Ann. Inst. Statist. Math.
Vol. 57, No. 1, 71–81 (2005)

OPTIMISATION OF LINEAR UNBIASED INTENSITY ESTIMATORS FOR POINT PROCESSES

TOMÁŠ MRKVIČKA[1]* AND ILYA MOLCHANOV[2]

[1] *Mathematical Department, Pedagogical Faculty, University of South Bohemia, Jeronýmova 10, 17001 České Budějovice, Czech Republic*
[2] *Department of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH- 3012 Bern, Switzerland*

(Received December 1, 2003; revised April 27, 2004)

Abstract. A general non-stationary point process whose intensity function is given up to an unknown numerical factor λ is considered. As an alternative to the conventional estimator of λ based on counting the points, we consider general linear unbiased estimators of λ given by sums of weights associated with individual points. A necessary and sufficient condition for a linear, unbiased estimator for the intensity λ to have the minimum variance is determined. It is shown that there are "nearly" no other processes than Poisson and Cox for which the unweighted estimator of λ, which counts the points only, is optimal. The properties of the optimal estimator are illustrated by simulations for the Matérn cluster and the Matérn hard-core processes.

Key words and phrases: Intensity estimation, Poisson process, linear estimators, Matérn cluster process; Matérn hard-core process.

1. Introduction

A typical task of spatial statistic is to study properties of estimators of parameters of point processes. One of the most important parameters of a point process is its intensity. In this work we consider a point process Φ with a known non-stationary structure. It means that the intensity measure of Φ has the form

$$\Lambda(B) = \int_B \lambda\gamma(x)dx,$$

where λ is unknown and $\gamma(x)$ is a known function that determines the non-stationary structure of Φ. Φ is said to be a *process with unknown scaling*.

If $\gamma(x)$ is constant, then Φ is stationary. To prevent the over-parametrisation of the model, assume that $\int_W \gamma(x)dx = |W|$, where $W \subset \mathbb{R}^d$ is the observation window and $|W|$ is the d-dimensional Lebesgue measure of W.

Processes with unknown scaling are important itself but even more because they are related to stationary processes of compact sets with a known distribution of primary grain, see Stoyan *et al.* (1995), Molchanov (1997). In fact the reference points of compact sets which hit the observation window form a process with unknown scaling where the function γ may be computed from the distribution of the primary grain.

*This research was partially supported by Grant Agency of Czech Republic, project No. 201/03/D062.

There are various estimators of the intensity λ. One often used is

$$\widehat{\lambda} = \frac{\Phi(W)}{|W|}, \tag{1.1}$$

where $\Phi(W)$ denotes the number of points of Φ in W. The estimator $\widehat{\lambda}$ is strongly consistent as $W \uparrow \mathbb{R}^d$ if the process Φ is ergodic. Other estimators based on the interpoint distance method are discussed in Byth (1982) and Diggle (1983). An estimation method in which a fraction of the points is independently marked (and thus counted) followed by consideration of the ranks of the nearest marked point is described by Särkkä (1992).

In this work we consider a family of unbiased estimators (called first order or linear, unbiased estimators)

$$\widehat{\lambda}_f = \sum_{x\in\Phi\cap W} f(x), \tag{1.2}$$

where $f : W \to \mathbb{R}$ is a weight function. Condition

$$\int_W f(x)\gamma(x)dx = 1$$

ensures that $\widehat{\lambda}_f$ is unbiased. The estimator (1.1) corresponds to $f(x) = |W|^{-1}$.

An important task is to compare the existing estimators and find an optimal estimator. For stationary Poisson and mixed Poisson processes (Stoyan *et al.* (1995)) the exact likelihood is available and the maximum likelihood estimator and the minimum variance unbiased estimator is $\widehat{\lambda}$. However for other more complicated processes the problem of determining the best estimator is still open.

In Section 2 we find a necessary and sufficient condition on f for estimator (1.2) to have the minimal variance among all linear, unbiased estimators. Then the question, when the common constant estimator $\widehat{\lambda}$ given by (1.1) is optimal, is discussed. It is shown that $\widehat{\lambda}$ is a minimum variance unbiased estimator for non-stationary Poisson and mixed Poisson point processes and that there are "nearly" no other processes than Poisson and Cox for which $\widehat{\lambda}$ is optimal. In Section 3 we present simulation studies which compare the constant estimator with the optimal one.

2. Necessary and sufficient condition for optimal estimator

Assume that the process Φ is a second-order point process in $\mathbb{R}^d$. Let $\mu^{(2)}(d(x,y))$ (respectively $\alpha^{(2)}(d(x,y))$) denote the second-order moment (respectively factorial moment) measure of Φ. Let $\mathcal{F}_\Phi$ be the family of real-valued measurable functions on W such that

$$J_\Phi(f) = \int_{W\times W} f(x)f(y)\mu^{(2)}(d(x,y)) < \infty.$$

Furthermore, let

$$\mathcal{G}_\Phi = \left\{ f \in \mathcal{F}_\Phi : \int_W f(x)\gamma(x)dx = 1 \right\}$$

and

$$\mathcal{H}_\Phi = \left\{ f \in \mathcal{F}_\Phi : \int_W f(x)\gamma(x)dx = 0 \right\}.$$

The variance of $\widehat{\lambda}_f$ is $\mathrm{Var}(\widehat{\lambda}_f) = J_\Phi(f) - \lambda^2$, where

$$\begin{aligned} J_\Phi(f) &= \int_{W\times W} f(x)f(y)\mu^{(2)}(d(x,y)) \\ &= \lambda \int_W f^2(x)\gamma(x)dx + \int_{W\times W} f(x)f(y)\alpha^{(2)}(d(x,y)). \end{aligned}$$

The aim is to minimize the variance of the estimator given by (1.2) for $f \in \mathcal{G}_\Phi$. Then minimisation of $\mathrm{Var}(\widehat{\lambda}_f)$ corresponds to minimisation of the functional $J_\Phi(f)$ over $f \in \mathcal{G}_\Phi$. Note that $J_\Phi(f)$ is a non-negative definite bilinear form hence it is convex on $\mathcal{F}_\Phi$.

LEMMA 2.1. *Let Φ be a second-order point process. The directional derivative $\delta J_\Phi(f,v)$ of $J_\Phi(f)$ at $f \in \mathcal{F}_\Phi$ in the direction $v \in \mathcal{F}_\Phi$ exists, is finite and is given by*

$$\delta J_\Phi(f,v) = 2\int_{W\times W} f(x)v(y)\mu^{(2)}(d(x,y)).$$

PROOF. The definition of the directional derivative (Zeidler (1986)) yields

$$\begin{aligned} \delta J_\Phi(f,v) &= \frac{\partial}{\partial \epsilon}\int_{W\times W} (f(x)+\epsilon v(x))(f(y)+\epsilon v(y))\mu^{(2)}(d(x,y)) \Big|_{\epsilon=0} \\ &= 2\int_{W\times W} f(x)v(y)\mu^{(2)}(d(x,y)). \end{aligned}$$

It suffices to prove that $\delta J_\Phi(f,v)$ is finite for every $f,v \in \mathcal{F}_\Phi$. The Schwartz inequality and Campbell's theorem yield

$$\begin{aligned} \int_{W\times W} f(x)v(y)\mu^{(2)}(d(x,y)) &= \mathbb{E}\left[\sum_{x_i\in\Phi}\sum_{y_j\in\Phi} f(x_i)v(y_j)\right] = \mathbb{E}\left[\sum_{x_i\in\Phi} f(x_i)\sum_{y_j\in\Phi} v(y_j)\right] \\ &\leq 2\left(\mathbb{E}\left[\sum_{x_i\in\Phi} f(x_i)\right]^2 \mathbb{E}\left[\sum_{y_j\in\Phi} v(y_j)\right]^2\right)^{1/2} \\ &= (J_\Phi(f)J_\Phi(v))^{1/2} < \infty. \end{aligned}$$

□

The following lemma provides a factorisation of the second-order factorial moment measure into the Lebesgue measure and the kernel $\mathcal{K}_x(dy)$ which is non-stationary version of the second reduced moment measure $\mathcal{K}(dh)$ defined in the stationary case in Stoyan *et al.* ((1995), p. 126).

LEMMA 2.2. *Let Φ be a point process with unknown scaling. Then*

$$\alpha^{(2)}(A\times B) = \lambda^2 \int_A\int_B \mathcal{K}_x(dy)\gamma(x)dx,$$

where $\mathcal{K}_x(dy)$ is a measure on $\mathbb{R}^d$ for every $x \in \mathbb{R}^d$.

PROOF. Let P denote the distribution of the process Φ and, let P_x denote the Palm distribution of Φ. The refined Campbell theorem (Stoyan *et al.* (1995), Section 4.4) yields

$$\begin{aligned}\alpha^{(2)}(A\times B) &= \mathbb{E}\left[\sum_{x,y\in\Phi, x\neq y} I_A(x)I_B(y)\right] = \int\sum_{x\in\Phi} I_A(x)\phi(B\setminus\{x\})P(d\phi)\\ &= \lambda\iint I_A(x)\phi(B\setminus\{x\})P_x(d\phi)\gamma(x)dx = \lambda^2\int_A\int_B \mathcal{K}_x(dy)\gamma(x)dx,\end{aligned}$$

where $I_A(x)$ is the indicator function and the measure $\mathcal{K}_x$ is defined for a Borel set B by

$$\lambda\mathcal{K}_x(B) = \int \phi(B\setminus\{x\})P_x(d\phi). \qquad \square$$

THEOREM 2.1. *Let Φ be a second-order point process with unknown scaling. Then $J_\Phi(f)$, $f\in\mathcal{G}_\Phi$, is minimal for $f=f_{\min}$ if and only if*

$$\lambda f_{\min}(x) + \lambda^2\int_W f_{\min}(y)\mathcal{K}_x(dy) = K \quad \textit{for almost all } x\in W, \tag{2.1}$$

where K is a constant such that $\int_W f_{\min}(x)\gamma(x)dx = 1$.

PROOF. First, note that $J_\Phi(f)$ attains its minimum for $f = f_{\min}$ if and only if $\delta J_\Phi(f_{\min}, v) = 0$ for every $v\in\mathcal{H}_\Phi$.

Necessity. Choose $v\in\mathcal{H}_\Phi$ given by $v(x) = \gamma(x)^{-1}I_{W_+}(x) - \gamma(x)^{-1}I_{W_-}(x)$ for a subset $W_+\subset W$ with $|W_+| = |W|/2$ and $W_- = W\setminus W_+$. Then by Lemma 2.1 and definition of $\mu^{(2)}(d(x,y))$ we get the second equality and by Lemma 2.2 we get the third equality,

$$\begin{aligned}0 = \delta J_\Phi(f_{\min}, v) &= 2\left[\lambda\int_W f_{\min}(x)v(x)\gamma(x)dx + \int_{W\times W} f_{\min}(y)v(x)\alpha^{(2)}(d(x,y))\right]\\
&= 2\left[\int_{W_+}\lambda f_{\min}(x)dx - \int_{W_-}\lambda f_{\min}(x)dx\right]\\
&\quad + 2\left[\int_{W_+}\int_W \lambda^2 f_{\min}(y)\mathcal{K}_x(dy)dx - \int_{W_-}\int_W\lambda^2 f_{\min}(y)\mathcal{K}_x(dy)dx\right]\\
&= 2\left[\int_{W_+}\left(\lambda f_{\min}(x) + \int_W\lambda^2 f_{\min}(y)\mathcal{K}_x(dy)\right)dx\right]\\
&\quad - 2\left[\int_{W_-}\left(\lambda f_{\min}(x) + \int_W\lambda^2 f_{\min}(y)\mathcal{K}_x(dy)\right)dx\right]\\
&= 2\left[\int_{W_+} g(x)dx - \int_{W_-} g(x)dx\right],\end{aligned}$$

where $g(x) = \lambda f_{\min}(x) + \int_W \lambda^2 f_{\min}(y)\mathcal{K}_x(dy)$. It is easy to see that $g(x) = K$ for almost all $x\in W$.

Sufficiency. Assuming (2.1) and by Lemma 2.1, Lemma 2.2 and definition of $\mu^{(2)}(d(x,y))$ we have

$$\begin{aligned}\delta J_\Phi(f_{\min}, v) &= \lambda \int_W f_{\min}(x) v(x)\gamma(x)dx + \lambda^2 \int_W \int_W f_{\min}(y)v(x)\mathcal{K}_x(dy)\gamma(x)dx \\ &= \int_W v(x)\gamma(x) \left[\lambda f_{\min}(x) + \lambda^2 \int_W f_{\min}(y)\mathcal{K}_x(dy)\right] dx \\ &= K \int_W v(x)\gamma(x)dx = 0,\end{aligned}$$

for every $v \in \mathcal{H}_\Phi$. $\square$

It can be shown that the optimal function $f_{\min}$ is continuous under a uniform continuity condition on the product density.

Now determine when the equally weighted estimator

$$\widehat{\lambda} = \sum_{x\in\Phi\cap W} \frac{1}{|W|} = \frac{\Phi(W)}{|W|}$$

is optimal.

THEOREM 2.2. *Let Φ be a non-stationary Poisson or mixed Poisson point process with an arbitrary unknown intensity function. Then the constant estimator is the minimum variance unbiased estimator.*

LEMMA 2.3. *Let Φ be a Poisson point process with unknown scaling. Then the statistic $\Phi(W)$ is complete and sufficient for the intensity λ.*

PROOF. $\Phi(W)$ has a Poisson distribution with parameter $\lambda|W|$. The Poisson distribution is a complete family of distributions, hence $\Phi(W)$ is complete statistic.

Since $\Lambda(W) < \infty$, Φ is a finite point process in a bounded region. Thus we can consider Φ to be a random element taking values in $\cup_{j=0}^{\infty} W^j$. Define μ as a measure on $\cup_{j=0}^{\infty} W^j$ by $\mu = \sum_{j=0}^{\infty} (\lambda^d)^j$, where $W^0 = \{\emptyset\}$, λ^d is the d-dimensional Lebesgue measure and $(\lambda^d)^0 = \delta_\emptyset$. Then the density of Φ with respect to μ is

$$f(x_1, \dots, x_j) = \exp(-\lambda|W|)\frac{(\lambda|W|)^j}{j!}\frac{\gamma(x_1)\cdots\gamma(x_j)}{|W|^j}.$$

It can be seen from the process density and from factorisation criterion (Lehmann (1991)) that the statistic $\Phi(W)$ is sufficient for λ. $\square$

PROOF OF THEOREM 2.2. Because the equally weighted estimator does not depend on the non-stationary structure $\gamma(x)$, we can assume for a while that $\gamma(x)$ is known and that we work with a point process with unknown scaling. Let Φ be a Poisson process. The Rao-Blackwell theorem, Lemma 2.3 and the fact that $\widehat{\lambda}$ is an unbiased estimator yield that

$$\widehat{\lambda}(\Phi) = \mathbb{E}[\widehat{\lambda}(\Phi) \mid \Phi(W)]$$

is the minimum variance unbiased estimator for λ for Poisson process. For mixed Poisson process (Stoyan *et al.* (1995)) it is straightforward consequence of the previous. □

The equally weighted estimator for estimating the global intensity of cyclic Poisson process is studied in Helmers and Mangku (1999). Theorem 2.2 says that in this framework the equally weighted estimator is the minimum variance unbiased estimator.

The following theorem is a characterisation of the optimality of the constant estimator among linear unbiased estimators.

THEOREM 2.3. *Let Φ be a stationary second-order point process with the second reduced moment measure $\mathcal{K}$.*

i) *If $\mathcal{K}$ is proportional to the Lebesgue measure, then the constant estimator is optimal for every observation window $W \subseteq \mathbb{R}^d$.*

ii) *If the equally weighted estimator is optimal on the observation window $W = [0, A]^d$ for every $A > 0$, then $\mathcal{K}$ is proportional to the Lebesgue measure.*

PROOF. By Theorem 2.1, the necessary and sufficient condition for the constant estimator to be optimal among all linear, unbiased estimators is

$$\mathcal{K}(\{h : x + h \in W\}) = K \quad \text{for almost all} \quad x \in W, \tag{2.2}$$

where $\mathcal{K} = \mathcal{K}_x$ is the second reduced moment measure. The implication i) follows immediately from (2.2).

ii) The stationarity implies that (2.2) is satisfied for every d-dimensional cube in $\mathbb{R}^d$. Note that the translates of cubes generate the Borel σ-algebra on $\mathbb{R}^d$. Thus (2.2) is satisfied for every Borel set. Therefore $\mathcal{K}$ is translation invariant, hence is proportional to the Lebesgue measure. □

3. Simulation study

In this section we will work only with processes for which the product density $\rho^{(2)}(x, y)$ of the second-order factorial moment measure exists. Note however that the optimal function satisfying equation (2.1) can also be found if the second-order factorial moment measure is discrete. The following corollary reformulates Theorem 2.1 for point process with existing product density.

COROLLARY 3.1. *Let Φ be a point process with unknown scaling and product density $\rho^{(2)}(x, y)$. Then $\widehat{\lambda}_f$ given by (1.2) has the smallest variance if and only if $f = f_{\min}$ with*

$$\lambda f_{\min}(x) + \int_W f_{\min}(y) \frac{\rho^{(2)}(x, y)}{\gamma(x)} dy = K \quad \textit{for almost all} \quad x \in W. \tag{3.1}$$

The equation (3.1) is a Fredholm integral equation of the second type. The theory of these integral equations is well developed, see for example Kanwal (1971), Kolmogorov and Fomin (1970) and Kress (1989).

If Φ is stationary, then $\rho^{(2)}(x, y) = \rho^{(2)}(x - y)$ depends only on the difference $(x - y)$ and $\gamma(x)$ identically equals 1. The following corollary concerns the stationary case with existing product density.

Corollary 3.2. *Let Φ be a stationary point process with product density $\rho^{(2)}(h)$. Then $\widehat{\lambda}_f$ has the smallest variance over $f \in \mathcal{G}_\Phi$ if and only if $f = f_{\min}$, where*

$$\lambda f_{\min}(x) + \int_W f_{\min}(y)\rho^{(2)}(x-y)dy = K \qquad \text{for almost all} \quad x \in W. \tag{3.2}$$

The numerical study has been carried over in two frameworks.

1. Assume that the distribution of the process is known up to the intensity λ. For each simulation λ is estimated by the equally weighted estimator $\widehat{\lambda}$. Then the optimal function $f_{\min}$ is calculated by (3.2) and the optimal estimator $\widehat{\lambda}_{f_{\min}}$ is constructed.

2. Assume that the distribution of the process belongs to a certain parametric family. For each simulation, the process parameters are estimated in order to determine the product density. Then the optimal function $f_{\min}$ is calculated by (3.2) and the optimal estimator $\widehat{\lambda}_{f_{\min}}$ is constructed.

The solution of the Fredholm integral equation of the second type (3.2) was found by a standard numerical algorithm, namely the quadrature method (Kress (1989)).

Unless otherwise stated, simulations have been carried over in a square window of side 1 in dimension 2. For every model, 1000 simulations have been performed in Framework 1 and 400 in Framework 2. There is a possibility of the third framework. That involves estimating the product density directly by a kernel estimator. But simulations showed that the estimator of product density is too unstable to achieve any improvement by our method.

Simulations reported below have been carried over for the Matérn cluster process and Matérn hard core process described in (Stoyan *et al.* (1995)).

Example 1. (Matérn cluster process) It can be shown that the optimal function $f_{\min}(x)$ for any Neyman-Scott process does not depend on the intensity λ which, in fact, we want to estimate. Figure 1 shows the optimal functions for Matérn cluster process computed in dimensions 1 and 2 for an arbitrary intensity, the cluster with ball shape with radius $R = 0.2$ and number of points in a cluster with Poisson distribution with parameter 10. The optimal function here gives to the points near the boundary bigger weight than to the middle points. We can observe a small wave in the middle of the function which we can not clearly explain. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ in

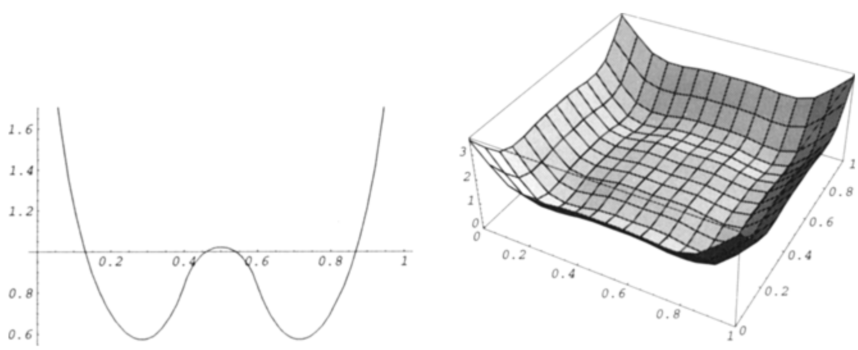

Fig. 1. The optimal functions for Matérn cluster process in one and two dimensions.

Table 1. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ for stationary Matérn cluster process.

parameters	Fr.	$\widehat{\lambda}$ mean	$\widehat{\lambda}$ Var	$\widehat{\lambda}_{f_{\min}}$ mean	$\widehat{\lambda}_{f_{\min}}$ Var	r
$\lambda = 100$, $R = 0.1$, $\mu = 10$	1	99.16	1018.78	99.4856	963.577	5.4
$\lambda = 200$, $R = 0.2$, $\mu = 20$	1	197.243	3308.25	197.542	2871.2	13.2
$\lambda = 100$, $R = 0.05$, $\mu = 10$	1	100.273	1063.76	100.101	1028.26	3.3
$\lambda = 100$, $R = 0.1$, $\mu = 10$	2	103.783	823.083	101.399	737.614	10.4

Frameworks 1 and 2 is shown in Table 1. The ratio r which appears in the tables is

$$r = 100(1 - \mathrm{Var}(\widehat{\lambda}_{f_{\min}})/\,\mathrm{Var}(\widehat{\lambda})).$$

In Framework 2 we estimate the diameter of the cluster R by half of the length where the estimated product density $\widehat{\rho(h)}$ changes for the first time, from the decreasing to the increasing behaviour. We estimate the mean number of points per cluster μ by minimising (over μ) the approximated integral

$$\int_{2b}^{2R} [\rho^{(2)}(h, \mu, \widehat{R}, \widehat{\lambda}) - \widehat{\rho(h)}]^2 dh,$$

where $\rho^{(2)}(h, \mu, \widehat{R}, \widehat{\lambda})$ is the theoretical product density, $\widehat{R}$ and $\widehat{\lambda}$ are estimated process parameters and b is the bandwidth used to calculate the kernel estimator $\widehat{\rho(h)}$ of the product density. We used the rectangular kernel because it gives smaller variance then Epanechnikov kernel, see Stoyan and Stoyan (2000).

Because the method mainly corrects the edge effects a simulation study has been carried over also for a non-square window being the map of middle Bohemia without Prague as this particular area was studied to determine the risk of being infected by encefalitida. Figure 2 shows the non-square window which was used for simulations and an optimal function for Matérn cluster process computed in dimension 2 for an arbitrary intensity, the cluster with ball shape with radius $R = 0.2$ and number of points in a cluster with Poisson distribution with parameter 20. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ based on 1000 simulations in Framework 1 shown in Table 2 confirms the achieved improvement.

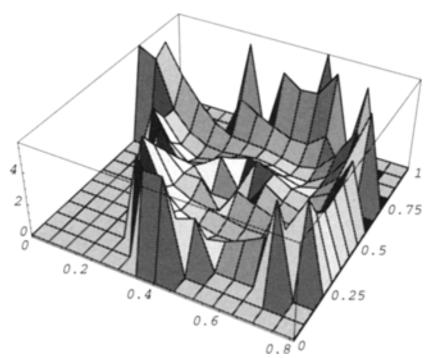

Fig. 2. The observation window and the corresponding optimal function for Matérn cluster process.

Table 2. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ for stationary Matérn cluster process in the non-square window shown on Fig. 2.

		$\widehat{\lambda}$		$\widehat{\lambda}_{f_{\min}}$		
parameters	Fr.	mean	Var	mean	Var	r
$\lambda = 100,\ R = 0.1,\ \mu = 10$	1	100.075	1902.25	99.7353	1694.95	11
$\lambda = 200,\ R = 0.2,\ \mu = 20$	1	201.293	5765.63	201.609	4897.58	15
$\lambda = 100,\ R = 0.05,\ \mu = 10$	1	97.9502	1953.47	98.3655	1828.97	6.3

Example 2. (Matérn hard-core process) Unfortunately the form of the product density here does not eliminate λ in solving the necessary and sufficient condition. But the dependence of the solution on λ is weak, it has even no influence on the shape of the solution, therefore we present here the optimal function for one parameter λ only. Figure 3 shows the optimal functions for Matérn hard-core process computed in dimension 1 and 2 with intensity 6 in dimension 1 and 50 in dimension 2 and with hard-core distance 0.075. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ in Frameworks 1 and 2 is shown in Table 3.

For hard-core processes the points near the border are getting less weight than in the middle and conversely for cluster processes. This is caused by behaviour of the optimal estimator which aims to incorporate the border effect. When a point from a repulsive point process appears near to the window's border, then we can expect no points near it and outside the observation window. Therefore the estimator gives to this point less

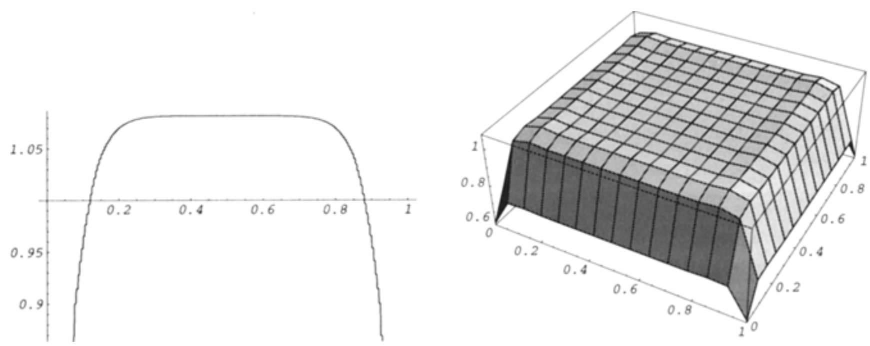

Fig. 3. The optimal functions for Matérn hard-core process in one and two dimensions.

Table 3. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ for stationary Matérn hard-core process.

		$\widehat{\lambda}$		$\widehat{\lambda}_{f_{\min}}$		
parameters	Fr.	mean	Var	mean	Var	r
$\lambda = 50,\ h = 0.075$	1	49.879	14.9773	49.891	14.5921	2.6
$\lambda = 7,\ h = 0.2$	1	6.947	2.88307	6.94688	2.71923	5.7
$\lambda = 7,\ h = 0.075$	1	7.02	6.30791	7.0288	6.30451	0.05
$\lambda = 50,\ h = 0.075$	2	50.0775	16.2822	50.088	15.7745	3.1

weight. In case of a cluster process, a point located near to the border is likely to have a number of neighbors from outside the window. Then the estimator attaches to this point more weight. Tables 1 and 3 show that the amount of the improvement of the optimal estimator depends on the size of the part of the window possibly influenced by outside points. In fact, the described method is an edge correction method, which incorporates the second order behavior of the process, in stationary case. Note that when we estimate the product density parametrically then the results are even better than for known product density. This is probably caused by the fact that when the parameters are estimated from particular observation, it suits better for particular observation then the real parameters, thus the method can better estimate the behaviour of this observation outside the window.

Example 3. (Non-stationary Matérn cluster process) Consider a non-stationary Matérn cluster process with the intensity function $\gamma(x_1, x_2) = \pi \sin(\pi x_1/2)/2$ in dimension 2. Similarly, as in the stationary case, the non-parametric kernel estimator of the intensity function is too unstable to achieve any improvement by our method. Therefore γ has to be estimated parametrically or supposed to be known. In this example we supposed that γ is known. We used the estimate of the product density of the second-order intensity-reweighted stationary point process which was introduced in Baddeley *et al.* (2000) to estimate the parameters of the cluster model in the Framework 2.

Figure 4 shows the intensity function and the optimal function for Matérn cluster process computed in dimension 2 for an arbitrary intensity, the cluster with ball shape

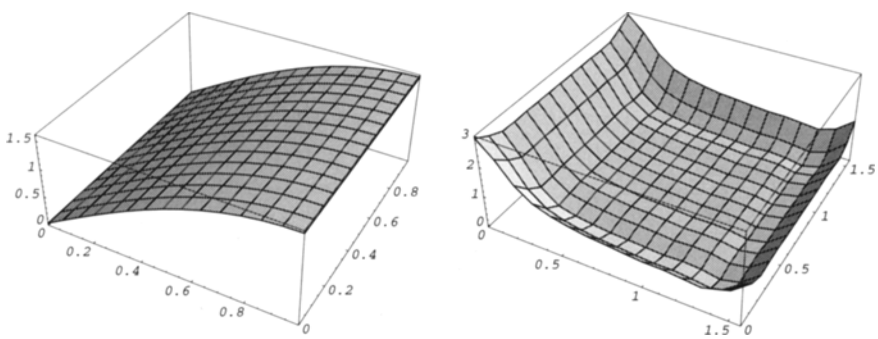

Fig. 4. The intensity function and the optimal function for the non-stationary Matérn cluster process.

Table 4. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ for Matérn cluster process with intensity function proportional to $\gamma(x_1, x_2) = \pi \sin(\pi x_1/2)/2$.

		$\widehat{\lambda}$		$\widehat{\lambda}_{f_{\min}}$		
parameters	Fr.	mean	Var	mean	Var	r
$\lambda = 100$, $R = 0.1$, $\mu = 10$	1	99.234	812.169	99.6242	712.918	12.2
$\lambda = 200$, $R = 0.2$, $\mu = 20$	1	200.738	601.072	200.835	525.252	13
$\lambda = 100$, $R = 0.05$, $\mu = 10$	1	99.903	873.649	100.393	772.606	11.6
$\lambda = 100$, $R = 0.1$, $\mu = 10$	2	102.107	836.422	101.36	726.126	13.2

with radius $R = 0.2$ and number of points in a cluster being Poisson distributed with mean 20. The comparison of $\widehat{\lambda}$ and $\widehat{\lambda}_{f_{\min}}$ in Frameworks 1 and 2 shown in Table 4 confirms improvement for the non-stationary case against the stationary case.

References

Baddeley, A. J., Møller, J. and Waagepetersen, J. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, *Statistica Neerlandica*, **54**, 329–350.

Byth, K. (1982). On robust distance-based intensity estimators, *Biometrics*, **38**, 127–135.

Diggle, P. J. (1983). *Statistical Analasis of Spatial Point Patterns*, Academic Press, London.

Helmers, R. and Mangku, I. W. (1999). Statistical estimation of Poisson intensity function, *Proceedings of Conference on Mathematics and Its Application*, Yogiokarta, Indonesia.

Kanwal, R. P. (1971). *Linear Integral Equations Theory and Technique*, Academic Press, New York and London.

Kolmogorov, A. N. and Fomin, S. V. (1970). *Introductory Real Analysis, Revised ed.*, Englewood Cliffs, New York.

Kress, R. (1989). *Linear Integral Equations*, Springer-Verlag, Berlin.

Lehmann, E. L. (1991). *Theory of Point Estimation*, Wadsworth & Brooks, California.

Molchanov, I. S. (1997). *Statistics of the Boolean Model for Practitioners and Mathematicians*, John Wiley & Sons, Chichester.

Särkkä, A. (1992). A note on robust intensity estimation for point processes, *Biometrical Journal*, **34**, 757–767.

Stoyan, D. and Stoyan, H. (2000). Improving ratio estimators of second order point process characteristics, *Scandinavian Journal of Statistics*, **27**, 641–656.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*, 2nd ed., John Wiley & Sons, Chichester.

Zeidler, E. (1986). *Nonlinear Functional Analysis and Its Applications I*, Springer-Verlag, New York.