Alternate Models to CSR

Bailey and Gatrell

Chapter 3

Lecture 10

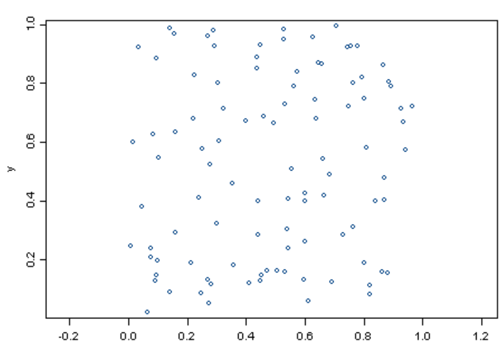
October 6, 2009

1

Alternate Models to CSR

To explain a particular pattern of clustering or regularity we need to find alternate models to CSR

2



Example of a CSR or homogeneous Poisson process

3

Homogeneous Poisson Process

Provides a benchmark of CSR against which patterns can be tested

- Homogeneous means that λ is constant over R
- There is spatial independence in the occurrence of the events – no correlation in the number of events in neighboring regions – there is no interaction (repulsively (regularity) or attractively (clustering).

If there is a violation of either of these properties and we reject CSR as a possible model

- we need to investigate alternate models to explain a particular clustering or regularity in the observed data

4

Alternate Models to CSR

- **For clustered patterns**

  First order effects only:
  - Heterogeneous Poisson Process
  - Cox Process

  Second order effects only:
  - Poisson Cluster Process
- **For regular patterns**
  - Simple Inhibition Process
- **Either**
  - Markov Point Processes

5

Heterogeneous Poisson Process

- A constant intensity λ is replaced by spatially variable intensity function λ(s) but any event remains independent of any other event
- A simple non-stationary point process
- The process has a Poisson distribution with mean

$$\int_A \lambda(s)ds$$

- Given n events in A they form an independent random sample from the distribution on A with PDF proportional to λ(s)

6

Simulation of Heterogeneous Poisson Process

$$\lambda(x_1, y_1) = \exp(-2x_1 - y_1)$$

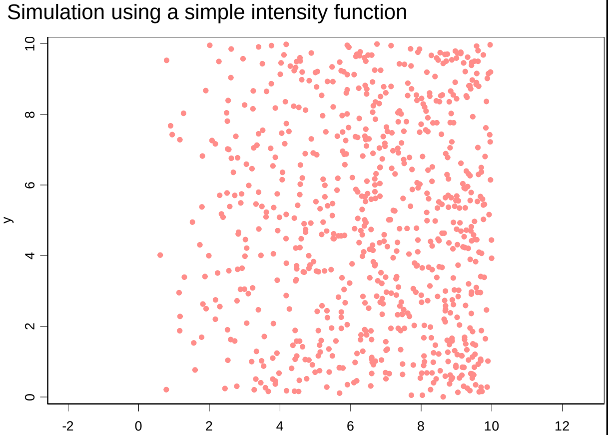
x	y	λ
0.1	0.2	0.67032
0.2	0.2	0.548812
0.2	0.3	0.496585
0.3	0.1	0.496585
0.04	0.2	0.755784
0.5	0.1	0.332871
0.2	0.6	0.367879
0.4	0.7	0.22313
0.1	0.8	0.367879
0.8	0.1	0.182684
0.05	0.5	0.548812

Simulate CSR on R with intensity λ_0 equal to maximum intensity within R

Retain an event at x with probability $\lambda_{(x)} / \lambda_0$

7

Simulation using a simple intensity function



$\lambda(s) = 1.5x$

8

Heterogeneous Poisson Process

A heterogeneous Poisson process provides a framework to include covariates into the analysis through an intensity function

$$\lambda(x) \equiv \lambda\{z_1(x), z_2(x), \ldots z_p(x)\}$$

Where covariates for example might be pH, soil fertility, or elevation where the events are trees

Variables suspected of influencing the intensity of the process

$\lambda(x) = \exp\{\alpha + \beta z(x)\}$ Where z(x) might be elevation above sea level at location x

Referred to as a modulated Poisson process

9

Cox Processes

A heterogeneous Poisson process with intensity function λ(x) can produce apparent clusters of events in regions of relatively high intensity.

Cox processes assume a source of aggregation (clustering) is environmental heterogeneity and that this environmental heterogeneity can be stochastic in nature.

Referred to as class of “doubly stochastic processes”

10

Cox Processes

Cox process is formed as an heterogeneous Poisson process with a stochastic intensity function.

$\{\Lambda(x) : x \in \mathrm{R}^2\}$ a non-negative valued stochastic process

Conditional on $\{\Lambda(x) = \lambda(x) : x \in \mathrm{R}^2\}$ events form an heterogeneous Poisson process with intensity function λ(x)

The point process is stationary if and only if the intensity process is stationary, and similarly for isotropy

In the stationary case the intensity $\lambda = E[\Lambda(x)]$

11

Poisson Cluster Process

- Incorporates an explicit spatial clustering mechanism
- Parent events form a Poisson process with intensity ρ
- Each parent produces a random number of offspring S realized independently and identically for each parent according to a probability distribution $\{p_s, s = 0,1,\ldots\}$
- The position of the offspring relative to their parents are independently and identically distributed according to a bivariate PDF g()

$$g(x, y) = (2\pi\sigma^2)^{-1} \exp\{-(x^2 + y^2)/2\sigma^2\}$$

- Final pattern consists of offspring only

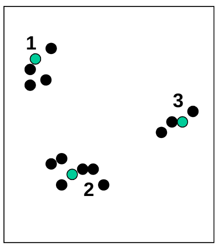
12

Simulation of Poisson Cluster Process

1. Simulate parents as Poisson process with intensity ρ

2. For each parent independently simulate number of offspring according to f()

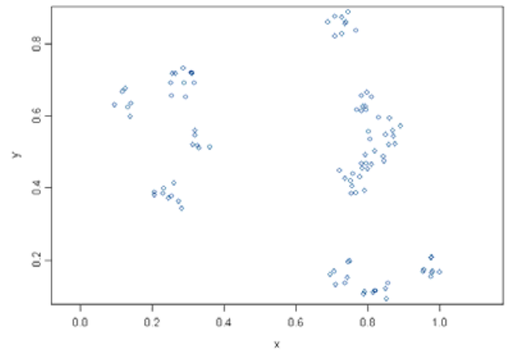
Parent 1	4
Parent 2	6
Parent 3	3



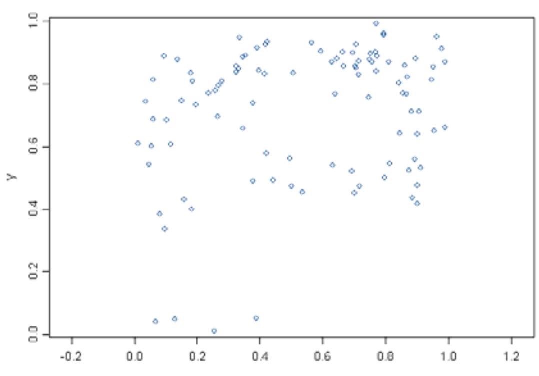
3. Independently locate the offspring around a parent according to g()

g() is a bivariate probability density function)

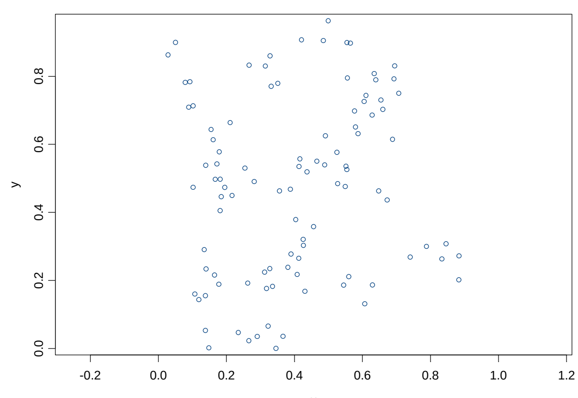
Final pattern consists of offspring only



Simulated Poisson cluster process, n= 100, σ = .05, ρ = 15 (ρ = intensity of parent process)



Simulated Poisson cluster process, n= 100, σ = .15, ρ = 15



plot(make.pattern(100,process="cluster", radius =.10, cpar = 20)

Poisson Cluster Process

Poisson cluster processes are stationary with intensity

$$\lambda = \rho\mu,\ \mu = E[S]$$ Where S is the random number of offspring

They are isotropic if the bivariate probability distribution g() is radially symmetric

17

Poisson Cluster Process

Second order properties

probability density function of the vector difference between the positions of 2 offspring from the same parent

$$g_2(x) = \int g(x)g(x-z)dx$$

G_2() is corresponding cumulative distribution function

The expected number of ordered pairs of such offspring is E[(S(S-1)]

The expected number of related events within distance h of an arbitrary event is E[(S(S-1)] G_2(h) /μ

The expected number of unrelated events (events from different clusters) within distance h of an arbitrary event is $\lambda\pi h^2$

18

Summing the contributions from related and unrelated events gives:

$$\lambda K(h) = \lambda\pi h^2 + E[(S(S-1)]G_2(h)/\mu$$

Divide by $\lambda = \rho\mu$

$$K(h) = \pi h^2 + E[(S(S-1)]G_2(h)/\rho\mu^2$$

Gives the K function for a Poisson cluster model

19

Process ambiguity

In certain cases the Poisson cluster process and certain Cox processes are statistically indistinguishable.

If the offspring distribution f() in the Poisson cluster model is a Poisson distribution it can be shown that it is possible to have a probability distribution for λ(s) in the Cox process that will produce exactly the same effect as the cluster process for any given g()

Although the interpretations are different: one being first order heterogeneous and second order independent and the other having stationary second order effects - dependence – they can not be distinguished statistically

20

Simple Inhibition Process

Regular patterns arise by imposing some minimum permissible distance δ between any two events

A CSR process of intensity ρ is thinned by the deletion of all pairs of events a distance less than δ apart. The probability that an arbitrary event survives is therefore

$$\exp(-\pi\rho\delta^2)$$

The intensity is $\lambda = \rho\exp(-\pi\rho\delta^2)$

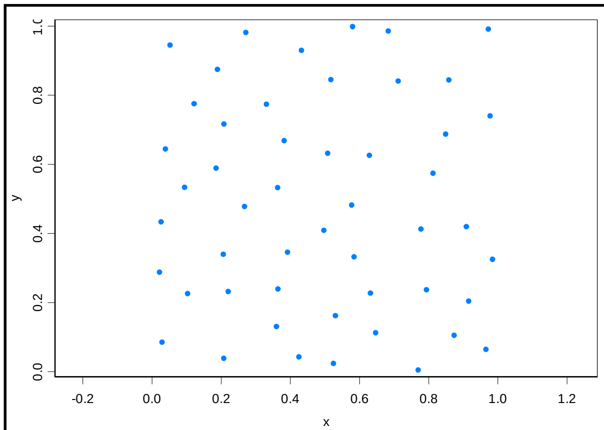
21

Simple Inhibition Process

Alternate version

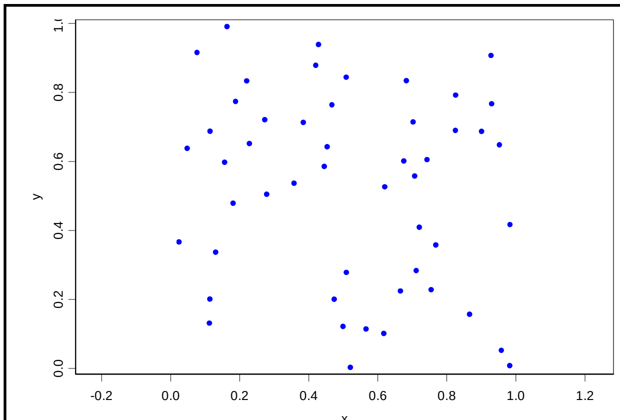
Generate a CSR process one point at a time, discard points if they lie within distance δ of any previously retained point

22



Simple inhibition process with δ = .1 on unit square

23



Simple inhibition process with δ = .05 on unit square

24

Markov Point Processes

Alternative to simple inhibition

Strict inhibition is not very flexible since it could be unlikely but not impossible that two events could be in close proximity

A Markov point process is characterized by its likelihood ratio f(.) with respect to a Poisson process of unit intensity.

If $S = \{s_1, \ldots s_n\}$ denotes a finite set of events in A then f(S) indicates essentially how much more likely is the configuration of events S for the particular process than for a Poisson process with unit intensity.

25

Markov Point Processes

Straus Process – pairwise interaction process

Points considered neighbors if less than δ distance apart

Joint density function for n point locations in R which contain m distinct pairs of neighbors

$$f(s_1, \cdots, s_n) = \alpha \beta^n \gamma^m \quad \beta > 0 \qquad 0 \le \gamma \le 1$$

α is normalizing constant

β is intensity of the process

n is the number of events

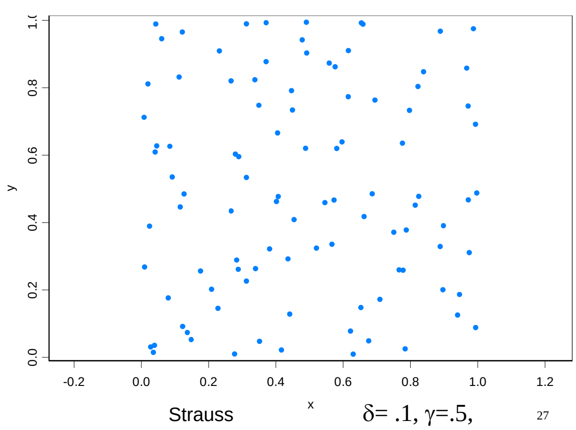
m is the number of distinct pairs of neighbors

γ describes interaction among neighbors

γ=1 yields Poisson process with intensity β

γ=0 yields a simple inhibition process

26



Strauss δ= .1, γ=.5,

27

Fitting Models other than CSR

Need to estimate unknown parameters and assess the goodness of fit

Methods of parameters estimation can employ K(h) and its estimator or similarly G(w) and F(x)

Several approaches use K(h) since the theoretical form of K is known for a number of classes of spatial point processes.

Plots of K(h) can be used to suggest statistically plausible models and provide initial parameter estimates

28

Least Squares Estimation

Suppose a model incorporates a vector of parameters θ.

Let K(h;**θ**) denote the theoretical K-function and $\hat{K}(h)$ the estimator calculated from the data.

To model the discrepancy between the model and data define the discrepancy to be

$$D(\theta)=\int_0^{h_0}\left[\left\{\hat{K}(h)\right\}^c-\left\{K(h;\theta)\right\}^c\right]^2 dh$$

Estimate θ to be the value $\hat{\theta}$ that minimizes $D(\theta)$

29

Least Squares Estimation

Choose values for tuning constants c and h_0

From empirical tests, suggested values for c are .5 for regular patterns and .25 for clustered patterns and .25 for h_0

Assume a Poisson cluster process for which the theoretical K function is:

$$K(h)=\pi h^2+\frac{(1-\exp\left\{-h^2/4\sigma^2\right\}}{\rho}$$

ρ is the mean number of parents per unit area and σ is the dispersion parameter of the radially symmetric normal distribution for the offspring.

Distribution of distance between offspring from same parent $G_2(h) = 1-\exp(-h^2/4\sigma^2)$

30

Least Squares Estimation

In this case the unknown parameters that need to be estimated are

$$\theta=(\rho,\sigma)^T$$

These are estimated by minimizing D

Goodness of Fit

Monte Carlo simulations similar to those used to check the fit to the CSR model can be used

Simulate an appropriate model in place of CSR

31