Edge corrected non-parametric intensity function estimators for heterogeneous Poisson point processes

M.N.M. van Lieshout

CWI and *EURANDOM*

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

July 2007

Abstract
We introduce a non-parametric mass preserving estimator for the intensity function of a Poisson point process. The new estimator's (integrated) mean squared error is compared to that of the classic Berman–Diggle estimator, both pointwise and cumulatively.

Keywords & Phrases: heterogeneous Poisson point process, intensity function, mass preservation.

2000 Mathematics Subject Classification: 60G55, 62M30.

1. Preliminaries and notation

Let Φ be a Poisson point process observed in a bounded open Borel subset $\emptyset \neq W \subseteq \mathbb{R}^2$ of the plane with locally finite intensity function $\lambda : \mathbb{R}^2 \to [0, \infty)$. If no specific parametric model is assumed, it is natural to apply ideas from kernel estimation theory to the problem of estimation of $\lambda(\cdot)$ as proposed by Berman and Diggle [1]:

$$\widehat{\lambda_{BD}(x_0)} := \frac{N(b(x_0, h) \cap W)}{|b(x_0, h) \cap W|}, \quad x_0 \in W. \tag{1.1}$$

Here $b(x_0, h)$ denotes the open ball around x_0 with radius $h > 0$, and, for Borel sets A, $N(A)$ denotes the number of points of Φ falling in A, $|A|$ its area. The choice of bandwidth h determines the amount of smoothing. Note that as W is open, one never divides by zero. In fact, a stronger statement can be made. The function $x \mapsto |b(x, h) \cap W|$ is continuous and attains its minimum on the closure $\bar{W}$. Since any point on the boundary ∂W always has a neighbour within distance h in W, $\inf_{x \in W} |b(x, h) \cap W| > 0$. Further details may be found e.g. in [2, 3, 5].

Although (1.1) is a natural estimator, it does not necessarily preserve the total mass in W, nor is it based on a generalised weight function [4]. The purpose of this paper is to propose an alternative estimator that does possess these two properties.

The plan of this paper is as follows. In Section 2, we introduce a non-parametric, mass preserving estimator for the intensity function of an inhomogeneous Poisson process, show it can be interpreted as a generalised weight function estimator, and compute its mean and variance. In Section 3, we give upper bounds on the difference in absolute bias and variance between the new and classic estimator, and present examples to show that neither of the two estimators is universally better than the other in terms of integrated mean squared error. In Section 4, we quantify the bias in total mass of the classic estimator. A discussion section concludes the paper.

2. Intensity function estimation

The Berman–Diggle estimator (1.1) may be written as

$$\widehat{\lambda_{BD}(x_0)} = \sum_{x \in \Phi \cap W} \frac{1\{||x - x_0|| < h\}}{|b(x_0, h) \cap W|}.$$

Hence, each point $x \in b(x_0, h) \cap W$ is assigned a weight $1/|b(x_0, h) \cap W|$. If x_0 is close to the boundary of W, the weight may be higher than h^{-2}/π to compensate the relative shortage of h-close points $x \in \Phi \cap W$. A, perhaps, more natural way to correct for such edge effects is to assign a weight $1/|b(x, h) \cap W|$ to each $x \in W$ with $||x - x_0|| < h$, as in the definition below.

Definition 1. *Let Φ be a Poisson point process observed in a bounded open Borel subset W of the plane with locally finite intensity function $\lambda : \mathbb{R}^2 \to [0, \infty)$, and define a non-parametric estimator by*

$$\widehat{\lambda(x_0)} := \sum_{x \in \Phi \cap W} \frac{1\{||x - x_0|| < h\}}{|b(x, h) \cap W|}, \quad x_0 \in W. \tag{2.1}$$

Note that (2.1) is well-defined. Moreover, if $b(x_0, 2h) \subseteq W$, the Berman–Diggle estimator and the estimator of Definition 1 coincide, as there is no need for edge correction of any kind. At the other extreme, if h is larger than the diameter of W, $|b(x, h) \cap W| = |W|$ for all $x \in W$, and again there is no difference between the estimators.

In contrast to (1.1), (2.1) preserves the total mass and is based on a weight function. To see this, note that

$$\int_W \frac{1\{||x - x_0|| < h\}}{|b(x, h) \cap W|} \, dx_0 \equiv 1$$

for all $x \in W$, that is, $\widehat{\lambda(\cdot)}$ is a generalised weight function estimator. Furthermore, for any realised point pattern $\varphi \cap W$ in W,

$$\int_W \left[\sum_{x \in \varphi \cap W} \frac{1\{||x - x_0|| < h\}}{|b(x, h) \cap W|} \right] dx_0 = \sum_{x \in \varphi \cap W} \int_W \frac{1\{||x - x_0|| < h\}}{|b(x, h) \cap W|} \, dx_0 = \#\{\varphi \cap W\} \tag{2.2}$$

is equal to the number of points in $\varphi \cap W$.

In order to assess the quality of the new estimator (see Section 3 below), we proceed to compute its mean and variance. Recall that for (1.1), since $N(b(x_0, h) \cap W)$ is Poisson distributed,

$$\mathbb{E}\left[\widehat{\lambda_{BD}(x_0)}\right] = \frac{\Lambda(b(x_0, h) \cap W)}{|b(x_0, h) \cap W|}, \tag{2.3}$$

$$\mathrm{Var}\left[\widehat{\lambda_{BD}(x_0)}\right] = \frac{\Lambda(b(x_0, h) \cap W)}{|b(x_0, h) \cap W|^2}, \tag{2.4}$$

where Λ is the first order moment measure of Φ, that is, for Borel subsets $A \subseteq W$, $\Lambda(A) = \int_A \lambda(x)\, dx < \infty$.

Theorem 1. *The estimator of Definition 1 has mean and variance*

$$\mathbb{E}\left[\widehat{\lambda(x_0)}\right] = \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x,h)\cap W|}\, dx, \tag{2.5}$$

$$\mathrm{Var}\left[\widehat{\lambda(x_0)}\right] = \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x,h)\cap W|^2}\, dx. \tag{2.6}$$

The integrals (2.5)–(2.6) are finite and should be compared to (2.3)–(2.4).

Proof: By the Campbell–Mecke theorem (see e.g. [2, 5]),

$$\mathbb{E}\left[\widehat{\lambda(x_0)}\right] = \int_W \frac{1\{||x-x_0||<h\}}{|b(x,h)\cap W|}\, \lambda(x)\, dx = \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x,h)\cap W|}\, dx.$$

Regarding the second moment, note that

$$\begin{aligned}\mathbb{E}\left[\widehat{\lambda(x_0)}^2\right] &= \mathbb{E}\left\{\sum_{x,y\in\Phi}^{\neq}\left[1_W(x)1_W(y)\frac{1\{||x-x_0||<h\}}{|b(x,h)\cap W|}\,\frac{1\{||y-x_0||<h\}}{|b(y,h)\cap W|}\right]\right\}+\\ &+\ \mathbb{E}\sum_{x\in\Phi}\left[1_W(x)\frac{1\{||x-x_0||<h\}}{|b(x,h)\cap W|^2}\right].\end{aligned}$$

Since the second order product density $\rho^2(x,y)$ exists and is equal to $\lambda(x)\lambda(y)$ for a Poisson process [5], the cross term on the right hand side is equal to

$$\int_W\int_W \frac{1\{||x-x_0||<h\}}{|b(x,h)\cap W|}\,\frac{1\{||y-x_0||<h\}}{|b(y,h)\cap W|}\,\lambda(x)\lambda(y)\, dxdy = \left(\mathbb{E}\left[\widehat{\lambda(x_0)}\right]\right)^2.$$

Another appeal to the Campbell–Mecke theorem yields

$$\mathbb{E}\sum_{x\in\Phi}\left[1_W(x)\frac{1\{||x-x_0||<h\}}{|b(x,h)\cap W|^2}\right] = \int_{b(x_0,h)\cap W}\frac{\lambda(x)}{|b(x,h)\cap W|^2}\, dx.$$

Hence,

$$\mathbb{E}\left[\widehat{\lambda(x_0)}^2\right] = \left(\mathbb{E}\left[\widehat{\lambda(x_0)}\right]\right)^2 + \int_{b(x_0,h)\cap W}\frac{\lambda(x)}{|b(x,h)\cap W|^2}\, dx,$$

from which (2.6) follows. □

3. Comparison of intensity function estimators

As noted in the previous section, for x_0 and h such that $b(x_0, 2h) \subseteq W$, the estimators defined by (1.1) and (2.1) are identical. The aim of this section is to quantify the difference between them for arbitrary $x_0 \in W$ and $h > 0$.

Theorem 2. *Let Φ be a Poisson point process observed in a bounded Borel subset W of the plane with locally finite intensity function $\lambda : \mathbb{R}^2 \to [0, \infty)$. Then, for fixed $x_0 \in W$, the absolute difference in absolute bias between (1.1) and (2.1) is bounded from above by*

$$\int_{b(x_0,h)\cap W} \left| \frac{1}{|b(x_0,h)\cap W|} - \frac{1}{|b(x,h)\cap W|} \right| \lambda(x)\, dx.$$

The absolute difference in variance is bounded by

$$\int_{b(x_0,h)\cap W} \left| \frac{1}{|b(x_0,h)\cap W|^2} - \frac{1}{|b(x,h)\cap W|^2} \right| \lambda(x)\, dx.$$

The upper bounds are finite in both cases.

Proof: Fix $x_0 \in W$. Write $b_{BD}(x_0) = \mathbb{E}\left[\widehat{\lambda_{BD}(x_0)} - \lambda(x_0)\right]$ for the bias of (1.1), $b_M(x_0)$ for that of the estimator of Definition 1. By (2.3),

$$b_{BD}(x_0) = \frac{\Lambda(b(x_0,h)\cap W)}{|b(x_0,h)\cap W|} - \lambda(x_0) = \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x_0)}{|b(x_0,h)\cap W|}\, dx.$$

The right hand side can be rewritten as

$$\int_{b(x_0,h)\cap W} \left[\frac{\lambda(x)}{|b(x,h)\cap W|} - \frac{\lambda(x_0)}{|b(x_0,h)\cap W|} + \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x)}{|b(x,h)\cap W|} \right] dx,$$

which, by Theorem 1, is equal to

$$b_M(x_0) + \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x)}{|b(x,h)\cap W|}\, dx.$$

Consequently

$$\begin{aligned} |b_{BD}(x_0)| &= \left| b_M(x_0) + \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x)}{|b(x,h)\cap W|}\, dx \right| \\ &\leq |b_M(x_0)| + \int_{b(x_0,h)\cap W} \left| \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x)}{|b(x,h)\cap W|} \right| dx. \end{aligned}$$

Similarly,

$$|b_M(x_0)| \leq |b_{BD}(x_0)| + \int_{b(x_0,h)\cap W} \left| \frac{\lambda(x)}{|b(x_0,h)\cap W|} - \frac{\lambda(x)}{|b(x,h)\cap W|} \right| dx,$$

which completes the proof of the upper bound on the difference in absolute bias.

Next turn to the variance. Write $\sigma^2_{BD}(x_0)$ for the variance of (1.1), $\sigma^2_M(x_0)$ for that of (2.1). Now, by (2.4),

$$\begin{aligned} \sigma^2_{BD}(x_0) &= \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x_0,h)\cap W|^2}\, dx \\ &= \sigma^2_M(x_0) + \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x_0,h)\cap W|^2} - \frac{\lambda(x)}{|b(x,h)\cap W|^2}\, dx \end{aligned}$$

by Theorem 1, and similar arguments as for the absolute bias complete the proof. □

A widely used criterion to assess the quality of an estimator is the *integrated mean squared error* (MISE)

$$\int_W \mathbb{E} \left[\widehat{\lambda(x_0)} - \lambda(x_0)\right]^2 dx_0 = \int_W \left[\mathrm{Var}(\widehat{\lambda(x_0)}) + \mathrm{bias}^2(\widehat{\lambda(x_0)})\right] dx_0, \tag{3.1}$$

which balances bias and variance.

Theorem 3. *Let Φ be a Poisson point process observed in a bounded Borel subset W of the plane with locally finite intensity function $\lambda : \mathbb{R}^2 \to [0,\infty)$. Then, the estimator of Definition 1 has smaller MISE than the Berman–Diggle estimator (1.1) if and only if*

$$\int_W \frac{\Lambda(b(x,h)\cap W) + \Lambda^2(b(x,h)\cap W)}{|b(x,h)\cap W|^2} dx >$$

$$> \int_W \left\{ \frac{\lambda(x)}{|b(x,h)\cap W|} + \left[\int_{b(x,h)\cap W} \frac{\lambda(x_0)}{|b(x_0,h)\cap W|} dx_0\right]^2 \right\} dx. \tag{3.2}$$

Proof: For both estimators,

$$\int_W \lambda(x_0)\mathbb{E}\left[\widehat{\lambda(x_0)}\right] dx_0 = \int_W \int_W \frac{\lambda(x)\lambda(x_0)}{|b(x,h)\cap W|} 1\{||x-x_0||<h\}\, dx dx_0$$

(by (2.3) and (2.5)). Hence a comparison in integrated mean squared error (3.1) is equivalent to a comparison in

$$\int_W \mathrm{Var}(\widehat{\lambda(x_0)}) + \left[\mathbb{E}\widehat{\lambda(x_0)}\right]^2 dx_0 = \int_W \mathbb{E}\left[\widehat{\lambda(x_0)}^2\right] dx_0.$$

By (2.3)–(2.4) and Theorem 1, the Berman–Diggle estimator has larger integrated mean squared error than (2.1) if and only if inequality (3.2) holds. □

The inequality (3.2) involves the unknown intensity function, the bandwidth h, and the geometry of the observation window W. Indeed, neither of the two estimators we are interested in is universally better than the other. To substantiate this claim, we proceed to give specific examples.

3.1 Global edge correction leads to small MISE

Let Φ be a homogeneous Poisson process with intensity $\lambda > 0$. Recall that under the assumption of homogeneity, the usual estimator $\hat{\lambda}^* = N(W)/|W|$ is unbiased with variance $\lambda/|W|$. In the absence of such information, a non-parametric estimator must be used.

The integrated variance of both (1.1) and (2.1) is equal to $\lambda \int_W |b(x,h)\cap W|^{-1} dx$. It is interesting to observe that the expression is decreasing in h. Moreover, for all $h > 0$, $|b(x,h)\cap W| \leq |W|$, so the non-parametric estimators never outperform $\hat{\lambda}^*$.

The bias of the Berman–Diggle estimator is zero, whereas by (2.5) the integrated squared bias of (2.1) is given by

$$\lambda^2 \int_W \left[\int_W \frac{1\{||x - x_0|| < h\}}{|b(x,h) \cap W|} \, dx - 1 \right]^2 dx_0 \geq 0$$

with equality if and only if $\int_{b(x_0,h)\cap W} |b(x,h) \cap W|^{-1} dx \equiv 1$ for almost all $x_0 \in W$.

In summary, the two types of edge correction are indistinguishable in integrated variance. If $\int_{b(x_0,h)\cap W} |b(x,h) \cap W|^{-1} dx \neq 1$ for x_0 in a non-null set, (1.1) has smaller integrated squared bias, but, as we shall see in Example 1 below, at the cost of a higher variance for its cumulative counterpart.

3.2 Local edge correction leads to small MISE

Let Φ be a Poisson process on W with intensity function $\lambda(x) = \lambda \, |b(x,h) \cap W|$, for some $\lambda > 0$. Then, the estimator of Definition 1 is unbiased with integrated variance $\lambda \, |W|$.

Write

$$g(x_0) := \int_{b(x_0,h)\cap W} \frac{|b(x,h) \cap W|}{|b(x_0,h) \cap W|^2} \, dx.$$

Note that $g(x_0) \leq |W|/|b(x_0,h) \cap W|$ is finite for all x_0 in the open set W, and depends only on the geometry of W and the bandwidth parameter h. Now, the mean (2.3) of the Berman–Diggle estimator evaluated at $x_0 \in W$ can be expressed as $\lambda(x_0) g(x_0)$, so its integrated squared bias is zero if and only if $g(x_0) \equiv 1$ for almost all $x_0 \in W$. Its integrated variance is $\lambda \int_W g(x_0) \, dx_0$, which reduces to $\lambda |W|$ if $g(x_0) \equiv 1$ for almost all $x_0 \in W$. Thus, under this condition, the two estimators are indistinguishable. Otherwise, the estimator of Definition 1 outperforms the Berman–Diggle estimator in integrated mean squared error. To see this, note that

$$\int_W g(x_0) \, dx_0 = |W| \mathbb{E} \left[\frac{|b(X,h) \cap W|}{|b(X_0,h) \cap W|}^2 \right] \geq |W| \mathbb{E}^2 \left[\frac{|b(X,h) \cap W|}{|b(X_0,h) \cap W|} \right] = |W|,$$

where the expectation is taken with respect to the probability distribution defined by its density $p(x,x_0) := 1\{||x - x_0|| < h\}/(|W| \, |b(x,h) \cap W|)$ on $W \times W$. In other words, the integrated variance of (1.1) is at least as large as that of (2.1). As the latter is unbiased, the same remark is true for the integrated mean squared error.

4. Estimation of the moment measure

A natural estimator for the total mass $\Lambda(W)$ placed in W based on observation of a point process Φ in W is $N(W)$. Under the Poisson assumption, the estimator is unbiased with variance $\Lambda(W)$, see for example [2, 3, 5]. From (2.2), we know that $N(W) = \int_W \widehat{\lambda(x_0)} \, dx_0$ is equal to the integrated estimated intensity function. An alternative estimator based on (1.1) may be defined as

$$\widehat{\Lambda_{BD}(W)} := \int_W \widehat{\lambda_{BD}(x_0)} \, dx_0 = \sum_{x \in \Phi \cap W} \int_W \frac{1\{||x - x_0|| < h\}}{|b(x_0,h) \cap W|} \, dx_0. \tag{4.1}$$

As an aside, upon integration over $A \subseteq W$ rather than W for some Borel set A, an estimator for $\Lambda(A)$ is obtained.

Theorem 4. *Let Φ be a Poisson point process observed in a bounded open Borel subset W of the plane with locally finite intensity function $\lambda : \mathbb{R}^2 \to [0, \infty)$. Then,*

$$\mathbb{E}\left[\widehat{\Lambda_{BD}(W)}\right] = \int_W f(x)\,\lambda(x)\,dx,$$
$$\mathrm{Var}\left[\widehat{\Lambda_{BD}(W)}\right] = \int_W f^2(x)\lambda(x)\,dx,$$

where the function $f : W \to (0, \infty)$ is defined by

$$f(x_0) := \int_{b(x_0,h)\cap W} |b(x,h) \cap W|^{-1}\,dx. \tag{4.2}$$

Note that (4.2) is a measurable function that depends on the geometry of W and the bandwidth h only. In particular, $b(x_0, 2h) \subseteq W$ implies $f(x_0) = 1$.

Proof: By the Campbell–Mecke theorem,

$$\mathbb{E}\left[\widehat{\Lambda_{BD}(W)}\right] = \int_W \left[\int_W \frac{1\{||x - x_0|| < h\}}{|b(x_0, h) \cap W|}\,dx_0\right] \lambda(x)\,dx = \int_W f(x)\,\lambda(x)\,dx.$$

To compute the variance, observe that

$$\mathbb{E}\left[\widehat{\Lambda_{BD}(W)}^2\right] = \mathbb{E}\left[\sum_{x,y\in\Phi\cap W}^{\neq} f(x)\,f(y)\right] + \mathbb{E}\left[\sum_{x\in\Phi\cap W} f(x)^2\right].$$

Since the second order product density $\rho^2(x, y)$ exists and is equal to $\lambda(x)\lambda(y)$ for a Poisson process [5], the cross term on the right hand side is equal to

$$\int_W \int_W f(x)f(y)\,\lambda(x)\lambda(y)\,dxdy = \left(\int_W f(x)\,\lambda(x)\,dx\right)^2 = \left(\mathbb{E}\left[\widehat{\Lambda_{BD}(W)}\right]\right)^2.$$

Another appeal to the Campbell–Mecke theorem yields

$$\mathbb{E}\left[\sum_{x\in\Phi\cap W} \left(\int_W \frac{1\{||x - x_0|| < h\}}{|b(x_0, h) \cap W|}\,dx_0\right)^2\right] = \int_W f^2(x)\,\lambda(x)\,dx.$$

Consequently, the variance of $\widehat{\Lambda_{BD}(W)}$ is given by $\int_W f^2(x)\,\lambda(x)\,dx$, and the proof is complete. □

Corollary 1. *Define $W^+ := \{x_0 \in W : f(x_0) > 1\}$ and $W^- := \{x_0 \in W : f(x_0) < 1\}$. Then*

1. *$\widehat{\Lambda_{BD}(W)}$ is unbiased if and only if $\int_{W^+} [f(x) - 1]\,\lambda(x)\;dx = \int_{W^-} [1 - f(x)]\,\lambda(x)\;dx$, or, equivalently, if and only if $\mathrm{Cov}(f(U), \lambda(U)) = 0$ where U is a uniformly distributed random variable on W.*

2. *If $\int_{W^+} \left[f(x)^2 - 1\right] \lambda(x)\;dx \geq \int_{W^-} \left[1 - f(x)^2\right] \lambda(x)\;dx$, the count estimator $N(W)$ is at least as good as (4.1) in terms of variance and mean squared error.*

The converse statement of 2. with respect to the mean squared error is not necessarily true in the case that f is not identically 1 almost everywhere, as can be seen by considering $\lambda(x) = c\,1\{x \in W^-\}$ for c above or below $\int_{W^-}(1-f^2(x))\,dx\ \left[\int_{W^-}(1-f(x))\,dx\right]^{-2}$.

Example 1. *Let Φ be a homogeneous Poisson process with intensity $\lambda > 0$, and f be given by (4.2). Then, as $\int_W f(x)\,dx = |W|$, Theorem 4.1 implies that $\widehat{\Lambda_{BD}(W)}$ is unbiased. Moreover,*

$$\lambda \int_W \left[f^2(x) - 1\right] dx = \lambda \int_W \left[f(x) - 1\right]^2 dx \geq 0,$$

hence, by Theorem 4.2, the variance of $\widehat{\Lambda_{BD}(W)}$ is at least as large as that of $N(W)$, with equality if and only if $\int_{b(x_0,h)\cap W} |b(x,h)\cap W|^{-1} dx = 1$ for almost all $x_0 \in W$. Recall this is exactly the condition for (2.1) to have zero integrated squared bias, cf. Section 3.1.

Example 2. *Consider the intensity function of Section 3.2. Then, by Theorem 4,*

$$\mathbb{E}\left[\widehat{\lambda_{BD}(W)}\right] = \int_W f(x)\lambda(x)\,dx = \lambda \int_W \int_W \frac{|b(x,h)\cap W|}{|b(x_0,h)\cap W|} 1\{||x - x_0|| < h\}\,dx dx_0.$$

Now, write $W \times W$ as the union of three disjoint sets on which $|b(x,h)\cap W|$ is less than, equal, or larger than $|b(x_0,h)\cap W|$, and use a symmetry argument as well as the fact that $y + 1/y > 2$ for positive $y \neq 1$ to obtain

$$\mathbb{E}\left[\widehat{\lambda_{BD}(W)}\right] \geq \lambda \int_W \int_W 1\{||x - x_0|| < h\}\,dx dx_0 = \lambda \int_W |b(x_0,h)\cap W|\,dx_0 = \Lambda(W)$$

with equality if only if $|b(x_0,h)\cap W|$ is constant for almost all $x_0 \in W$. Otherwise, $\widehat{\Lambda_{BD}(W)}$ overestimates $\Lambda(W)$. The variance $\mathrm{Var}(\widehat{\Lambda_{BD}(W)}) = \int_W f^2(x)\lambda(x)\,dx$ can be expressed as $\Lambda(W)\,\mathbb{E}_\lambda\left[f^2(X)\right]$ where the expectation is with respect to the normalised intensity function. Since

$$\mathbb{E}_\lambda\left[f^2(X)\right] \geq \left[\mathbb{E}_\lambda f(X)\right]^2 = \frac{1}{\Lambda(W)^2}\mathbb{E}^2\left[\widehat{\Lambda_{BD}(W)}\right] \geq 1,$$

the variance of $\widehat{\Lambda_{BD}(W)}$ exceeds that of the count estimator $N(W)$.

5. Discussion

As the distribution of a heterogeneous Poisson point process is completely determined by the intensity function λ, provided it exists, we have focused on this model. However, most of the work presented here can be carried over to the case of a locally finite point process whose first order moment measure exists and allows for a Radon–Nikodym derivative with respect to Lebesgue measure. Kernel estimators for this intensity function may be defined exactly as for the Poisson case, (2.2) remains valid, and the equations for the mean carry over verbatim. Regarding the variance, we must assume that a second order product density $\rho^{(2)}$ exists [5]. Doing so, the variance of (2.1) can be expressed as

$$\int_{b(x_0,h)\cap W} \int_{b(x_0,h)\cap W} \frac{\rho^{(2)}(x,y) - \lambda(x)\lambda(y)}{|b(x,h)\cap W|\,|b(y,h)\cap W|}\,dx dy + \int_{b(x_0,h)\cap W} \frac{\lambda(x)}{|b(x,h)\cap W|^2}\,dx$$

with a similar adaptation for (1.1). Generalisation to $\mathbb{R}^d$ is straightforward.

The estimators discussed in this paper involve a bandwidth parameter h; the larger h, the smoother the estimated intensity function. For specific models, h may be chosen by optimisation of the (integrated) mean squared error [3]. In practice, Diggle [3] recommends to choose h proportional to $n^{-1/2}$, where n is the observed number of points.

References

1. M. Berman and P. J. Diggle. Estimating weighted integrals of the second-order intensity of a spatial point process. *J. Roy. Statist. Soc. Ser. B*, 51:81–92, 1989.
2. N. A. C. Cressie. *Statistics for spatial data.* John Wiley and Sons, New York, 1991. Revised edition, 1993.
3. P. J. Diggle. *Statistical analysis of spatial point patterns.* Academic Press, London, 1983. Revised edition, 2003.
4. B. W. Silverman. *Density estimation for statistics and data analysis.* Chapman and Hall, London, 1986.
5. D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* John Wiley and Sons, Chichester, 1987. Revised edition, 1995.