75563 Spatial Statistics,
Assignment 2:Point processes

Exercise 1: Inhomogeneous Poisson process

We will study an inhomogeneous Poisson process at the unit square. The process has the intensity function:

$$\lambda_\theta(x, y) = \mu \exp(-\alpha x - \beta y)$$

where $\theta = (\mu, \alpha, \beta)$.

a. The number of points in the unit square is a random variable. Compute its mean and variance.

b. Simulate the inhomogeneous Poisson process conditioned to the number of points in the unit square. Let the intensity function be as given with $\theta = (200, 2.0, 1.0)$.

Vary the number of points in the unit square. (f.ex: $n = (20, 50, 100)$).

Splus : Use `par(pty=''s'')` to plot at a square.

c. Expand the algorithm in b, to simulate unconditional to the number of points and do the simulation.

d. An alternative algorithm to sample the inhomogeneous Poisson process,is based on p - thinning.

To simulate at the unit square do the following.

1. put
$$\lambda^* = \max_{(x,y)\in[0,1]^2} \lambda_\theta(x, y)$$
2. Simulate a homogeneous point process with intensity λ^*.
3. Keep a simulated point (x_n, y_n) with probability
$$p = \frac{\lambda_\theta(x_n, y_n)}{\lambda^*}$$

Implement this algorithm as well, and do the simulation. What is the main difference of the two algorithms. Generally speaking, what is the pros and the cons of this algorithm compared to the algorithm you used in c.

e. Identify the equations used to find the maximum likelihood estimate of $\theta = (\mu, \alpha, \beta)$ in the model above. The likelihood is given by:

$$L(x_1, y_1 \ldots, x_n, y_n; \theta) = \lambda_\theta(x_1, y_1) \ldots \lambda_\theta(x_n, y_n) \exp[-\int_0^1 \int_0^1 \lambda_\theta(x, y) dx dy],$$

at the unit square.

f. Choose a point pattern from the unconditional simulation (c or d). Determine the maximum likelihood estimates of $\theta = (\mu, \alpha, \beta)$ based on that data set. Plot the estimated intensity with the function `image`.

g. Copy the files intest.s and integral.o. from `/local/kurs/romstat` . This is code for a non-parametric estimate of the intensity, based on the kernel estimation, (ref. Cressie 1993, p 600) How does the function intest.s work ? Use it to estimate the intensity for the same set of data you used for the maximum likelihood estimation. Try different smoothing parameters. Plot the intensity, and compare the two estimates. (Use the same color range for the two intensities if you plot with image.)

Exercise 2: Neyman-Scott process

In this exercise we will study Neyman-Scott processes where the number of "child's" are Poisson distributed. The distribution relative to the mother, you can choose as a symmetric normal distribution (Thomas process), or the uniform distribution for a disk with radius R (Matern cluster process). (ref: Neyman-Scott processes; Stoyan & Stoyan (1994) ch.16.2)

a. Simulate a Neyman-Scott process at the unit square unconditional to the number of points. What would you do to avoid boundary effects ? Vary the parameters of the model. How does it affect the point patterns.

b.A Neyman-Scott process where the number of child's are Poisson distributed, is equivalent with a Cox process with intensity function:

$$\Lambda(r) = \omega \sum_{i=1}^{\infty} f(r - s_i) \ ,$$

where ω is expected number of child's for each mother, f is the distribution of the "children's" position relative to the mother and s_i is the location of mother number i.

Loosely one may say: A Cox- process is a "inhomogeneous Poisson distribution with stochastic intensity". Which features of $\Lambda(r)$ are stochastic ? How could You simulate $\Lambda(r)$ conditioned to the number of child's ? How would You progress in order to simulate a Neyman-Scott process conditional to the number of (child) points. You do not have to do the simulation.

c. Calculate the L-function for some of the simulated point patterns (ref exercise 3), and compare with the theoretical L- function, given under. (NB! Use the Splus commando `ppregion()` to initialize, the point process region.)

For the Thomas-process we have:

$$K(t) = \pi t^2 + \rho^{-1}(1 - \exp(-t^2/4\sigma^2))$$

Where ρ is the intensity of the mother process, and σ^2 is the variance of the "children's" distribution relative to the mother.

For the Matern cluster-process we have

$$K(t) = \pi t^2 + \rho^{-1} \times \begin{cases} 2 + \pi^{-1}[(8z^2 - 4)\arccos z - 2\arcsin z \\ +4z\sqrt{(1-z^2)^3} - 6z\sqrt{1-z^2}] & (t \leq 2R) \\ 1 & (t > 2R) \end{cases}$$

where $z = t/2R$ and other notation as earlier.

Exercise 3: Strauss process

Load the "Spatial" library using the commando `library(spatial)`, this Library contains among other things a function for calculating the L -function. $(L(t) = \sqrt{K(t)/\pi})$ with "envelopes", and a function for simulating the Strauss process, conditional to the number of points. Relative to the binomial process, is the unnormalized density of a Strauss process defined as $c^{s(x)}$ where $c > 0$ and $s(x)$ is the number of point pairs $(\mathbf{x}_i, \mathbf{x}_j)$ where $\|\mathbf{x}_i - \mathbf{x}_j\| < r$.

a. Simulate the Strauss-process with 50 points, $r = 0.1$ and $c = 0.01$, $c = 0.5$, $c = 1.0$. Compare the point patterns, make sure the plots are on the same scale. Calculate the L-function of your point patterns, Plot it together with "envelopes" based on 99 simulations from a Binomial process. What does the plots tell.(NB! use `ppregion()`) What happens if $c > 1$ if 1) we condition to the number of points, 2) we do not condition to the number of points.

b. Describe a Metropolis-Hastings algorithm to simulate from a Strauss-process conditioned to the number of points. (You do not have to do the simulation.)

c. Describe how to simulate an unconditional Strauss process by the use of a Metropolis-Hastings algorithm. (You do not have to do the simulation.)

Exercise 4: Analysis of point patterns.

In this exercise we will use three different point patterns. "cells.dat","redwood.dat" and "pines2.dat". Use the command `ppinit` from the "spatial" library to load the files.

a. Plot the point patterns. Does the the Poisson assumption apply for any of the point patterns? Use different tests to investigate. (f.ex L-function with Poisson envelopes, dispersion index, others?)

b. Implement an estimate of the D-function. D is the distribution of the distance from a point to the nearest neighboring point. Plot this for the three examples. How did you solve the boundary problems ? Useful Splus functions `dist()`, `dist2full` as described in help entry for `dist()`, and `apply( , ,min)`.

c. Choose a point pattern that is not Poisson distributed and suggest a better model. Try to estimate the parameters of the model you suggest. Check the model fit by use of the L and D function. Useful Splus commands. `nlmin` ,`integrate` ,`pplik` You do not have to use any of theese.

References:

Cressie; N.; 1993: Statistics for Spatial Data, Wiley & Sons.

Stoyan; D. and Stoyan; H.; 1994: Fractals, Random Shapes and Point Fields; Wiley & Sons. Kap 11,12,13,15,16