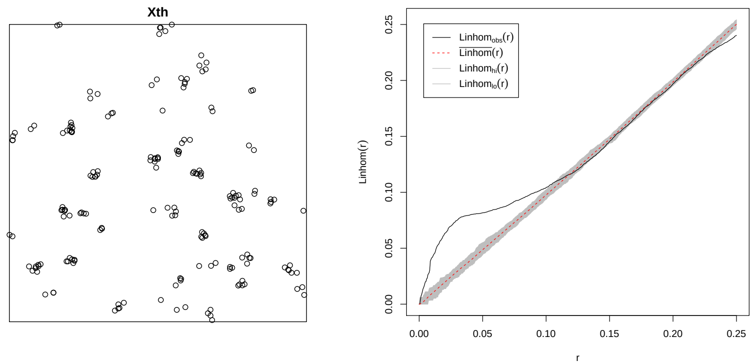
Example of L function for a repulsive point pattern



- ⇒ the point pattern does not come from the realization of a homogeneous Poisson point process.
- exhibits repulsion at short distances ($r \leq .05$)

Notes

Example of L function for a clustered point pattern



- ⇒ the point pattern does not come from the realization of a homogeneous Poisson point process.
- exhibits attraction at short distances ($r \leq .08$).

Notes

Statistics based on distances : F, G and J functions

Assume X is stationary (definitions can be extended in the general case)

Definition

- The empty space function is defined by

$$F(r) = P(d(0, X) \leq r) = P(N(B(0, r)) > 0), \qquad r > 0.$$

- The nearest-neighbour distribution function is

$$G(r) = P(d(0, X \setminus 0) \leq r | 0 \in X)$$

- J-function : $J(r) = (1 - G(r))/(1 - F(r)), \qquad r > 0.$

- Poisson case : $\forall r > 0$, $F(r) = G(r) = 1 - e^{-\pi r^2}$, $J(r) = 1$.
- $F(r) < F_{pois}(r)$, $G(r) > G_{pois}(r)$, $J(r) < 1$: attraction at dist. $< r$.
- $F(r) > F_{pois}(r)$, $G(r) < G_{pois}(r)$, $J(r) > 1$: repulsion at dist. $< r$.

Notes

Non-parametric estimation of F, G and J

As for the K and L functions, several edge corrections exist. We focus here only on the border correction. We assume that X is observed on a bounded window W with positive volume.

Definition

- Let $I \subseteq W$ be a finite regular grid of points and $n(I)$ its cardinality. Then, the (border corrected) estimator of F is

$$\widehat{F}(r) = \frac{1}{n(I_r)} \sum_{u \in I_r} \mathbf{1}(d(u, X) \leq r)$$

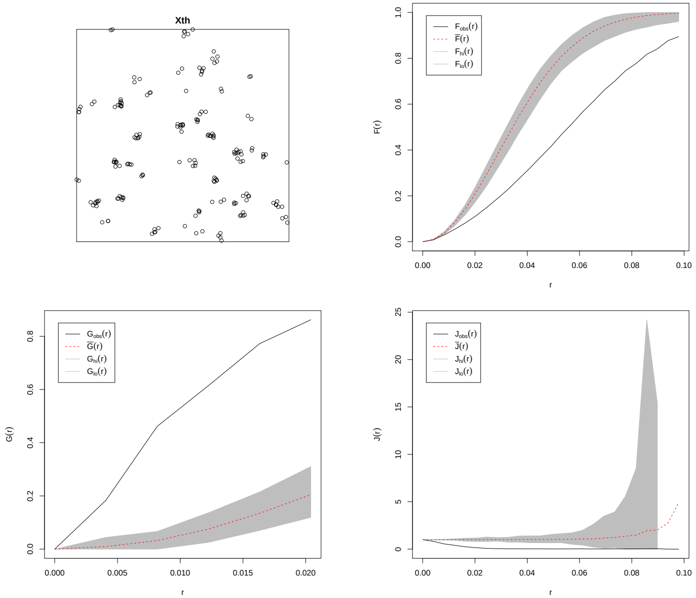
where $I_r = I \cap W_{\ominus r}$.

- The (border corrected) estimator of G is

$$\widehat{G}(r) = \frac{1}{N(W_{\ominus r})} \sum_{u \in X \cap W_{\ominus r}} \mathbf{1}(d(u, X \setminus u) \leq r)$$

Notes

Application to a clustered point pattern data



Notes

Objective

The main objectives of this section are

- to present more realistic models than the too simple Poisson point process to take into account the spatial dependence between points.
- to present statistical methodologies to infer these models.

We can distinguish several classes of models for spatial point processes

1. point processes based on the thinning of a Poisson point processes, on the superimposition of Poisson point processes. [sometimes hard to relate the stochastic process producing the realization and the physical phenomenon producing the data]
2. Cox point processes (which include Cluster point processes,. . .).
3. Gibbs point processes.
4. Determinental point processes.

Notes

An attempt to classify these models . . .

Model	Allows to model	Are moments expressible in a closed form ?	Density w.r.t. Poisson ?
Cox	**attraction**	**yes**	**no**
Gibbs	**repulsion** but also attraction	**no**	**yes**
Determinental	**repulsion**	**yes**	**yes**

This course only focuses on the two first classes of point processes, i.e. on Cox and Gibbs point processes.

Notes

Definition

We let $S \subseteq \mathbb{R}^d$ throughout this section. B denotes any bounded domain $\subseteq S$.

Definition

Suppose that $Z = \{Z(u) : u \in S\}$ is a nonnegative random field so that with probability one, $u \to Z(u)$ is a locally integrable function. If the conditional distribution of X given Z is a Poisson process on S with intensity function Z, then X is said to be a *Cox process* driven by Z.

Remarks :

- Z is a random field means that $Z(u)$ is a random variable $\forall u \in S$.
- if $\mathrm{E}Z(u)$ exists and is locally integrable then w.p. 1, $Z(u)$ is a locally integrable function.

Notes

Basic properties

Proposition

1. Provided $Z(u)$ has finite expectation and variance for any $u \in S$

$$\rho(u) = \mathrm{E}Z(u),\ \rho^{(2)}(u,v) = \mathrm{E}[Z(u)Z(v)],\ g(u,v) = \frac{\mathrm{E}[Z(u)Z(v)]}{\rho(u)\rho(v)}.$$

2. The void probabilities are given by

$$v(B) = \mathrm{E}\exp\left(-\int_B Z(u)\mathrm{d}u\right)$$

for bounded $B \subseteq S$.

Proof : direct consequence of the fact that $X|Z$ is a Poisson point process with intensity function Z.

Notes

Over-dispersion of Cox processes

Proposition

Let A, B bounded sets of S, then

$$\mathrm{Cov}(N(A), N(B)) = \int_A \int_B \mathrm{Cov}(Z(u), Z(v))\mathrm{d}u\mathrm{d}v + \int_{A\cap B} \mathrm{E}Z(u)\mathrm{d}u$$

Consequence :

- In particular, $\mathrm{Var}N(A) \geq \mathrm{E}N(A)$ with equality only when X is a Poisson process.
- $\Rightarrow$ over-dispersion of the counting variables.

Other remarks :

- Most of models have pcf such that $g \geq 1$ (but a few exceptions $\exists$).
- If $S = \mathbb{R}^d$ and X is stationary and/or isotropic then X is stationary and/or isotropic.
- Explicit expressions of the F, G and J functions in the stationary case are in general difficult to derive.

Notes

A first example

Definition

A *mixed Poisson process* is a Cox process where $Z(u) = Z_0$ is given by a positive random variable for any $u \in S$, i.e. $X|Z_0$ follows a homogeneous Poisson process with intensity Z_0.

- Limited interest ...
- X is stationary and (provided Z_0 has first two moments)

$$\rho = \mathrm{E}Z_0 \quad \text{and} \quad g(u,v) = \frac{\mathrm{E}[Z_0^2]}{\mathrm{E}[Z_0]^2} \geq 1.$$

- The K and L functions are given by

$$K(r) = \beta\omega_d r^d \quad \text{and} \quad L(r) = \beta^{1/d} r \geq r$$

where $\omega_d = |B(0,1)|$ and $\beta = \frac{\mathrm{E}[Z_0^2]}{\mathrm{E}[Z_0]^2}$.
(recall that $K'(r) = d\omega_d g(r) r^{d-1}$).

Notes

Neymann-Scott processes

Definition

Let C be a stationary Poisson process on $\mathbb{R}^d$ with intensity $\kappa > 0$. Conditional on C, let $X_c, c \in C$ be independent Poisson processes on $\mathbb{R}^d$ where X_c has intensity function

$$\rho_c(u) = \alpha k(u - c)$$

where $\alpha > 0$ is a parameter and k is a kernel (i.e. for all $c \in \mathbb{R}^d$, $u \to k(u-c)$ is a density function). Then $X = \cup_{c \in C} X_c$ is a Neymann-Scott process with cluster centres C and clusters $X_c, c \in C$.

- X is also a Cox process on $\mathbb{R}^d$ driven by $Z(u) = \sum_{c \in C} \alpha k(u - c)$.
- Simulating a Neymann-Scott process (on W) is very simple (if k has compact support $T < \infty$)
  1. Generate $C \sim \text{Poisson}(W \oplus T, \kappa)$.
  2. For each $c \in C$, generate $X_c \sim \text{Poisson}(W, \rho_c)$.
  3. Concatenate all the X_c's.
- If k has unbounded support, an exact simulation is still possible.

Notes

Two classical NS pp

We obtain specific models by choosing specific kernel densities.

1. the *Matérn cluster process* where

$$k(u) = \mathbf{1}(\|u\| \leq R)\frac{1}{\omega_d R^d}$$

   is the uniform density on the $B(0,R)$.

2. the *Thomas process* where

$$k(u) = \left(\frac{1}{2\pi\sigma^2}\right)^{d/2} \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right)$$

   is the density of $\mathcal{N}(0, \sigma^2 I_d)$.

When R is small or when σ is small, then point pattern exhibit strong attraction.

Notes

Basic properties of NS pp

- κ is the mean number of cluster centres per unit square, α is the mean number of daughters points per cluster.
- X is stationary (since Z is stationary) and is isotropic if $k(u) = k(\|u\|)$.
- Intensity of X : $\rho(u) = \alpha\kappa$.
- The (stationary) pair correlation function is given by

$$g(u,v) = 1 + \frac{k * k(v-u)}{\kappa} \geq 1 \quad \text{where} \quad k*k(u) = \int k(c)k(v-u+c)\mathrm{d}c.$$

- The F, G and J functions are also expressible in terms of k. In particular

$$J(r) = \int k(u)\exp\left(-\alpha \int_{\|v\|\leq r} k(u+v)\mathrm{d}v\right)\mathrm{d}u$$

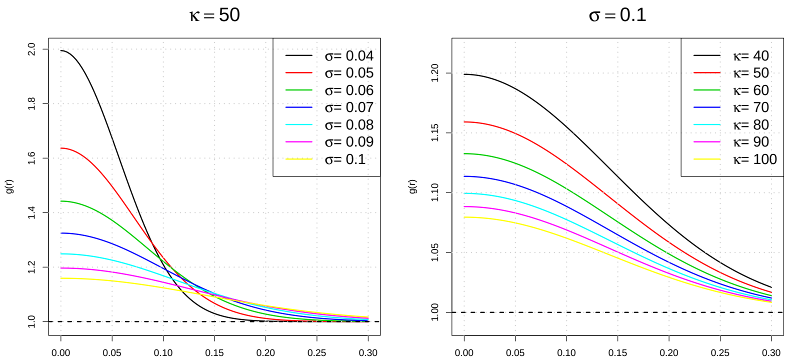
  whereby we deduce that $\exp(-\alpha) \leq J(r) \leq 1$.

Notes

Back to the Thomas process

Recall that k is the density of a $\mathcal{N}(0, \sigma^2 I_d)$. Applying the previous results, we get (for the pcf)

$$g(r) = 1 + \frac{1}{(4\pi\sigma^2)^{d/2}} \exp\left(-r^2/(4\sigma^2)\right)/\kappa$$

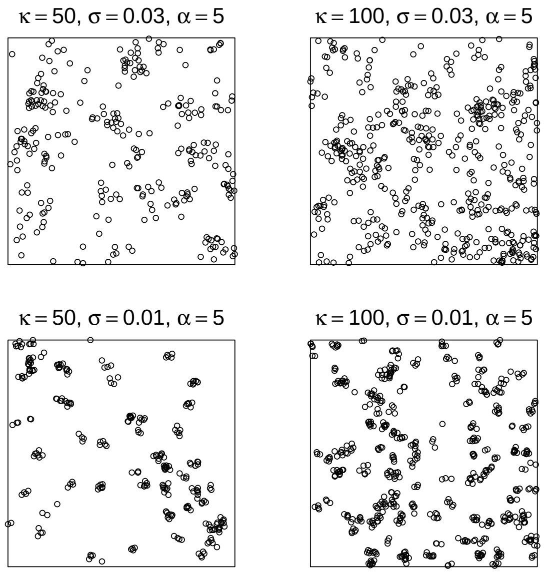


(similar developments can be done for the K, L, J functions and with more work for the Matérn process).

Notes

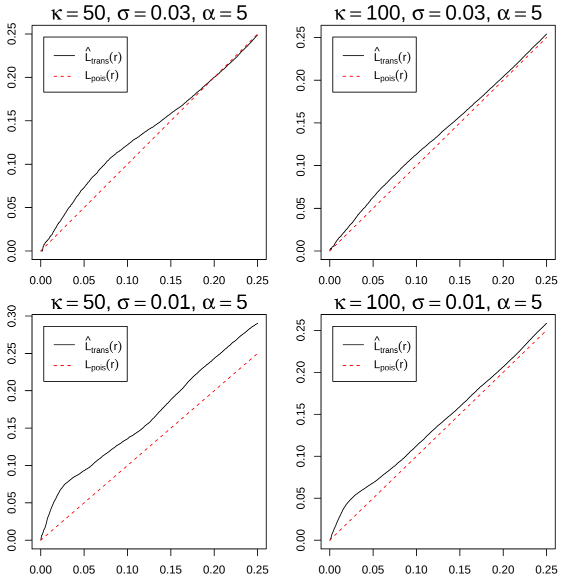
Examples Definitions, Poisson Summary statistics Modelling and inference

Four realizations of Thomas point processes



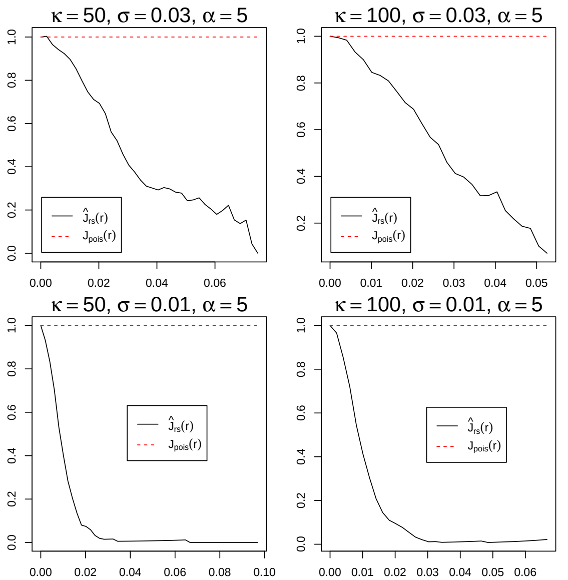
Notes

Correponding L estimates



Notes

Correponding J estimates



Notes

Complements

- Inhomogeneous Neymann-Scott processes can be obtained by replacing the intensity parameter κ by a spatial function $\kappa(u)$.
- The natural extension of NS processes is given by shot-noise Cox processes which is a Cox process driven by

$$Z(u) = \sum_{(c,\gamma)\in\Phi} \gamma k(c,u)$$

where $k(\cdot,\cdot)$ is a kernel and Φ is a Poisson point process on $\mathbb{R}^d \times (0,\infty)$ with a locally integrable intensity function ζ. (see e.g. Møller and Waagepetersen 2004 for complements).

Notes

Log-Gaussian Cox processes

Definition

Let X be a Cox process on $\mathbb{R}^d$ driven by $Z = \exp Y$ where Y is a Gaussian random field. Then, X is said to be a *log Gaussian Cox process* (LGCP).

Remarks :

- we could consider $Z = h(Y)$ for some non-negative function h, but the exp leads to tractable calculations.
- another possibility : using a χ^2 field, i.e. $Z(u) = Y_1(u)^2 + \ldots + Y_m(u)^2$ are the Y_i's are independent Gaussian fields with zero mean.
- LGCP are easy to simulate since the problem is transfered to generate a Gaussian field (which can be handled by several methods).
- The mean and covariance function of Y determine the distribution of X.

Notes

Particular cases

- In the following we let

$$m(u) = \mathrm{E}Y(u) \quad \text{and } c(u,v) = \mathrm{Cov}(Y(u), Y(v))$$

and we focus on the case where $c(u,v)$ depends only on $\|v - u\|$ (covariance function invariant by translation and by rotation).

- Conditions on c are needed to get a covariance function. Among functions satisfying these properties we find :
  - the *power exponential family* satisfies these conditions

$$c(u,v) = \sigma^2 r(\|v-u\|/\alpha) \text{ with } r(t) = \exp\left(-t^\delta\right), t \geq 0$$

  with $\alpha, \sigma > 0$. $\delta = 1$ is the exponential correlation function ; $\delta = 1/2$ is the stable correlation function ; $\delta = 2$ is the Gaussian correlation function.
  - the *cardinal sine correlation* :

$$c(u,v) = \sigma^2 r(\|v-u\|/\alpha) \text{ with } r(t) = \frac{\sin(t)}{t}, t \geq 0$$

Notes

Examples Definitions, Poisson Summary statistics Modelling and inference

Summary statistics for the LGCP

Proposition

Let X be a LGCP then under the previous notation

1. the intension function of X is

$$\rho(u) = \exp\left(m(u) + c(u,u)/2\right).$$

2. The pair correlation function g of X is

$$g(u,v) = \exp(c(u,v)).$$

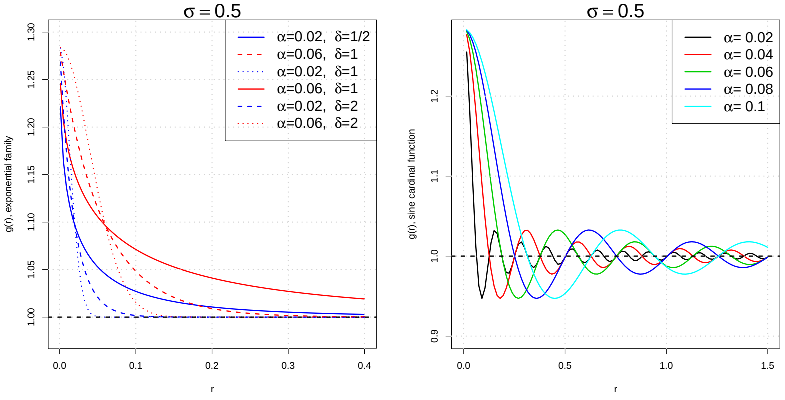
Proof : based on the fact that for $U \sim \mathcal{N}(\zeta, \sigma^2)$, the Laplace transform of U is $\mathrm{E}\exp(tU) = \exp(\zeta + \sigma^2 t/2)$.

- one to one correspondendce between (m,c) and (ρ, g).
- If c is translation invariant then X is second order reweighted stationary (stationary if m is constant, and isotropic if in addition $c(u,v)$ depends only on $\|v-u\|$).

Notes

A few plots of pair correlation function

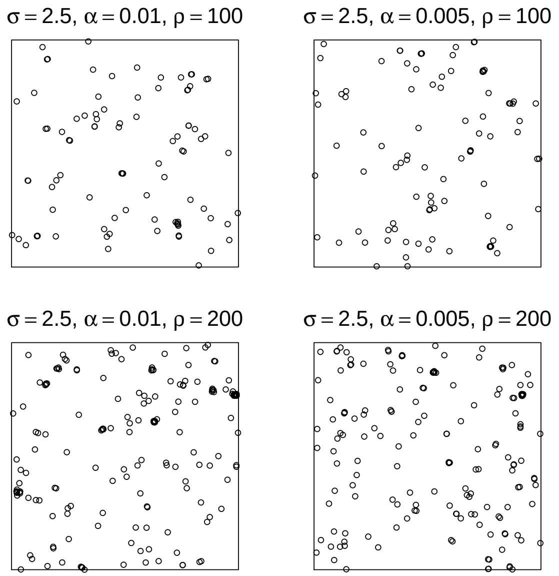
- pcf for the *power exponential family* : $\log g(r) = \sigma^2 \exp\left(-\left(\frac{r}{\alpha}\right)^{\delta}\right), \quad \alpha, \sigma, \delta > 0$
- pcf for the *cardinal sine correlation* : $\log g(r) = \sigma^2 \frac{\sin(r/\alpha)}{r/\alpha}, \quad \alpha, \sigma > 0$



Notes

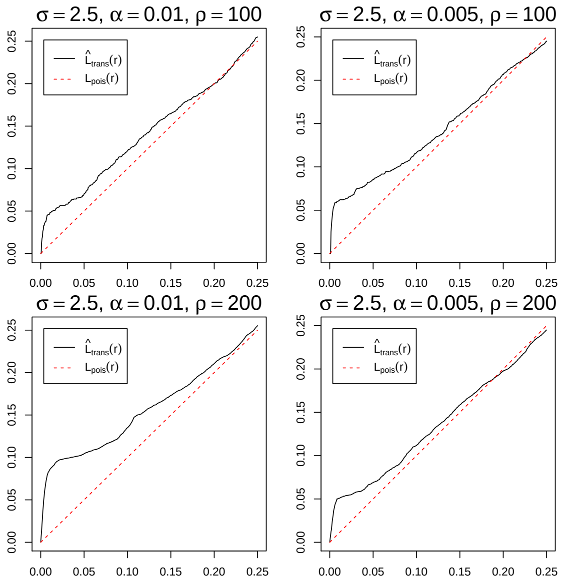
Four realizations of (stationary) LGCP point processes

- with exponential correlation function ($\delta = 1$).
- The mean m of the Gaussian process is such that $\rho = \exp(m + \sigma^2/2)$.



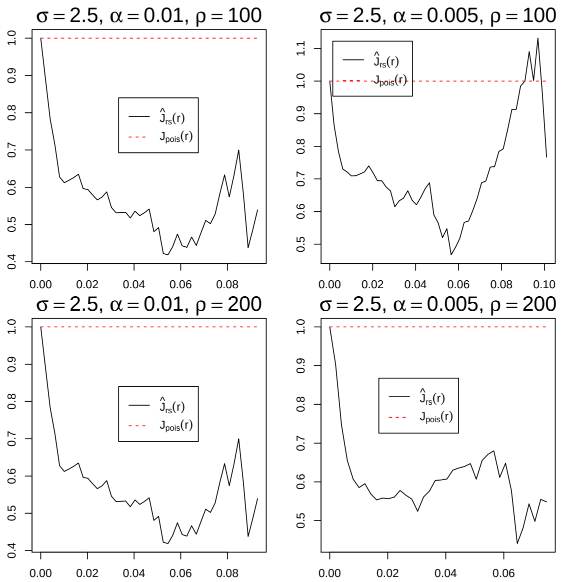
Notes

Correponding L estimates



Notes

Correponding J estimates



Notes

Is likelihood available ?

- Assume (only here) that S is a bounded domain, then the density of X_S w.r.t a Poisson processes with unit rate is given by

$$f(x) = \mathrm{E}\left[\exp\left(|S| - \int_S Z(u)\mathrm{d}u\right)\prod_{u\in x} Z(u)\right]$$

for finite point configurations $x \subset S$. Explicit expression of the expectation is usually unknown and the integral may be difficult to calculate.
$\Rightarrow$ MLE is usually impossible to calculate (approximations or Bayesian should be used)

- In most of applications, we only observe the realization of X.
$\Rightarrow$ Z should be considered as a latent process generating the point process, which is not observed.

Notes

General method based on minimum contrast estimation

- Assume we observe the realization of a stationary Cox point process which belongs to a parametric family with parameter θ (ex : $\theta = (\alpha, \kappa, \sigma^2)$ for the Thomas process, $\theta = (\mu, \alpha, \sigma^2)$ for a LGCP with exponential correlation function).
- For most of Cox point processes, $\rho = \rho_\theta$, $K = K_\theta$ or $g = g_\theta$ functions are expressible in a closed form, for instance :
  - for a planar ($d = 2$) **Thomas process** (NS process with Gaussian kernel) : $\rho = \alpha\kappa$ and

$$g_\theta(r) = 1 + \frac{1}{\sqrt{4\pi\sigma^2}}\exp\left(-r^2/(4\sigma^2)\right)/\kappa \quad \text{and} \quad K_\theta(r) = \pi r^2 + \left(1 - \exp\left(-r^2/(4\sigma^2)\right)\right)/\kappa$$

  - for a **LGCP with exponential correlation function**

$$\rho = \exp(m + \sigma^2/2) \quad \text{and} \quad \log g_\theta(r) = \sigma^2 \exp(-r/alpha).$$

Notes

General method based on minimum contrast estimation (2)

- Then the idea is then to estimate θ using a **minimum contrast approach** : i.e. define $\hat{\theta}$ as the minimizer of

$$\int_{r_1}^{r_2} \left|\widehat{K}(r)^q - K_\theta(r)^q\right|^2 \mathrm{d}r \quad \text{or} \quad \int_{r_1}^{r_2} \left|\widehat{g}(r)^q - g_\theta(r)^q\right|^2 \mathrm{d}r$$

  where
  - $\widehat{K}(r)$ and $\widehat{g}(r)$ are the nonparametric estimates of $K(r)$ and $g(r)$.
  - where $[r_1, r_2]$ is a set of r fixed values.
  - q is a power parameter (adviced in the literature to be set to $q = 1/4$ or $1/2$).

Notes

A short simulation

- we generated 200 replications of a Thomas process with parameters $\kappa = 100$, $\sigma^2 = 10^{-4}$ and $\alpha = 5$
- we estimated the parameters σ^2 and κ using the minimimum contrast estimat based on the K function.
- Then α is estimated using $\widehat{\alpha} = \widehat{\rho}/\widehat{\kappa}$

	Parameter κ	
	$W = [0,1]^2$	$W = [0,2]^2$
Emp. mean	98.9	102.4
Emp. var.	251.9	78.1

	Parameter α	
	$W = [0,1]^2$	$W = [0,2]^2$
Emp. mean	4.9	4.9
Emp. var.	40.1	6.1

	Parameter σ^2	
	$W = [0,1]^2$	$W = [0,2]^2$
Emp. mean	1.01×10^{-4}	9.7×10^{-5}
Emp. var.	1.5×10^{-5}	8.2×10^{-6}

Notes

Introduction

- the objective of this section is to introduce a new class of point processes : the class of Gibbs point processes.
- **Gibbs point process** :
  - are mainly used to model **repulsion** between point (but a few models allows also to produce **aggregated models**). That's why this kind of models are widely used in statistical physics to model particles systems.
  - are defined (in a bounded domain) by a **density** w.r.t. a Poisson point process
    $\Rightarrow$ **very easy** to interpret the model and the parameters.
  - their main drawback : **moments are not expressible** in a closed form and density known up to a scalar
    $\Rightarrow$ **specific inference methods** are required.

Notes

Important restriction of this section

- Throughout this chapter : we assume that the point process X is defined in a bounded domain $S \subset \mathbb{R}^d$ ($|S| < \infty$).
- Gibbs point processes defined on $\mathbb{R}^d$ are of particular interest :
  - in statistical physics because they can model **phase transition** .
  - in asymptotic statistics : if for instance we want to prove the convergence of an estimator as the window expands to $\mathbb{R}^d$

  However, the formalism is more complicated and technical and this is not considered here.

$\Rightarrow$ from now, X is a **finite point process in** S **(bounded)** taking values in N_f (space of finite configurations of points)

$$N_f = \{x \subset S : n(x) < \infty\} .$$

Most of the results presented here have an extension to $S = \mathbb{R}^d$.

Notes

Definition of Gibbs point processes

Definition

A finite point process X on a bounded domain S ($0 < |S| < \infty$) is said to be a Gibbs point process if it admits a density f w.r.t. a Poisson point process with unit rate, i.e. for any $F \subseteq N_f$

$$P(X \in F) = \sum_{n \geq 0} \frac{\exp(-|S|)}{n!} \times \int_S \cdots \int_S \mathbf{1}(\{x_1, \ldots, x_n\} \in F) f(\{x_1, \ldots, x_n\}) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

where the term $n = 0$ is read as $\exp(-|S|)\mathbf{1}(\emptyset \in F)f(\emptyset)$.

- Gpp can be viewed as a perturbation of a Poisson point process.
- f is easily interpretable since it is in some sense a weight w.r.t. a Poisson process.

Notes

The simplest example . . .

is the inhomogeneous Poisson point process. Indeed for $X \sim \text{Poisson}(S, \rho)$ (such that $\mu(S) < \infty$), we recall that X admits a density w.r.t. to a Poisson point process with unit rate given for any $x \in N_f$ by

$$f(x) = \exp(|S| - \mu(S)) \prod_{u \in x} \rho(u).$$

In most of cases, f is specified up to a proportionality $f = c^{-1}h$ where $h : N_f \to \mathbb{R}^+$ is a known function.
$\Rightarrow$ c is given by

$$c = \sum_{n \geq 0} \frac{\exp(-|S|)}{n!} \int_S \cdots \int_S h(\{x_1, \ldots, x_n\}) \mathrm{d}x_1 \ldots \mathrm{d}x_n = \mathrm{E}[h(Y)]$$

where $Y \sim \text{Poisson}(S, 1)$.

Notes

Papangelou conditional intensity

Definition

The Papangelou conditional intensity for a point process X with density f is defined by

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

for any $x \in N_f$ and $u \in S$ $(u \notin x)$, taking $a/0 = 0$ for $a \geq 0$.

- λ does not depend on c.
- for Poisson(S,ρ), $\lambda(u, x) = \rho(u)$ does not depend on x!
- $\lambda(u, x)\mathrm{d}u$ can be interpreted as the conditional probability of observing a point in an infinitesimal region containing u of size $\mathrm{d}u$ given the rest of X is x.

Notes

Attraction, repulsion, heredity

Definition

We often say that X (or f) is

- **attractive** if
$$\lambda(u, x) \leq \lambda(u, y) \text{ whenever } x \subset y.$$
- **repulsive** if
$$\lambda(u, x) \geq \lambda(u, y) \text{ whenever } x \subset y.$$
- **hereditary** if
$$f(x) > 0 \Rightarrow f(y) > 0 \text{ for any } y \subset x.$$

- if f is hereditary, then $f \Leftrightarrow \lambda$ (one-to-one correspondence).

Notes

Existence of a Gpp in S ($|S| < \infty$)

Proposition

Let $\phi^\star : S \to \mathbb{R}^+$ be a function so that $c^\star = \int_S \phi^\star(u)\mathrm{d}u < \infty$. Let $h = cf$, we say that X (or f) satisfies the

- local stability property if for any $x \in N_f$, $u \in S$

$$h(x \cup u) \le \phi^\star(u)h(x) \Leftrightarrow \lambda(u,x) \le \phi^\star(u).$$

- the Ruelle stability property if for any $x \in N_f$ and for $\alpha > 0$

$$h(x) \le \alpha \prod_{u \in x} \phi^\star(u).$$

local stability condition $\Rightarrow$ Ruelle stability condition (and that f is hereditary) $\Rightarrow$ existence of point process in S.

Proof : the first implication is obvious ; for the last one it consists in checking that $c < \infty$.

Notes

Pairwise interaction point processes

For simplicity, we focus on the isotropic case.

Definition

A istotropic parwise interaction point process (PIPP) has a density of the form (for any $x \in N_f$)

$$f(x) \propto \prod_{u \in x} \phi(u) \prod_{\{u,v\} \subseteq x} \phi_2(\|v-u\|)$$

where $\phi : S \to \mathbb{R}^+$ and $\phi_2 : \mathbb{R}^+_* \to \mathbb{R}+$.

- If ϕ is constant (equal to β) then the Gpp is said to be homogeneous (note that $\prod_{u \in x} \phi(u) = \beta^{n(x)}$).
- ϕ_2 is called the interaction function.
- this class of models is hereditary
- f is repulsive if $\phi_2 \le 1$, in which case the process is locally stable if $\int_S \phi(u)\mathrm{d}u$.

Notes

Strauss point process

Among the class of PIPP, the main example is the Strauss point process defined by

$$f(x) \propto \beta^{n(x)} \gamma^{s_R(x)} \qquad \lambda(u,x) = \beta \gamma^{t_R(u,x)}$$

where $\beta > 0, R < \infty$, where $s_R(x)$ is the number of R-close pairs of points in x and $t_R(u,x) = s_R(x \cup u) - s_R(x)$ is the number of R-close neighbours of u in x

$$s_R(x) = \sum_{\{u,v\} \in x} \mathbf{1}(\|v-u\| \le R) \text{ and } t_R(u,x) = \sum_{v \in x} \mathbf{1}(\|v-u\| \le R).$$

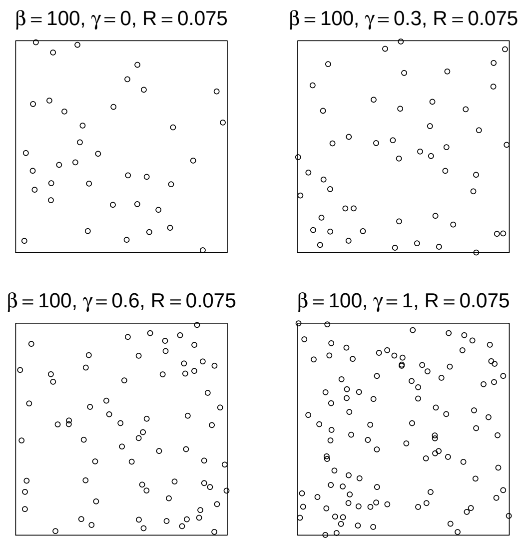
The parameter γ is called the **interaction parameter** :

- $\gamma = 1$: homogeneous Poisson point process with intensity β.
- $0 < \gamma < 1$: repulsive point process.
- $\gamma = 0$: hard-core process with hard-core R ; the points are prohibited from being closer han R.
- $\gamma > 1$: the model is not well-defined (if there exists a set $A \subset S$ with $|A| > 0$ and $diam(A) \le R$, then $c > \sum_{n \ge 0} \frac{(\beta |A|)^n}{n!} \gamma^{n(n-1)/2} = \infty$).

Notes

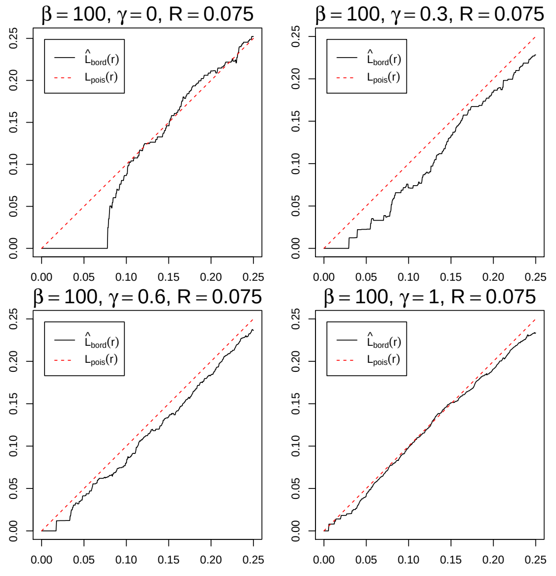
Realizations of Strauss point processes

(simulation of spatial Gibbs point processes can be done using spatial birth-and-death process or using MCMC with reversible jumps, see Møller and Waagepetersen for details)



Notes

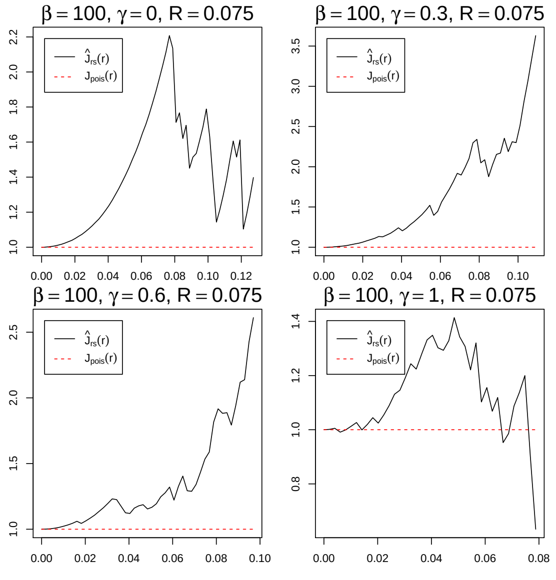
Corresponding L estimates



Notes

Examples Definitions, Poisson Summary statistics Modelling and inference

Corresponding J estimates



Notes

Finite range property (spatial Markov property)

Definition

A Gibbs point process X has a finite range R if the Papangelou conditional intensity satisfies

$$\lambda(u, x) = \lambda(u, x \cap B(u, R)).$$

- the probability to insert a point u into x depends only on some neighborhood of u.
- this definition is actually more general and leads to the definition of Markov point process (omitted here to save time).
- interesting property when we want to deal with edge effects.
- Finite range of the Strauss point process $= R$.

Notes

Other pairwise interaction point processes

- **Strauss** point process : $\phi_2(r) = \gamma^{\mathbf{1}(r \leq R)}$.
- **Piecewise Strauss** point process :

$$\phi_2(r) = \gamma_1^{\mathbf{1}(r \leq R_1)} \gamma_2^{\mathbf{1}(R_1 < r \leq R_2)} \dots \gamma_p^{\mathbf{1}(R_{p-1} < r \leq R)},$$

  with $\gamma_i \in [0, 1]$ and $0 \leq R_1 < \dots < R_p = R < \infty$ (finite range R) .
- **Overlap area** process :

$$\phi_2(r) = \gamma^{|B(u, R/2) \cap B(v, R/2)|},$$

  with $r = \|v - u\|$ with $\gamma \in [0, 1]$ (finite range R) .
- **Lennard-Jones** process :

$$\phi_2(r) = \exp(\alpha_1(\sigma/r)^6 - \alpha_2(\sigma/r)^{12}),$$

  with $\alpha \geq 0$, $\alpha_2 > 0$, $\sigma > 0$ (well-known example used in statistical physics, not locally stable but Ruelle stable) (infinite range) .

Notes

Non pairwise interaction point processes

- **Geyer's triplet** point process :

$$f(x) \propto \beta^{n(x)} \gamma^{s_R(x)} \delta^{u_R(x)}$$

$\beta > 0$, $s_R(x)$ is defined as in the Strauss case and

$$u_R(x) = \sum_{\{u,v,w\}} \mathbf{1}(\|v-u\| \leq R, \|w-v\| \leq R, \|w-u\| \leq R)$$

  - (i) $\gamma \in [0,1]$ and $\delta \in [0,1]$: locally stable, repulsive, finite range R.
  - (ii) $\gamma > 1$ and $\delta \in (0,1)$: locally stable, neither attractive nor repulsive, finite range R.

Notes

Non pairwise interaction point processes (2)

- **Area-interaction** point process :

$$f(x) \propto \beta^{n(x)} \gamma^{-|U_{x,R}|}$$

where $U_{x,R} = \cup_{u \in x} B(u,R)$, $\beta > 0$ and $\gamma > 0$. It is attractive for $\gamma \geq 1$ and repulsive for $0 < \gamma \leq 1$. In both cases, it is locally stable since

$$\lambda(u,x) = \beta \gamma^{-|B(u,R) \setminus \cup_{v \in x : \|v-u\| \leq 2R} B(v,R)|}$$

satisfies $\lambda(u,x) \leq \beta$ when $\gamma \geq 1$ and $\lambda(u,x) \leq \beta \gamma^{-\omega_d R^d}$ in the other case. (finite range $2R$)

Notes

GNZ formula

The following result is also a characterization of a Gibbs point process.

Georgii-Nguyen-Zeissin Formula

Let X be a finite and hereditary Gibbs point process defined on S. Then, for any function $h : S \times N_f \to \mathbb{R}^+$, we have

$$\mathrm{E}\Big[\sum_{u \in X} h(u, X \setminus u)\Big] = \int_S \mathrm{E}[h(u,X)\lambda(u,X)]\mathrm{d}u.$$

Proof : we know that $\mathrm{E}g(X) = \mathrm{E}[g(Y)f(Y)]$ where f is the density of a Poisson point process with unit rate Y. Apply this to the function $g(X) = \sum_{u\in X} h(u, X\setminus u)$

$$\begin{aligned}
\mathrm{E}[g(X)] &= \mathrm{E}\big[\sum_{u\in Y} h(u, Y\setminus u) f(Y)\big] \\
&= \int_S \mathrm{E}[h(u,Y)f(Y\cup u)]\mathrm{d}u \quad \text{from the Slivnyak-Mecke Theorem} \\
&= \int_S \mathrm{E}[h(u,Y)f(Y)\lambda(u,Y)]\mathrm{d}u \quad \text{since } X \text{ is hereditary} \\
&= \int_S \mathrm{E}[h(u,X)\lambda(u,X)]\mathrm{d}u.
\end{aligned}$$

Notes

First and second order intensities

Proposition

1. The intensity function is given by

$$\rho(u) = \mathrm{E}[\lambda(u,X)].$$

2. The second order intensity function is given by

$$\rho^{(2)}(u,v) = \mathrm{E}[\lambda(u,X)\lambda(v,X)]$$

- can be deduced from the GNZ formula.
- Except for the Poissonian case, moments are not expressible in a closed form, e.g.

$$\rho(u) = \frac{1}{c}\sum_{n\geq 0} \frac{\exp(-|S|)}{n!} \int_S \cdots \int_S \lambda(u, \{x_1,\ldots,x_n\}) h(\{x_1,\ldots,x_n\})\mathrm{d}x_1 \ldots \mathrm{d}x_n.$$

- Approximations can be obtained using a Monte-Carlo approach or using a saddle-point approximation (very recent).

Notes

Position of the problem

- we observe a realization of X on $W = S$ ($|S| < \infty$; edge effects occur when $W \subset S$) of a parametric Gibbs point process with density which belongs to a parametric family of densities $(f_\theta = h_\theta / c_\theta)_{\theta \in \Theta}$ for $\Theta \subset \mathbb{R}^p$.
- Problem : estimate the parameter θ based on a single realization.
- *MLE approach* : the log-likelihood is $\boxed{\ell_W(x;\theta) = \log h_\theta - \log c_\theta}$.
  Pbm : Given a model h_θ can be computed but c_θ cannot be evaluated even for a single value of θ ; asymptotic properties are only partial.
  $\Rightarrow$ several solutions exist
  1. Approximate c_θ using a Monte-Carlo approach.
  2. Bayesian approach, importance sampling method (to estimate a ratio of normalizing constants).
  3. Combine the MLE with the Ogata-Tanemura approximation.
  4. Find another method which does not involve c_θ.

Notes

Pseudo-likelihood

- To avoid the computation of the normalizing constant, the idea is to compute a likelihood based on conditional densities

$$PL_W(x;\theta) = \exp(-|W|) \lim_{i\to\infty} \prod_{j=1}^{m_i} f(x_{A_{ij}} | x_{W \setminus A_{ij}};\theta)$$

  where $\{A_{ij} : j = 1, \ldots, m_i\}$ $i = 1, 2, \ldots$ are nested subdivisions of W.
- By letting $m_i \to \infty$ and $m_i \max |A_{ij}|^2 \to 0$ as $i \to \infty$ and taking the log, Jensen and Møller (91) obtained

$$LPL_W(x;\theta) = \sum_{u \in x_W} \lambda(u, x \setminus u; \theta) - \int_W \lambda(u, x; \theta) \mathrm{d}u$$

Notes

Comments on the Pseudo-likelihood

The MPLE is the estimate $\widehat{\theta}$ maximizing

$$LPL_W(x;\theta) = \sum_{u \in x_W} \log \lambda(u, x \setminus u; \theta) - \int_W \lambda(u, x; \theta) \mathrm{d}u$$

1. **Independent on c_θ,** so the LPL is up to an integral discretization and up to edge effects very to compute.
2. If X has a finite range R, then since x is observed in W, we can replace W by $W_{\ominus R}$ so that for instance $\lambda(u, x; \theta)$ can always be computed for any $u \in W_{\ominus R}$ **(border correction)**.
3. If $\log \lambda(u, x; \theta) = \theta^\top v(u, x)$ (exponential family - class of all examples presented before), then LPL is a **concave** function of θ.
4. under suitable conditions $\widehat{\theta}$ is a **consistent** estimate and satisfies a **CLT** (and a fast covariance estimate is available) as the window W expands to $\mathbb{R}^d$. [Jensen and Künsch'94, Billiot Coeurjolly and Drouilhet'08-'10, Coeurjolly and Rubak'12].

Notes

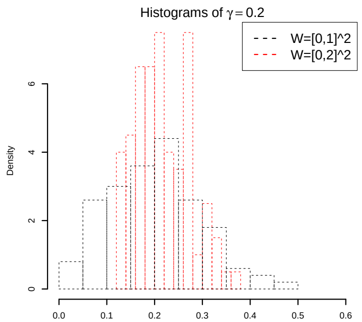
Simulation example

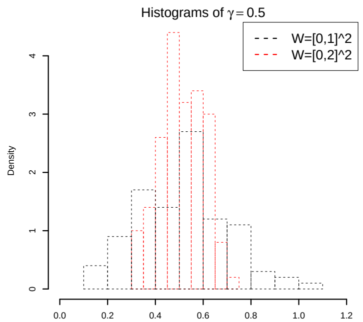
We generated 100 replications of Strauss point processes (a border correction was applied) :

1. mod1 : $\beta = 100$, $\gamma = 0.2$, $R = .05$.
2. mod2 : $\beta = 100$, $\gamma = 0.5$, $R = .05$.

	Estimates of β			
	$W = [0,1]^2$		$W = [0,2]^2$	
mod1	99.52	(17.84)	97.98	(9.24)
mod2	99.28	(20.48)	98.21	(8.53)

	Estimates of γ			
	$W = [0,1]^2$		$W = [0,2]^2$	
mod1	0.20	(0.09)	0.21	(0.06)
mod2	0.52	(0.19)	0.51	(0.09)





Notes

Takacs-Fiksel method

- Denote for any function h (eventually depending on θ)

$$L_W(X,h;\theta) = \sum_{u \in X_W} h(u, X \backslash u; \theta) \text{ and } R_W(X,h;\theta) = \int_W h(u,X;\theta)\lambda(u,X;\theta)\mathrm{d}u$$

- The GNZ formula states : $\mathrm{E}[L_W(X,h;\theta)] = \mathrm{E}[R_W(X,h;\theta)]$.
- **Idea** : if θ is a p-dimensional vector,
  1. choose p test function h_i and define the contrast

$$U_W(X,\theta) = \sum_{i=1}^{p} \left(L_W(X,h;\theta) - R_W(X,h;\theta)\right)^2.$$

  2. Define $\widehat{\theta}^{TF} = \mathrm{argmin}_\theta U_W(X,\theta)$.

Notes

Takacs-Fiksel (2)

General comments :

- **like the MPLE** :
  - independent of c_θ, border correction possible in case of X has a finite range
  - consistent and asymptotically Gaussian estimate (Coeurjolly et al.'12).
- **Another advantage** : interesting choices of test functions cal least to a decreasing of computation time.
  Ex : $h_i(u,X) = n(B(u,r_i))\lambda^{-1}(u,X;\theta) \Rightarrow R_W$ independent of θ.
- Actually : **MPLE = TFE** with $h = (h_1,\dots,h_p)^\top = \lambda^{(1)}(\cdot,\cdot;\theta)$.
  Indeed (assume $\log \lambda(u,X;\theta) = \theta^\top v(u,X)$ (for simplicity)

$$\nabla LPL_W(X;\theta) = \sum_{u \in X_W} v(u, X \setminus u) - \int_W v(u,X)\lambda(u,X;\theta)\mathrm{d}u.$$

Notes

A funny example for the Strauss point process

Recall that the Papangelou conditional intensity of a Strauss point process is

$$\lambda(u,X) = \beta\gamma^{t_R(u,X)} \text{ with } t_R(u,X) = \sum_{v\in X} \mathbf{1}(\|v-u\| \le R).$$

Choose $h_1(u,X) = \mathbf{1}(n(B(u,R)=0))$ and $h_2(u,X) = \mathbf{1}(n(B(u,R)=1))$, then

- $L_W(X,h_1) = L_1$ and $R_W(X,h_1) = \beta\int_W \mathbf{1}(n(B(u,R)=0)) = \beta I_1$.
- $L_W(X,h_2) = L_2$ and $R_W(X,h_2) = \beta\gamma\int_W \mathbf{1}(n(B(u,R)=1)) = \beta I_2$.

Then, the contrast function rewrites

$$U_W(X) = (L_1 - \beta I_1)^2 + (L_2 - \beta\gamma I_2)^2$$

which leads to the **explicit** solution

$$\widehat{\beta} = \frac{L_1}{I_1} \quad \text{and } \widehat{\gamma} = \frac{L_2}{I_2} \times \frac{I_1}{L_1}.$$

Notes

Complements

Other parametric approaches :

- Variational approach : (Baddeley and Dereudre'12).
- Method based on a logistic regression likelihood (Baddeley, Coeurjolly, Rubak, Waagepetersen'13).

Model fitting :

- Monte-Carlo approach : we can compare a summary statistic e.g. L with $L_{\widehat{\theta}}$.
  Pbm : L_θ not expressible in a closed form and must be approximated.
- We can still use the GNZ formula : given a test function h, we can construct

$$L_W(X,h;\widehat{\theta}) - R_W(X,h;\widehat{\theta}) =: \mathrm{Residuals(X,h)}.$$

If the model is correct, then $\mathrm{Residuals(X,h)}$ should be close to zero. (Baddeley et al.'05,08', Coeurjolly and Lavancier'12).

Notes

General Conclusion

The anaysis of **spatial point pattern**

- very large domain of research including probability, mathematical statistics, applied statistics
- own specific models, methodologies and software(s) to deal with.
- is involved in more and more applied fields : economy, biology, physics, hydrology, environmetrics,. . .

Still a lot of **challenges**

- Modelling : the “true model”, problems of existence, phase transition.
- Many classical statistical methodologies need to be adapted (and proved) to s.p.p. : robust methods, resampling techniques, multiple hypothesis testing.
- High-dimensional problems : $S = \mathbb{R}^d$ with d large, selection of variables, regularization methods,. . .
- Space-time point processes.

Notes

Notes