Exploratory analysis for spatial point processes

References: Diggle [1983], Cressie [1993], Cressie and Wikle [2011]

• Estimating the intensity function
  – Analyzing spatial point processes in R
  – Ex 1: A nonhomogeneous Poisson point process
  – Ex 2: Trees at Hyytiälä
  – Ex 3: Gorilla nesting sites

• Estimating the K function
  – Estimating the weighted $K$-function
  – Ex 1: A homogeneous Poisson point process
  – Ex 2: Trees at Hyytiälä
  – Ex 3: Gorilla nesting sites
Estimating the intensity function

- Remember the **first order intensity function** is
  \[
  \lambda_Z(s) = \lim_{|ds| \to 0} \frac{E(Z(ds))}{|ds|}.
  \]

- We can estimate this function from a given spatial point pattern via nonparametric smoothing. Based on \( m = Z(D) \) observations \( z = \{s_1, \ldots, s_m\} \) observed in \( D \), our estimate of the intensity function is
  \[
  \hat{\lambda}_Z(s) = \frac{\sum_{i=1}^m k_b(s - s_i)}{K}.
  \]

- The **kernel** function, \( k_b(\cdot) \), that carries out the smoothing, is a probability density function which is assumed to be symmetric about the origin.

- The parameter \( b \) (the **bandwidth**) controls the amount of smoothing.
  
  - A smaller value of \( b \) yields a rougher estimate.
  
  - A larger value of \( b \) yields a smoother estimate.

- The choice of \( b \) trade-offs between too much bias and too much variability in the estimate.
Estimating the intensity function, continued

- In the estimate of the intensity function

\[ K = \int_D k_b(s - t) dt, \]

is an edge correction.

- For 2D point processes Diggle [1985] suggests the use of the product of two Epanechnikov kernels

\[ p_b(s) = e_b(s_x)e_b(s_y), \]

where \( s_x \) denotes the \( x \) coordinate of \( s \) and \( s_y \) denotes the \( y \) coordinate, with

\[ e_b(x) = \frac{3}{4b} \left[ 1 - \frac{|x|}{b} \right] I(|x|/b \leq 1). \]

- A Gaussian kernel can also be used.
Analyzing spatial point processes in R

- We use the library `spatstat` for the statistical analysis of point processes in R.
  
  (http://cran.r-project.org/web/packages/spatstat/)

- The R function `density` will estimate the intensity function of a spatial point pattern `x`.

  \[
  \text{density}(x, \text{sigma})
  \]

  Here `sigma` is the bandwidth of a Gaussian kernel.

- To use the bandwidth proposed by Diggle [1985] use

  \[
  \text{density}(x, \text{bw.diggle})
  \]
Ex 1: A nonhomogeneous Poisson point process

- Here is the simulated example from the first set of notes.
Ex 2: Trees at Hyytiälä

- Comments:
Ex 3: Gorilla nesting sites

- Comments:
Estimating the K function

- Suppose that the spatial point process is stationary and isotropic with intensity $\lambda_Z$.

- Remember that **Ripley’s K-function** is defined to be

$$K_Z(t) = \frac{E_Z(t)}{\lambda_Z},$$

where $E_Z(t) =$ expected number of events within a distance $t$ of an arbitrary point.

- Based on $z = \{s_i \in D : i = 1, \ldots, m\}$ we estimate $E_Z(t)$ by

$$\frac{1}{m} \sum_{i=1}^{m} \sum_{j \neq i} I(||s_i - s_j|| \leq t),$$

and $\lambda_Z$ by

$$(m - 1)/|D|,$$

- The **estimated K-function** is then

$$\hat{K}_Z(t) = \frac{|D|}{m(m - 1)} \sum_{i=1}^{m} \sum_{j \neq i} I(||s_i - s_j|| \leq t).$$

- This estimator can have high bias for large values of $t$. 
Estimating the weighted $K$-function

- In practice we need to introduce weighting into the estimator to correct for edge effects (allowing for the fact when cannot observe points outside $D$).

Given weights $\{w_{ij}\}$ the estimated $K$-function becomes

$$
\hat{K}_Z(t) = \frac{|D|}{m(m-1)} \sum_{i=1}^{m} \sum_{j \neq i} w_{ij} I(||s_i - s_j|| \leq t).
$$

- We use the R function `Kest` (with Ripley’s isotropic correction) to estimate the $K$-function.

- We can view the departure from a CSR process by plotting $\hat{K}_Z(t) - \pi t^2$ versus $t$.

- At short distances $t$:
  1. We have **clustering** when $K_Z(t) - \pi t^2 > 0$.
  2. We have **regularity** when $K_Z(t) - \pi t^2 < 0$. 
Ex 1: A homogeneous Poisson point process

- An example with $\lambda_Z = 500$. 

![Estimated K function and Deviation from CSR graphs](image-url)
Ex 2: Trees at Hyytiälä

Estimated K function

\[ \hat{K}_{\text{iso}}(r) \]
\[ K_{\text{pola}}(r) \]

Deviation from CSR

\[ \hat{K}_{\text{iso}}(r) - K_{\text{pola}}(r) \]
Ex 3: Gorilla nesting sites

Estimated K function

\[ \hat{K}_{iso}(r) \]
\[ K_{polo}(r) \]

Deviation from CSR

\[ \hat{K}_{iso}(r) - K_{polo}(r) \]
References


