Methods for Spatial Point Patterns

Bailey and Gatrell

Chapter 3

Lecture 8

September 29, 2009

Point Patterns

- A **spatial point process** is any stochastic mechanism that generates a countable set of events (s_i) in a plane
- **Event** is the term used for an observation
- **Point** is the term used for an arbitrary location

mapped point pattern: all relevant events in a study area R have been recorded

sampled point pattern: events are recorded from a sample of different areas within a region

Objectives

- Determine if there is a tendency of events to exhibit a systematic pattern over an area as opposed to being randomly distributed
- Estimate how the intensity of the point pattern varies over an area (first order effect)
- Estimate the presence of spatial dependence among events (second order effect)
- Seek models to account for observed patterns

Analysis approach

- Patterns in event locations are the focus of the analysis
- Events may have attributes which can be used to distinguish types – but it is the location pattern that is analyzed
- Stochastic process of interest is where events are likely to occur
- Does a pattern exhibit clustering or regularity?
- Over what spatial scales do patterns exist?

Analysis approach

- Generally the study region R can be an arbitrary shape
- Some methods require a square or rectangular study region
- May require working with a subset and guard area to avoid edge effects
- Any selected area should be representative of the larger region from which it is selected

Analysis approach

For first order properties - use a measure of intensity

intensity = mean number of events per unit area at point **s** defined as the limit

$$\lambda(s) = \lim_{ds \to 0} \left\{ \frac{E(Y(ds))}{ds} \right\}$$

where **ds** is small region around s and ds is the area of this region

For stationary processes $\lambda(s)$ will be constant over R so $E(Y(A)) = \lambda(A)$ where Y(A) is the number of events occurring in area A

Analysis approach

Second order property, **second order intensity**, is the relationship between the number of events in pairs of areas in R.

$$\gamma(s_i, s_j) = \lim_{ds_i, ds_j \to 0} \left\{ \frac{E(Y(ds_i)Y(ds_j))}{ds_i, ds_j} \right\}$$

For stationary processes second order intensity

γ (si, sj) = γ (si - sj) = γ (**h**)

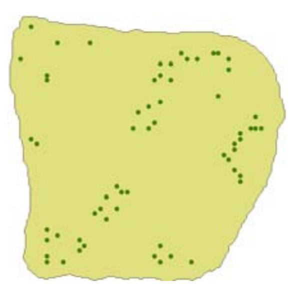
h is a vector of direction and distance

Depends only on the direction and distance between si and sj.

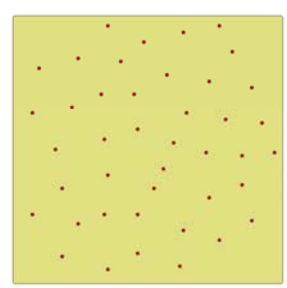
Cannot be estimated directly from the observed events

Visualizing Spatial Point Patterns

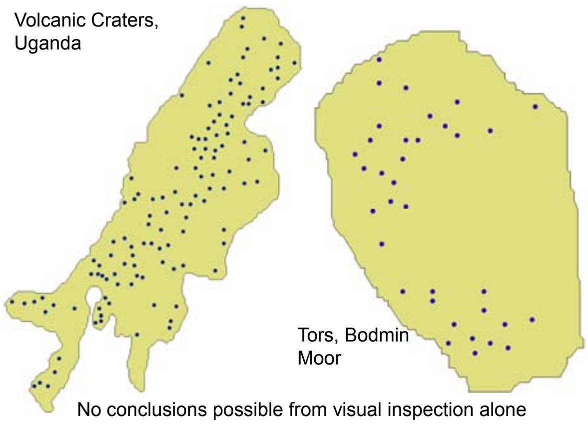
Use dot map to view pattern of events with respect to region



Redwood seedlings



Cell Centers



No conclusions possible from visual inspection alone

Visualization Issues

- Is there an underlying population distribution from which events arise in a region?
- If the underlying population varies we would expect events to cluster in areas of high population.
- Are the events more or less clustered than we would expect on the basis of population alone?

Can create event symbols inversely proportional to population density in event location and look for gaps in the maps

Exploring Point Patterns

Focus on methods for:

First order effects:

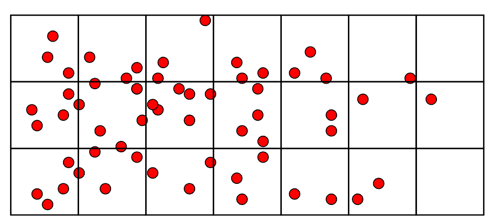
- quadrat methods
- kernel estimation

Second order effects:

- Nearest neighbor distances
- K function

Quadrat methods

Summarize number of events in each quadrat in region R.

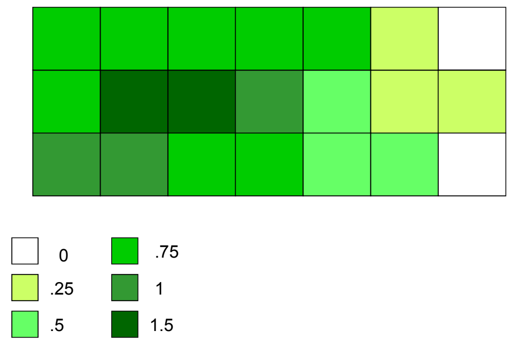


number							
	3	3	3	3	3	1	0
	4	6	6	4	2	1	1
	4	4	3	3	2	2	0

λ = n/A where n =number of events and A = 4 is area of each quadrat

Intensity λ							
	.75	.75	.75	.75	.75	.25	0
	1	1.5	1.5	1	.5	.25	.25
	1	1	.75	.75	.5	.5	0

Compare intensity variation over R to get a sense of first order effects



Quadrat methods

Disadvantages:

converts data to an area value

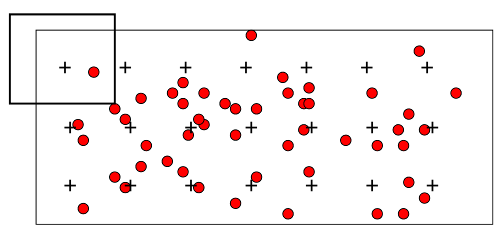
loses spatial detail

Making quadrats smaller to capture more spatial detail

generates higher variation

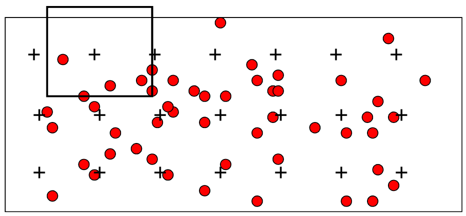
in the extreme many quadrats will have no data

Quadrat methods



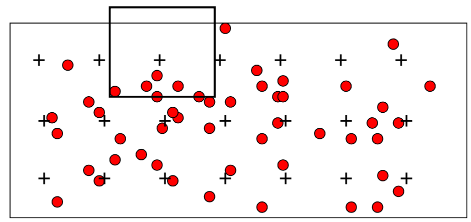
Moving window approach – a window is moved over a grid of points in R

Quadrat methods



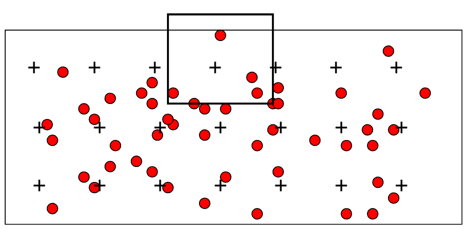
Moving window approach – a window is moved over a grid of points in R

Quadrat methods



Moving window approach – a window is moved over a grid of points in R

Quadrat methods



Moving window approach – a window is moved over a grid of points in R

What should be the size of the window?

Kernel Estimation

Adapted to provide estimates of intensity

s is general location in R

$s_1 \ldots s_n$ are the locations of n events in R

intensity $\lambda(s)$ is estimated by

$$\hat{\lambda}_\tau(s) = \frac{1}{\delta_\tau(s)} \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{(s-s_i)}{\tau}\right)$$

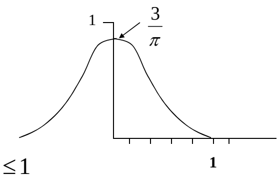
Edge correction   kernel   bandwidth

Kernel Estimation

k the kernel, a bivariate probability density function

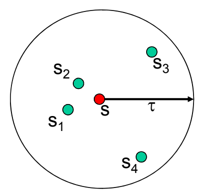
Quartic kernel

$$k(u) = \begin{cases} \frac{3}{\pi}\left(1 - u^T u\right)^2 & for\ u^T u \le 1 \\ 0 & otherwise \end{cases}$$

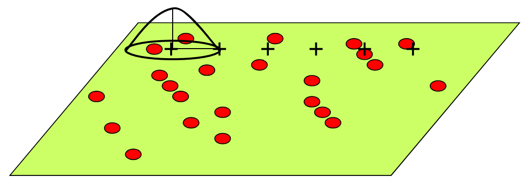


τ is the bandwidth, radius of disc centered on s

A smoothing parameter

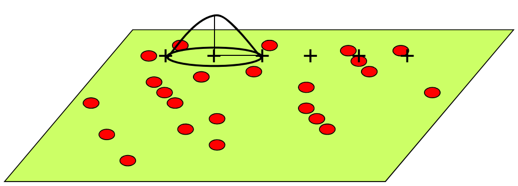


Kernel Estimation



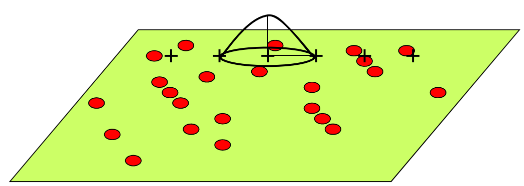
The kernel function visits each s point

Kernel Estimation



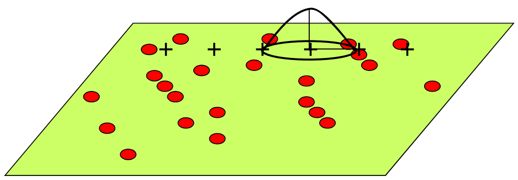
The kernel function visits each s point

Kernel Estimation



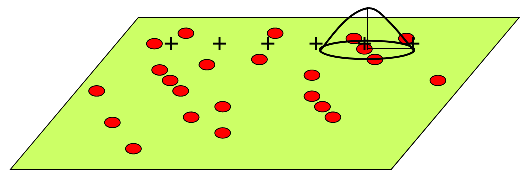
The kernel function visits each s point

Kernel Estimation



The kernel function visits each s point

Kernel Estimation



The kernel function visits each s point

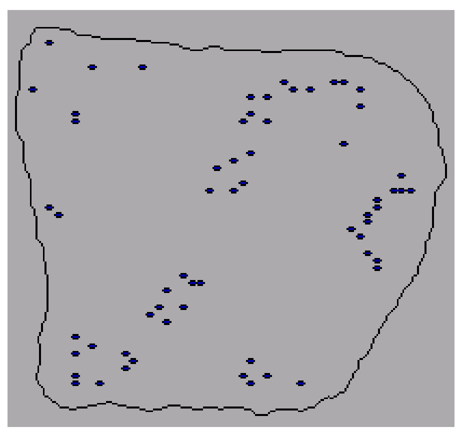
Events within the bandwidth contribute to the intensity based on weighting of kernel at that distance

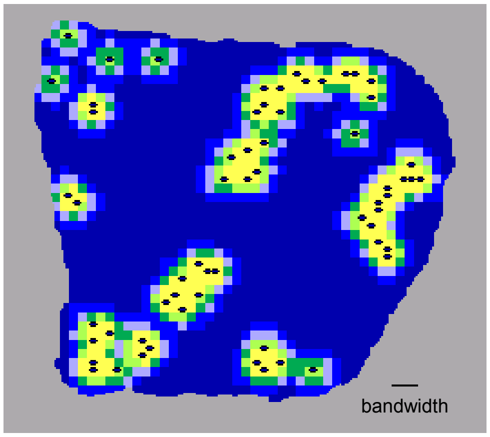
Kernel Estimation

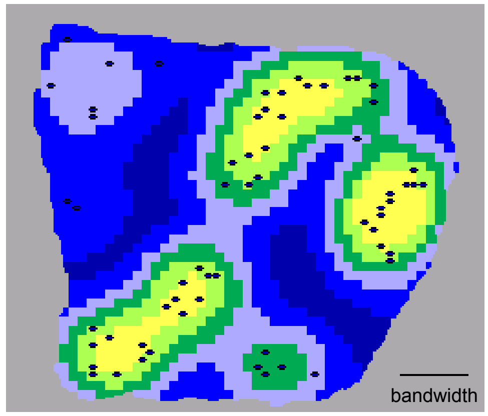
- When τ is large, the intensity will appear smooth with local details obscured
- When τ is small, the intensity appears as local spikes at the event locations

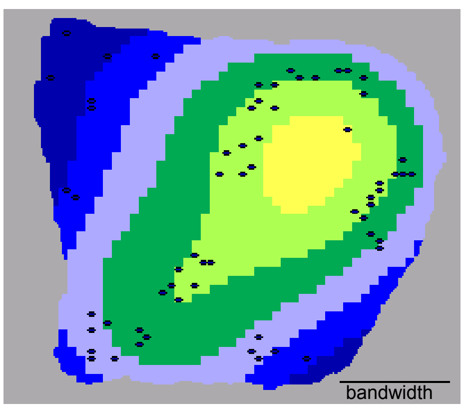
Changing the bandwidth τ allows you to look at the variation in intensity at different scales.

For exploratory purposes it is useful to test various bandwidths to examine the change in intensity at different scales





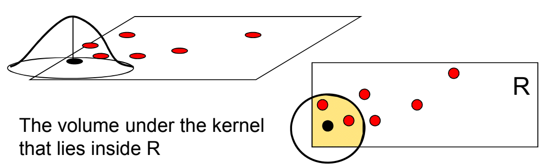




Kernel Estimation

Edge correction factor

$$\delta_\tau(s)=\int_R \frac{1}{\tau^2}k\left(\frac{(s-u)}{\tau}\right)\mathrm{d}u$$



Adaptive Kernel Estimation

For point patterns with substantial variation, an adaptive kernel estimation can be designed.

$$\hat{\lambda}_\tau(s) = \sum_{i=1}^{n} \frac{1}{\tau^2(s_i)} k\left(\frac{(s - s_i)}{\tau(s_i)}\right)$$

where $\tau(s_i)$ is a function of presence of events in the neighborhood of s_i

First get pilot estimate for τ_0

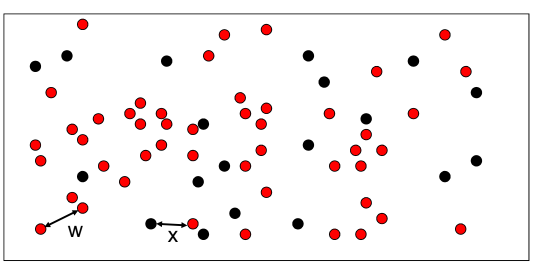
Nearest neighbor distances

Used to investigate second order properties

uses distances between events

- **W event - event nearest neighbor distances**
  distances between a randomly chosen event and its nearest neighbor event
- **X point - event nearest neighbor distances**
  distance between a randomly selected point and nearest neighbor event

Nearest neighbor distances



- event
- point

Nearest neighbor distances

Requires a completely enumerated set of events otherwise W is undefined

Spatial dependence can be explored by examining the **distribution** of nearest neighbor distances

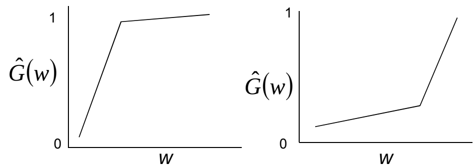
Use estimates of the empirical cumulative probability distribution of either W or X

$$\hat{G}(w) = \frac{\#(w_i \le w)}{n} \qquad \hat{F}(x) = \frac{\#(x_i \le x)}{m}$$

n is number of events in R   *m* is number of random points in R

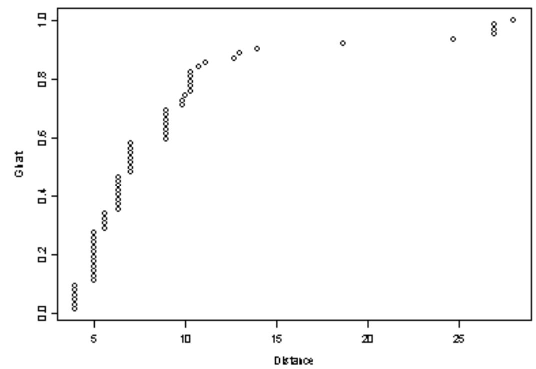
Nearest neighbor distances

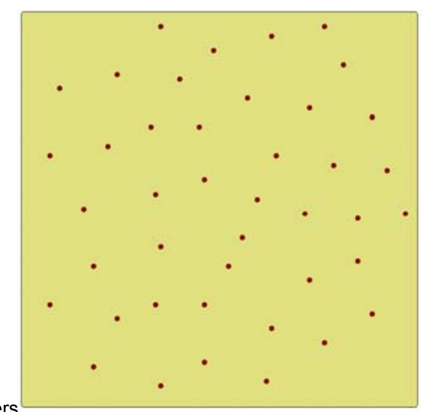
G(w) or F(x) can be plotted against values of w or x



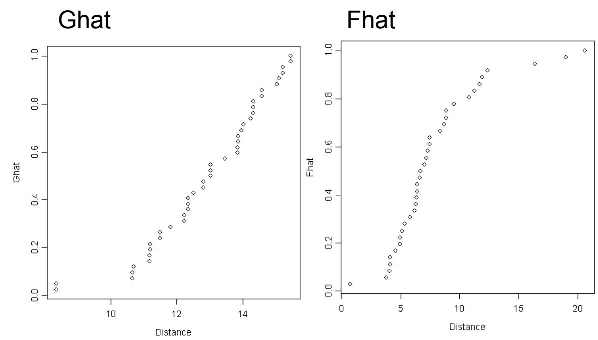
Early sharply rising function could indicate clustering - inter-event interaction

Late sharply rising function could indicate repulsion or a regular pattern

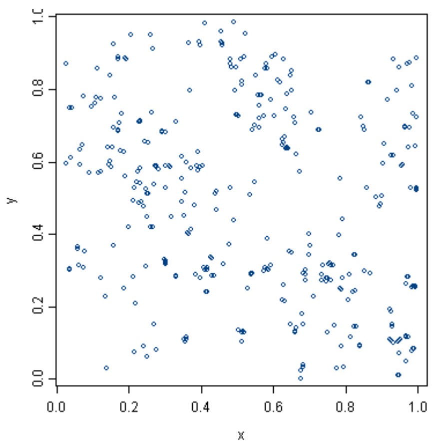


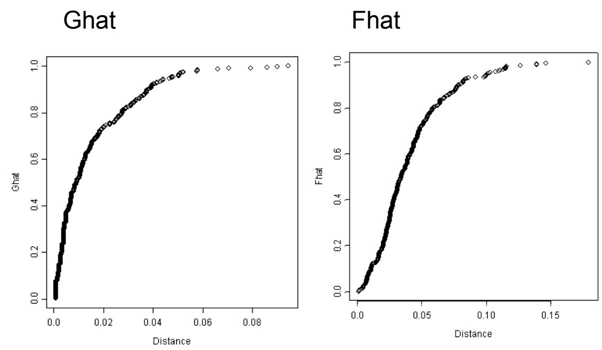


Cell Centers



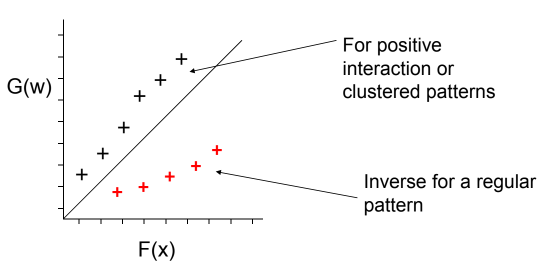
Cell Centers





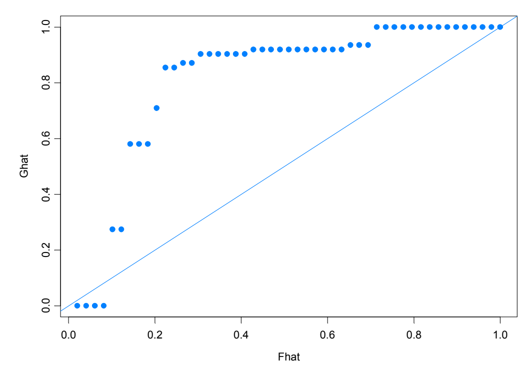
Nearest neighbor distances

G(w) plotted against F(x) should be linear if no interaction - values should be the same



x_i will tend to be large relative to w_i so G(w) would tend to be greater than F(x)

Nearest neighbor distances



Nearest neighbor distances

For boundary cases, because the nearest event may be located outside R, distance to the nearest event is unknown.

If the nearest neighbor is taken to be the closest event within the study area, expected nearest neighbor distances will be greater for events located near the boundary than for events located near the center of the study region

Thus estimates based on nearest neighbor statistics will be biased without some edge correction applied

Nearest neighbor distances

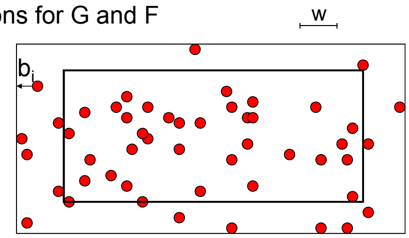
Edge corrections for G and F

Create guard area for R

or

Toroidal edge correction

or



$$\hat{G}(w) = \frac{\#(b_i > w \geq w_i)}{\#(b_i > w)}$$

The K Function

- Limitation of nearest neighbor distance method is that it uses only nearest distance - so considers only the shortest scales of variation
- Reduced second moment measure or K function is an alternative
- Provides an estimate of spatial dependence over a wider range of scales
- Describes second order properties of an isotropic process

The K Function

Need to be sure it is valid to examine second order effects over larger scales

We must assume isotropy over the region.

If not: second order properties are not constant and there is not sufficient repetition to estimate them and they may be confounded with first order effects

The K Function

Definition

$\lambda K(h) = E(\#(\text{events within distance h of an arbitrary event})$

λ is the intensity - assumed to be constant over R.

Advantage of K(h) is that it can be computed directly from event observations unlike the theoretical second order intensity.

$$\gamma(s_i, s_j) = \lim_{ds_i, ds_j \to 0} \left\{ \frac{E(Y(ds_i)Y(ds_j))}{ds_i, ds_j} \right\}$$

dsi, dsj > 0

The K Function

If R is area of *R* then **expected number of events** in *R* is λR

By definition of K, expected number of **ordered pairs** at most h distance apart is λ^2RK(h)

d_{ij} is distance between the ith and jth event in R

$I_h(d_{ij})$ is an indicator function $1 \; if \; d_{ij} \le h, \quad 0 \; otherwise$

the observed number of ordered pairs is $\sum\sum_{i \ne j} I_h(d_{ij})$

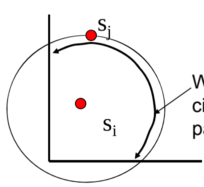
Estimate for K is

$$\hat{K}(h) = \frac{1}{\lambda^2 R} \sum\sum_{i \ne j} I_h(d_{ij})$$

The K Function

Edge correction

$$\hat{K}(h) = \frac{1}{\lambda^2 R} \sum\sum_{i \ne j} \frac{I_h(d_{ij})}{w_{ij}}$$



W_{ij} is proportion of circumference of circle centered on ith event and passing through the jth event

Conditional probability that an event is observed in R given that it is a distance d_{ij} from the ith event

The K Function

Need an estimate for λ - can use n/R

$$\hat{K}(h) = \frac{1}{\lambda^2 R} \sum\sum_{i \ne j} \frac{I_h(d_{ij})}{w_{ij}}$$

$$\hat{K}(h) = \frac{R}{n^2} \sum\sum_{i \ne j} \frac{I_h(d_{ij})}{w_{ij}}$$

The K Function

K(h) can be plotted against different values of h

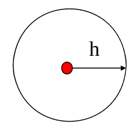
But what should K look like for no spatial dependence

Consider what K(h) should look like for a random point process

Random implies that the probability of an event at any point in R is independent of what other events have occurred and equally likely anywhere in R.

The K Function

For a random process, the expected number of events within a distance h of a randomly chosen event is $\lambda\pi h^2$



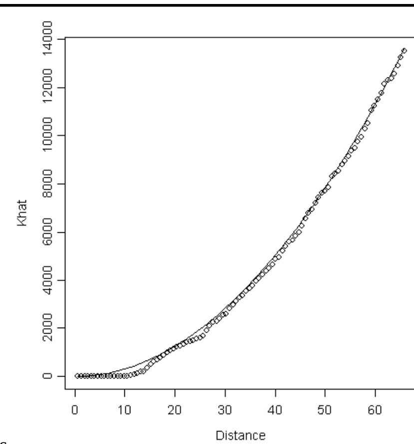
$\lambda K(h) = \lambda\pi h^2$

so $K(h) = \pi h^2$ for a random point process

The K Function

under regularity K(h) would be less than πh^2 and for clustering K(h) would be greater than πh^2

now we can compare $\hat{K}(h)$ with πh^2



Cell Centers

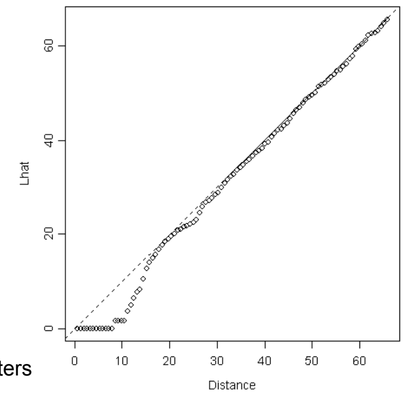
The L Function

compute the L(h) function

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}} - h$$ Bailey and Gatrell

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}}$$ S-PLUS

The L Function - SPlus



Cell Centers

The L Function – Bailey and Gatrell