On estimation of the intensity function of a point process

M.N.M. van Lieshout

colette@cwi.nl

CWI & Eindhoven University of Technology
The Netherlands
Let $N(A)$ be the number of points in set $A$, and $M(A) = \mathbb{E}N(A)$. Often

$$M(A) = \int_A \lambda(x) \, dx$$

for some function $\lambda(x) \geq 0$.

**Goal:** estimate $\lambda$ based on a realisation $\Phi \cap A$ in a bounded Borel set $A$ (assumed to be open and convex).
The Berman–Diggle kernel estimator

For \( x_0 \in A \), set

\[
\lambda_{BD}(x_0) := \frac{N(b(x_0, h) \cap A)}{|b(x_0, h) \cap A|} = \sum_{x \in \Phi \cap A} k_h(x_0 \mid x),
\]

with kernel

\[
k_h(x_0 \mid x) = \frac{1_{\{|x - x_0| < h\}}}{|b(x_0, h) \cap A|}.
\]

The bandwidth \( h > 0 \) controls the amount of smoothing.

**Disadvantages:**

- mass preservation need not hold;
- \( k_h \) is not necessarily a weight function.
Mass preserving border correction

For $x_0 \in A$, set

$$
\hat{\lambda}_K(x_0) := \sum_{x \in \Phi \cap A} \frac{1\{|x - x_0| < h\}}{|b(x, h) \cap A|}
$$

with kernel

$$
k_h(x_0 \mid x) = \frac{1\{|x - x_0| < h\}}{|b(x, h) \cap A|}.
$$

Then $\hat{\lambda}_K(x_0)$

- preserves total mass: $\int_A \hat{\lambda}_K(x_0) \, dx_0 = N(A)$;
- is a generalised weight function estimator: $\int_A k_h(x_0 \mid x) \, dx_0 \equiv 1$;
- bandwidth $h > 0$ determines smoothness.
Moments of the mass preserving kernel estimator

First moment

\[ \mathbb{E} \left[ \lambda_K(x_0) \right] = \int_{A \cap b(x_0, h)} \frac{\lambda(x)}{|b(x, h) \cap A|} \, dx. \]

Second moment

\[ \mathbb{E} \left[ \lambda_K(x_0)^2 \right] = \int_{(b(x_0, h) \cap A)^2} \frac{\rho^{(2)}(x, y)}{|b(x, h) \cap A| \, |b(y, h) \cap A|} \, dx \, dy \\
+ \int_{b(x_0, h) \cap A} \frac{\lambda(x)}{|b(x, h) \cap A|^2} \, dx, \]

where \( \rho^{(2)}(x, y) \, dx \, dy \) is the probability of points at \( dx \) and \( dy \).
Voronoi and Delaunay tessellations

Suppose that realisations of a point process are a.s. in general position. For any realisation \( \varphi \), define the Voronoi cell

\[
C(x_i \mid \varphi) := \{ y \in \mathbb{R}^d : ||x_i - y|| \leq ||x_j - y|| \quad \forall x_j \in \varphi \}
\]

of \( x_i \in \varphi \). The ensemble of cells forms the Voronoi tessellation.

- Voronoi cells are closed and convex, not necessarily bounded.
- Intersections between \( k = 2, \ldots, d + 1 \) different Voronoi cells are either empty or of dimension \( d - k + 1 \). In particular,

\[
\bigcap_{i=1}^{d+1} C(x_i \mid \varphi) \neq \emptyset \iff b(x_1, \ldots, x_{d+1}) \cap \varphi = \emptyset
\]

where \( b(x_1, \ldots, x_{d+1}) \) is the open ball spanned by \( x_1, \ldots, x_{d+1} \) on its boundary, and in that case is a single point, a vertex.
- The convex hull \( D(x_1, \ldots, x_{d+1}) \) is a Delaunay cell.
Contiguity and neighbourhood

$x_1, x_2 \in \varphi$ are **Voronoi neighbours** if

$$C(x_1 \mid \varphi) \cap C(x_2 \mid \varphi) \neq \emptyset,$$

i.e. their cells share a $d - 1$ dimensional border.

**Note:** the Delaunay tessellation $D(\varphi)$ is the graph with edges between Voronoi neighbours.

The union of Delaunay cells containing $x_i \in \varphi$ is the **contiguous Voronoi cell** $W(x_i \mid \varphi)$ of $x_i$ in $\varphi$. 
The Delaunay tessellation field estimator (DTFE)

Schaap and Van de Weygaert (2000, 2007)

For $x \in \Phi \cap A$, define

$$\lambda(x) := \frac{d + 1}{|W(x \mid \Phi \cap A)|}.$$ 

For any $x_0 \in A$ in the interior of some Delaunay cell, define

$$\lambda(x_0) := \frac{1}{d + 1} \sum_{x \in \Phi \cap D(x_0 \mid \Phi \cap A)} \lambda(x)$$

by averaging over the vertices of the Delaunay cell $D(x_0 \mid \Phi \cap A)$ containing $x_0$.

Note: DTFE preserves total mass and is an adaptive kernel estimator.
DTFE as adaptive kernel estimator

Set
\[ g(x_0 \mid x, \varphi) := \frac{\sum_{D_j \in D(\varphi \cap A)} 1\{x_0 \in D_j^o; x \in D_j\}}{|W(x \mid \varphi \cap A)|} , \]
for \( x_0 \in A \setminus \varphi, x \in \varphi, \) and
\[ g(x \mid x, \varphi) := \frac{d + 1}{|W(x \mid \varphi \cap A)|} \]
if \( x \in \varphi \cap A. \)

Then
\[ \widehat{\lambda}(x_0) = \sum_{x \in \Phi \cap A} g(x_0 \mid x, \Phi) \]
and \( \int_A g(x_0 \mid x, \Phi) \, dx_0 = 1. \)
Example

On estimation of the intensity function of a point process – p. 10/1
Moments of Delaunay tessellation field estimation

**First moment**

\[
\mathbb{E} \left[ \lambda(x_0) \right] = \int_A \mathbb{E}_x \left[ g(x_0 \mid x, \Phi) \right] \lambda(x) \, dx,
\]

where \( \mathbb{E}_x \) is the conditional distribution of \( \Phi \) given a point at \( x \).

**Second moment**

\[
\mathbb{E} \left[ \lambda(x_0)^2 \right] = \int_A \int_A \mathbb{E}_{x,y} \left[ g(x_0 \mid x, \Phi) g(x_0 \mid y, \Phi) \right] \rho^{(2)}(x, y) \, dx \, dy + \int_A \mathbb{E}_x \left[ g^2(x_0 \mid x, \Phi) \right] \lambda(x) \, dx
\]

where \( \mathbb{E}_{x,y}^{(2)} \) is the conditional distribution of \( \Phi \) given points at \( x \).
The Poisson process

\( \Phi = \{x_1, x_2, \ldots \} \) is a Poisson process on \( \mathbb{R}^d \) with intensity function \( \lambda(x) \geq 0, x \in \mathbb{R}^d \), if

- for every bounded Borel set \( A \), \( N(A) := \sum_{i=1}^{\infty} 1\{x_i \in A\} \) is Poisson distributed with mean
  \[
  M(A) = \int_A \lambda(x) \, dx;
  \]
- for any \( k \) disjoint bounded Borel sets \( A_1, \ldots, A_k \), the random variables \( N(A_1), \ldots, N(A_k) \) are independent.

**Properties:** For a Poisson process,

- \( P_x^l = P \);
- \( \rho^{(2)}(x, y) = \lambda(x) \lambda(y) \).
- the Berman–Diggle estimator has mean \( M(b(x_0, h) \cap A)/|b(x_0, h) \cap A| \) and variance \( M(b(x_0, h) \cap A)/|b(x_0, h) \cap A|^2 \).
The stationary Poisson process: Mean

Except for stationary Poisson processes, little is known about the distribution of $W(x|\Phi)$. However, results obtained are approximately valid for Poisson processes with slowly varying intensity function.

**Theorem**

If $\Phi$ is a stationary Poisson process on $\mathbb{R}^d$ with intensity $\lambda > 0$, the Delaunay tessellation field estimator $\hat{\lambda}(0)$ and the kernel estimators $\hat{\lambda}_{BD}(0)$ and $\hat{\lambda}_{K}(0)$ are unbiased.
The stationary Poisson process: Variance

**Theorem**

If \( \Phi \) is a stationary Poisson process on \( \mathbb{R}^d \) with intensity \( \lambda > 0 \), \( \lambda(0) \) has variance \( c_d \lambda^2 \) with \( c_d \) given by

\[
E_1 \left[ \frac{1}{|W(0 | \Phi \cup \{0\})|} \left( 1 + \sum_{y \in N(0|\Phi \cup \{0\})} \frac{|W(0 | \Phi \cup \{0\}) \cap W(y | \Phi \cup \{0\})|}{|W(y | \Phi \cup \{0\})|} \right) \right]^{-1}.
\]

**Notes:** the kernel estimators are unbiased with variance

\[ \lambda \omega_d^{-1} h^{-d}, \]

i.e. more efficient whenever \( EN(b(0, h)) > c_d^{-1} \).

For \( d = 1, c_1 = 2 (2 - \pi^2 / 6) \approx 0.7. \)
Example 1

Intensity function (solid line) with estimates of expectation of kernel (finely dashed) and DTFE (coarsely dashed) estimator.

Mean integrated absolute (squared) error 16(18) vs 9.6(22). We prefer DTFE.
Example 2

Intensity function (solid line) with estimates of expectation of kernel (finely dashed) and DTFE (coarsely dashed) estimator.

Mean integrated absolute (squared) error $1.3(3)$ vs $1.0(10)$. Average estimated standard deviation of DTFE is up to the first decimal equal to $\bar{\lambda}\sqrt{c_1} = 0.7$. We prefer kernel estimation.
We considered the problem of estimating the intensity function of a point process. We

- recalled and modified a classic kernel estimator;
- discussed the DTFE as an adaptive kernel estimator;
- gave general expressions for mean and variance in terms of the first and second order factorial moment measures of the underlying point process;
- focussed on the stationary Poisson process and showed that the DTFE and kernel estimators are asymptotically unbiased; with variance $c_d \lambda^2$ and $\lambda \omega_d^{-1} h^{-d}$ respectively.
(Schaap, 2007)