Point Pattern Analysis: Processes (2)

SERGIO REY

GPH 483/598
Geographic Information Analysis
School of Geographical Sciences
Arizona State University
Spring 2010

Outline

1. Clustered Patterns
   - Neyman Scott
   - Inhomogeneous Poisson Process

2. Regular Pattern
   - Inhibition: Matern Processes
   - Markov Point Processes

Clustered Pattern

More Grouped than CSR

- some higher densities, aggregated
- many points at shorter distances

Overdispersion

- variance > mean
- greater variation in densities than CSR

Sources of Clustering

Contagion

- presence of events at x affects probability of event at y
- correlated point processes

Heterogeneity

- intensity $\lambda(s)$ varies with s
- heterogeneous Poisson point process

Contagious Distributions

Two stages

- point pattern for parents
- point pattern for offspring centered on parent locations
- parents may or may not be included

Examples

- Poisson cluster process (Neyman-Scott)
- Matern cluster process

Poisson Cluster Process

Parent Events

- Poisson process with intensity λ

Number of Offspring Events S

- identical distribution for each parent
- $E[S] = \mu$

Location of Offspring Events

- independent and identically distributed
- following a bivariate density h

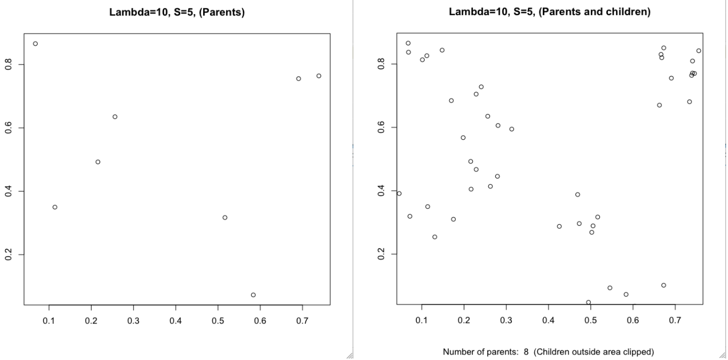
Example

Parent Process Poisson

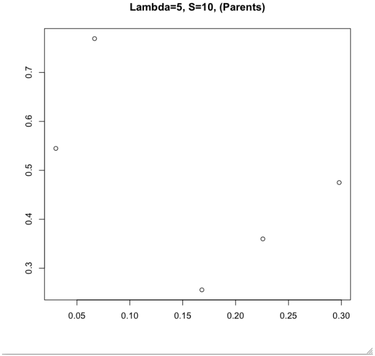
- homogeneous, intensity λ constant

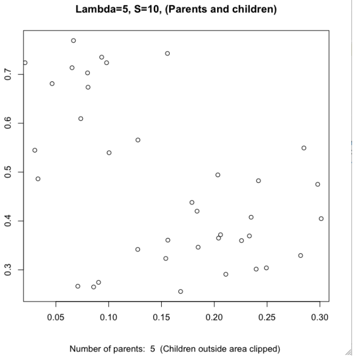
Child Process

- uniform points in circle centered on parent
- fixed number of points in circle centered on parent
- points outside window eliminated


Lambda=10, S=5, (Parents)
0.2
0.4
0.6
0.8
0.1 0.2 0.3 0.4 0.5 0.6 0.7
Lambda=10, S=5, (Parents and children)
0.2
0.4
0.6
0.8
0.1 0.2 0.3 0.4 0.5 0.6 0.7
Number of parents: 8 (Children outside area clipped)

Neyman Scott $\lambda = 5, S = 10$





Inhomogeneous Poisson Process

Implications

- Apparent clusters can occur solely due to heterogeneities in the intensity function $\lambda(s)$.
- Individual event locations still remain independent of one another.
- Process is not stationary due to intensity heterogeneity

HPP vs. IPP

HPP is a special case of IPP with a constant intensity

CSR vs. Constant Risk Hypotheses

CSR

- Intensity is spatially constant
- Population at risk assumed spatially uniform
- Useful null hypothesis if these conditions are met

Constant Risk Hypothesis

- Population density variable
- Individual risk constant
- Expected number of events should vary with population density
- Clusters due to deviation from CSR
- Clusters due to deviation from CSR and Constant Risk

Inhomogeneous Poisson Process

Non-Stationary

- spatially varying intensity $\lambda(s)$
- mean is $\int_A \lambda(s)ds$
- an integral of the location-specific intensities over the region

Properties

- $N(A) \sim Poi(\int_A \lambda(s)ds$
- $N(A) = n, \ n$ events independent sample with pdf proportional to $\lambda(s)$

Sources of Variability

Deterministic

- function for variability of $\lambda(s)$
- introduce covariates: $\lambda(s) = f(z)$

Stochastic

- doubly stochastic process
- distribution for $\lambda(s) \sim \Lambda(s)$

Examples

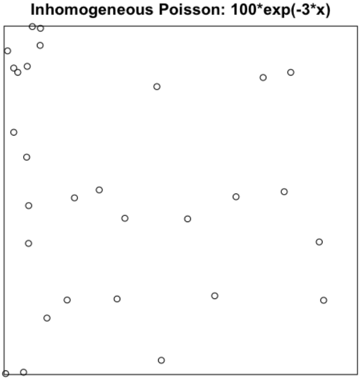
Intensity Varies with a Covariate

- trend surface
- $\lambda(s) = exp(\alpha + \beta s)$

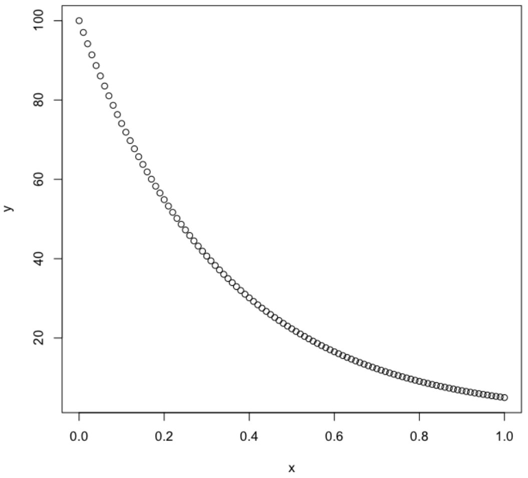
Intensity Varies with Distance to Focus

- $\lambda(s) = \lambda 0(s).f(||s - s_0||, \theta)$

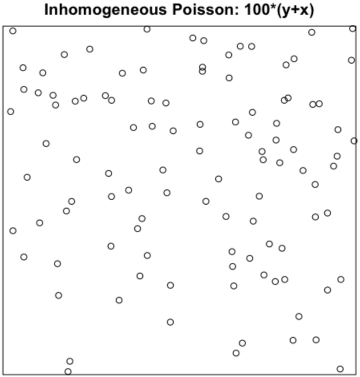
Inhomogeneous Poisson Process: $\lambda(x, y) = 100 * exp(-3x)$



Inhomogeneous Poisson Process: $\lambda(x, y) = 100 * exp(-3x)$



Inhomogeneous Poisson Process: $\lambda(x, y) = 100 * (x + y)$



Thinning

From Homogeneous to Heterogeneous

- remove points

Types

- *p*-thinning: constant probability
- $p(s)$-thinning: probability varies with s
- Π-thinning: thinning function is random

Simulation

Start with homogeneous Poisson

- $\lambda = max[\alpha(s)]$

Apply $p(s)$ Thinning

- keep points with probability $p(s)$
- $p(s) = \alpha(s)/\lambda$
- e.g., keep if generated uniform random number < $p(s)$

Cox Process

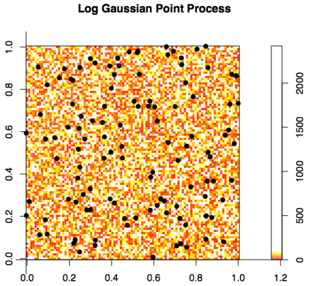
Doubly Stochastic Process

- $\Lambda(s)$ is stochastic process over A
- events inhomogeneous Poisson process with $\lambda(s) = \Lambda(s)$ (a realization)

Log-Gaussian Process

- $\Lambda(s) = exp[Z(s)]$ with $Z(s) \sim N(\mu, \sigma^w)$
- $E[\lambda] = exp(\mu + 0.5\sigma^2)$

Cox process



$\ln Z \sim N(4.1, 1)$ $E[\lambda] \approx 100$ $\lambda = 113$

Identification

Inverse problem

- identify process from pattern

True Contagion - Apparent Contagion

- impossible to distinguish contagious process from heterogeneous process

Identification

Bartlett Equivalence

- Cox process (heterogeneity) and Poisson Cluster process (contagion) yield equivalent patterns

Identification Strategies

- repeated observation, covariates
- heterogeneous in same location, contagious not

Regular Pattern

Less Grouped than CSR

- fewer high densities, empty space
- dispersed
- repulsion, competition

Underdispersion

- variance < mean
- less variation in densities than CSR

Inhibition Process

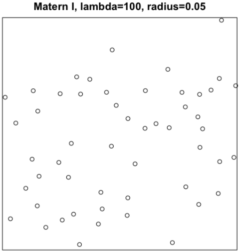
Minimum Permissible Distance

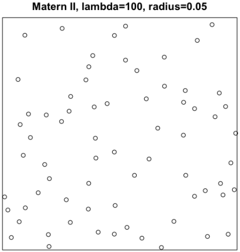
- no two points closer than δ
- packing intensity $\tau = \lambda \pi \delta^2 / 4$

Matern Process

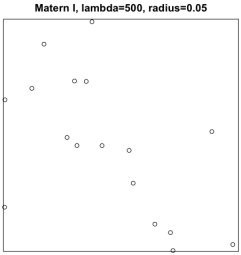
- I: thinned Poisson process using δ
- II: sequential inhibition process, generates points conditional on previous points and distance (denser than I)

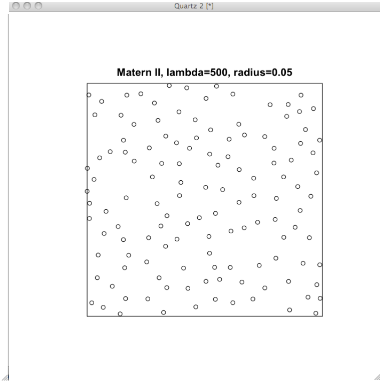
Matern I and II $\lambda = 100$





Matern I and II $\lambda = 500$





Markov Point Processes

Allow for Interaction

- neighbors $||x - y|| < \delta$, range of process
- likelihood relative to Poisson with $\lambda = 1$
- $f(x) = exp(-|A|)$

Strauss Process

- $f(x) = \alpha\beta^n\gamma^s$
- β intensity, n number of points
- γ interactions, $\gamma < 1$ is inhibition, s pairs

Interaction Processes

Pairwise Interaction Point Processes

- $f(x) = \alpha\beta^n \prod_i \prod_{j \neq i} h(||x_i - x_j||)$
- $h = 0$ for $\delta > 0$

Area Interaction Point Processes

- $f(x) = \alpha\beta^n \alpha^{-A(X,\delta)}$
- $\delta = 1$: Poisson
- $\delta < 1$: inhibition
- $\delta > 1$: aggregated