Ann. Inst. Statist. Math.
Vol. 45, No. 2, 211–221 (1993)

ESTIMATION VARIANCES FOR ESTIMATORS OF PRODUCT DENSITIES AND PAIR CORRELATION FUNCTIONS OF PLANAR POINT PROCESSES

D. STOYAN[1], U. BERTRAM[2] AND H. WENDROCK[2]

[1]*Fachbereich Mathematik, Bergakademie Freiberg, 09596 Freiberg, Germany*
[2]*Institut für Strukturforschung und Festkörperphysik Dresden, 01171 Dresden, Germany*

(Received November 22, 1991; revised July 13, 1992)

Abstract. Approximations of the estimation variances of kernel estimators of the pair correlation function and the product density of a planar Poisson process are given. Furthermore, a heuristic approximation of the estimation variance of an estimator of the pair correlation function of a "general" planar point process is suggested. All formulae have been tested by simulation experiments.

Key words and phrases: Approximation, estimation variance, pair correlation function, Poisson process, product density, simulation.

1. Introduction

Second order characteristics play an important role in point process statistics. Usually, Ripley's K-function and the L-function (see Ripley (1981) and Stoyan *et al.* (1987)) are used for goodness-of-fit tests and parameter estimation, while the product density $\varrho(t)$ and the pair correlation function $g(t)$ are used in exploratory data analysis. The form of these functions helps to understand the type of interaction in the point pattern and to find suitable point process models. In particular, minima and maxima (if existing) of the pair correlation function may give valuable information on the strength of order.

Since estimated second order characteristics deviate from their theoretical counterparts because of statistical fluctuations, error bounds for these functions are very important. For example, they are needed to distinguish between statistical fluctuations in an estimated pair correlation function and peaks which are due to real properties of the point process under study. Until now, variances of estimation for second order characteristics are known only in particular cases. Ripley (1988) has given such variances for a series of estimators of the K-function for the Poisson process.

This paper gives estimation variances for product densities and pair correlation functions. First, in analogy with to Ripley's (1988) calculations, estimation variances in the Poisson process case are derived. The formulae obtained are quite accurate as shown by simulations. Second, a heuristic approximation of the estimation variance for the pair correlation function of a "general" planar point process is suggested, using a Poisson approximation. Again, simulation experiments have shown that the formula gives acceptable values, which can be used in practice.

2. Estimators of the product density and the pair correlation function

Let N be a planar stationary and isotropic second order point process of intensity λ. Its pair correlation function $g(t)$ can be explained in two ways; see also Penttinen and Stoyan (1989).

First, it is connected with Ripley's K-function $K(t)$ by

$$g(t) = \frac{dK(t)}{dt}/(2\pi t), \qquad t \geq 0.$$

Second, it is given by the following probability $P(t)$. Consider two infinitesimally small disks of areas dF_1 and dF_2 with intercenter distance t. Let $P(t)$ be the probability of finding a point in each of the disks. Then, up to quantities of higher order, it can be written in the form

$$P(t) = \lambda dF_1 \lambda dF_2 g(t).$$

The (second order) product density $\varrho(t)$ is given by

$$\varrho(t) = \lambda^2 g(t), \qquad t \geq 0.$$

Usually, first $\varrho(t)$ is estimated, and estimates of $g(t)$ are obtained dividing it by estimates of λ^2. Estimators of λ^2 are

$$\hat{\lambda}^2 = \left(\frac{n}{a}\right)^2 \tag{2.1}$$

and

$$\widehat{\lambda^2} = \frac{n(n-1)}{a^2}. \tag{2.2}$$

Here a is the area of the window E of observation; $n = N(E)$ is the number of points in E. The estimator (2.2) is unbiased in the case of a Poisson process.

For estimating $\varrho(t)$ edge corrected kernel estimators are frequently used, which have a property similar to unbiasedness, see Fiksel (1988). Three estimators of this type are

$$\hat{\varrho}_F(t) = \frac{1}{nt} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{k(t - \|x_i - x_j\|)}{\nu(E_{x_i} \cap E_{x_j})}, \tag{2.3}$$

$$\hat{\varrho}_R(t) = \frac{1}{nta} \sum_{i=1}^{n} \sum_{j=i+1}^{n} k(t - \|x_i - x_j\|) b_{ij}, \tag{2.4}$$

$$\hat{\varrho}_O(t) = \frac{1}{nt\gamma_E(t)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} k(t - \|x_i - x_j\|). \tag{2.5}$$

The summation goes over all points x_i (x_j) in the window E.

Here $k(s)$ is a kernel function, for example the Epanechnikov kernel, which is used throughout this paper,

$$k(s) = \begin{cases} \frac{3}{4}\varepsilon\left(1 - \frac{s^2}{\varepsilon^2}\right), & |s| \le \varepsilon \\ 0, & \text{otherwise.} \end{cases}$$

The parameter ε is called the band width.

Furthermore, $E_z = E + z = \{x : x = y + z, y \in E\}$, $\nu(X)$ is the area of X, and b_{ij} is the proportion in E of the perimeter length of the circle of radius $\|x_i - x_j\|$ centred at x_i. Finally, $\gamma_E(t)$ is the isotropized set covariance function of E, i.e. the mean of $\nu(E \cap E_Z)$, where Z is uniformly distributed on the circle of radius t centred at the origin o. For convex E, a general approximation formula for $\gamma_E(t)$ for small t is

$$\gamma_E(t) \sim a - \frac{u}{\pi}t, \tag{2.6}$$

where u is the perimeter of E.

For circular E (radius R) it is

$$\gamma_E(t) = 2R^2 \arccos\frac{t}{2R} - \frac{t}{2}\sqrt{4R^2 - t^2}, \qquad 0 \le t \le 2R,$$

which can be approximated by

$$\gamma_E(t) \sim \pi R^2 - 2Rt - \frac{t^3}{6R}.$$

For rectangular E (side lengths α and β, $\alpha \le \beta$) and $t \le \alpha$ it is

$$\gamma_E(t) = \alpha\beta - \frac{2(\alpha + \beta)}{\pi}t + \frac{t^2}{\pi}.$$

Practical experience and simulation experiments for ε suggest the value

$$\varepsilon = \frac{e}{\sqrt{\lambda}}$$

with e between 0.1 and 0.25; Fiksel (1988) has suggested $e = 0.1\sqrt{5}$.

The term $\nu(E_{x_i} \cap E_{x_j})$ in (2.3), which replaces the area a in a naive estimator, ensures the edge correction. This estimator is approximately unbiased also in the anisotropic case, see Fiksel (1988). The second estimator is a counterpart to Ripley's edge corrected estimator of the K-function. Here the term b_{ij} ensures the edge correction. This estimator is approximately unbiased only in the case of an isotropic point process. (For large values of t a modification has been suggested by Ohser which is analogous to that of the estimator of the K-function; see Stoyan *et al.* ((1987), p. 125)). The third estimator is an isotropized variant of the first one; it has been suggested by Ohser (1991). This is perhaps the most elegant: here the correction term $\gamma_E(t)$ has to be calculated only once for fixed t. Its simple form is suitable both for programming and mathematical calculations.

Some further estimators have been studied by Doguwa (1990).

3. Simulation experiments and their general results

The estimators in Section 2 are so complicated that their distributional properties can be investigated by simulation only. In this paper simulation is used for testing the quality of the approximations which will be suggested in the Sections 4 and 5. We have used earlier results of Doguwa (1990) and Fiksel (1988) and some simulations of our own. In all three cases three types of point process models have been investigated: the Poisson process, cluster processes and hard core processes.

Both Fiksel and ourselves have simulated Matern cluster processes, see Stoyan *et al.* ((1987), p. 143). These processes have parent points forming a Poisson process of intensity λ_p and daughter points uniformly distributed in discs of fixed radius R centred at the parent points; the number of points per cluster has a Poisson distribution of parameter μ. The parameters for the simulations have been:

$$\lambda_p = 0.16, \qquad R = 1.5, \qquad \text{and} \qquad \mu = 5 \qquad \text{(Fiksel)}$$

and

$$\lambda_p = 0.2, \qquad R = 0.5, \qquad \text{and} \qquad \mu = 10.$$

Doguwa has considered a so-called Thomas process, see Stoyan *et al.* (1987). The intensity of the parent process has been 20, the mean number of daughter points per cluster 5 and the variance parameter $\sigma = 0.035$.

Both Fiksel and ourselves have simulated Matern's second hard core process, see Ripley (1977) and Stoyan *et al.* ((1987), p. 146). It results from a dependent thinning of a Poisson process of intensity λ_b. Interpoint distances smaller than $2R$ are impossible. The process parameters have been

$$\lambda_b = 1.25, \qquad \text{and} \qquad R = 0.25 \qquad \text{(Fiksel)}$$

and

$$\lambda_b = 0.2, \qquad \text{and} \qquad R = 0.15.$$

In contrast, Doguwa has used Matern's first hard core process, see Ripley (1977). His parameters have been, in Ripley's notation,

$$\alpha = 145 \qquad \text{and} \qquad 2R = 0.035.$$

In all three cases the windows have been rectangles with sizes as follows:

$$1 \times 1 \text{ (Doguwa)}, \qquad 15 \times 10 \text{ (Fiksel)} \qquad \text{and} \qquad 10 \times 10 \text{ (this paper)}$$

The band width has been chosen in Doguwa and Fiksel as $0.1\sqrt{5/\lambda}$, while we have used both this value and its half and double.

Doguwa has simulated 100 samples, Fiksel 25, and we 50.

All authors have considered pair correlation function estimators, while we have also considered product density estimators.

General tendencies

Doguwa's and our simulations have clearly shown that there are no essential differences between the three estimators $\hat{\varrho}_F(t)$, $\hat{\varrho}_O(t)$ and $\hat{\varrho}_R(t)$; only $\hat{\varrho}_F(t)$ has estimation variances a little larger than those of the other two. Thus for mathematical calculations we will consider the very simple estimator $\hat{\varrho}_O(t)$ and the corresponding pair correlation estimator.

All estimators turned out to have very small empirical biases with the exception of the case of very small values of t. They can be reduced by the reflection technique, see Doguwa (1990).

The estimation variances in the case of a Poisson process have a behaviour such as shown in Fig. 1. For small t they are decreasing with increasing t and then they remain constant for moderate t and increase again for large t. It is plausible that for small values of t large variances have to be expected: The estimators contain the term "$1/t$" and the double sum is formed by the few point pairs in the sample of a very short inter-point distance.

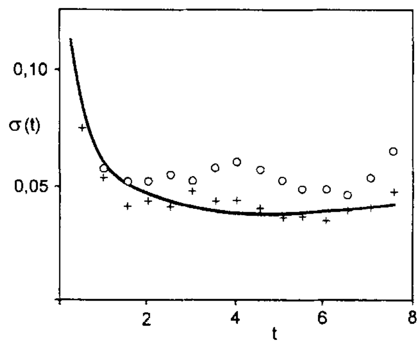

Fig. 1. Empirical standard deviations of the pair correlation function estimator based on $\hat{\varrho}_O(t)$ for our simulations compared with $\sigma(t)$ as given by (5.1). The estimation situation is the same both for calculations and simulations, $\lambda = 2$, 10×10 window and $\varepsilon = 0.158$. — $= \sigma(t)$, $\circ$ = Poisson process, $+$ = Matern hard core process. The values for the cluster process vary between 0.85 for small t and 0.27 for t greater than 2.

For the hard core processes, the behaviour of the estimation variances is more complicated. Relatively large values appear for t-values near to that t-value for which the discontinuity and maximum of the pair correlation function $g(t)$ appear. (Its theoretical form is shown for an example in Stoyan *et al.* ((1987), p. 121), and Fiksel has given values in his Table 3.) For those values of t, where the value of $g(t)$ is unity, the estimation variances tend to be constant. Their level is smaller than in the case of a Poisson process.

The estimation variances in the case of a cluster process are larger than for the Poisson process. They are very large for small t (where $g(t)$ is greater than unity) and decrease with increasing t. Even for those t-values, where $g(t)$ is unity, the estimation variances are still larger than in the Poisson process case. For details see the figures and tables in Doguwa (1990) and Fiksel (1988).

4. Asymptotic variance for Ohser's estimator in the case of a Poisson process

Ohser's estimator has the same structure as the quantity T on p. 30 of Ripley (1988),

$$T = \sum_{\substack{\text{sample points} \\ x \neq y}} \Phi(x, y).$$

Thus, the same method as there can be applied for calculating the variance. It consists in expressing the first two moments of T by means of factorial moment measures. They have a well-known simple form for the Poisson process. The result is

$$\text{var}\,\hat{\varrho}_O(t) = \frac{c(t)^2}{4}\left[4\lambda^3 S_1 + 2\lambda^2 S_2\right] \tag{4.1}$$

with

$$S_1 = \int_E \left\{ \int_E k(t - \|x-y\|)dy \right\}^2 dx,$$
$$S_2 = \int_E \int_E k(t - \|x-y\|)^2 dy dx,$$

and

$$c(t) = \frac{1}{\pi \gamma_E(t)}.$$

Similarly as in Ripley (1988), S_2 can be rewritten in polar coordinates as

$$S_2 = 2\pi \int_{t-\varepsilon}^{t+\varepsilon} k(t-r)^2 \gamma_E(r) r dr.$$

Using the formula for the kernel function $k(r)$ and the approximation (2.6) for the set covariance function $\gamma_E(r)$ yields the following (very precise) approximation for S_2:

$$S_2 = \frac{6}{5\varepsilon}\left(a\pi t - u\left(\frac{\varepsilon^2}{7} + t^2\right)\right). \tag{4.2}$$

A rough upper bound for S_1 can be obtained as in Ripley ((1988), p. 31). It is

$$\int_E k(t - \|x-y\|)dy \leq 2\pi \int_{t-\varepsilon}^{t+\varepsilon} k(t-r) r dr = 2\pi t.$$

Thus

$$S_1 \leq 4\pi^2 t^2 a. \tag{4.3}$$

A lower bound is obtained by restricting the integration for the outer integral of S_1 on $E \ominus b(o, t+\varepsilon)$:

$$S_1 \geq 4\pi^2 t^2 \nu(E \ominus b(o, t+\varepsilon)). \tag{4.4}$$

Here $\ominus$ denotes Minkowski subtraction and $b(x,v)$ is the disc of radius v centred at x.

In the case of a rectangular window of side lengths α and β it is

$$\nu(E \ominus b(o,t+\varepsilon)) = (\alpha - 2(t+\varepsilon))(\beta - 2(t+\varepsilon)),$$

for such values of t for which both factors are positive.

As it seems, these bounds are useful for practical applications. Our simulations have led to empirical estimation standard deviations which are well included by the approximations obtained by (4.1), (4.2), (4.3), and (4.4), as shown in Table 1. Only the values for $t = 1$ are given. This has two reasons: first, with increasing t the approximations lose their quality, in particular the lower bound. Second, the general experience reported in Section 3 suggests similar deviations also for larger values of t, and thus for practical use it is sufficient to calculate the bounds once for a small value of t. The approximation may be refined by using better approximations for S_1, which can be obtained in a similar way as the following approximations of estimation variances for pair correlation function estimators.

Table 1. Empirical estimation standard deviations of $\hat{\varrho}_O(1)$ from our simulations and corresponding bounds in the case of a Poisson process.

ε	lower bound	empirical value	upper bound
0.079	0.608	0.672	0.729
0.158	0.551	0.612	0.690
0.316	0.506	0.584	0.670

For approximating the variance of the pair correlation function estimator based on $\hat{\varrho}_O(t)$ again Ripley's method can be used. It is

$$\hat{g}_O(t) = \frac{a^2}{n^2}\hat{\varrho}_O(t), \qquad t \geq 0.$$

The form of this estimator is the same as that of $\hat{K}_O(t)$ on p. 35 of Ripley (1988), and it is therefore possible to use similar arguments for the calculation of its variance. Thus

$$\mathrm{var}_n\, \hat{g}_O(t) = \frac{c(t)^2}{4} a^2 \frac{n-1}{n^3}\left(\frac{4n-8}{a}S_1 - \frac{4n-6}{a^2}S^2 + 2S_2\right),$$

where var_n, as in Ripley (1988), denotes the variance in the case of a binomial process.

Here S_1 and S_2 are defined as above, while

$$S = \int_E \int_E k(t - \|x-y\|)dxdy.$$

In polar coordinates this integral can be rewritten as

$$S = 2\pi \int_{t-\varepsilon}^{t+\varepsilon} k(t-r)\gamma_E(r)r dr.$$

Similarly as S_2, S can be approximated by

$$S = 2\pi ta - 2u\left(\frac{\varepsilon^2}{5} + t^2\right). \tag{4.5}$$

Also as above, upper and lower bounds for S_1 can be used for deriving bounds for $\operatorname{var} \hat{g}_O(t)$. Here we give an approximation which is in the sense of Ripley. Obviously,

$$S_1 = 4\pi^2 t^2 \nu(E \ominus b(o, t+\varepsilon)) + \int_{E \setminus E \ominus b(o,t+\varepsilon)} \{\cdots\}^2 dx.$$

For a rectangle the first term is

$$4\pi^2 t^2 [a - u(t+\varepsilon) + 4(t+\varepsilon)^2].$$

The following integral is approximated by

$$\frac{1}{t+\varepsilon} \int_0^{t+\varepsilon} b(r) dr [u(t+\varepsilon) - 4(t+\varepsilon)^2],$$

where $b(r)$ is the length of the arc in E of the circle of radius $t+\varepsilon$ centred in E at a point of distance $t+\varepsilon-r$ from the boundary. This yields

$$\operatorname{var}_n \hat{g}_O(t) = \frac{c(t)^2}{4} a^2 \frac{n-1}{n^3}\left(X + \frac{n}{a}Y + \frac{n}{a^2}Z\right) \tag{4.6}$$

with

$$\begin{aligned} X &= \frac{12}{5\varepsilon}\left[a\pi t - u\left(t^2 + \frac{\varepsilon^2}{7}\right)\right] - 8\pi^2 t^2 + \frac{6}{a^2}C + \frac{32}{a}BD, \\ Y &= -16BD, \\ Z &= -4C, \end{aligned}$$

and

$$\begin{aligned} B &= \pi^2 t^2 - (t+\varepsilon)^2(\pi-1)^2, \\ C &= 4u^2\left(\frac{\varepsilon^2}{5} + t^2\right)^2 - 8\pi tau\left(\frac{\varepsilon^2}{5} + t^2\right)^2, \\ D &= u(t+\varepsilon) - 4(t+\varepsilon)^2. \end{aligned}$$

Taking expectations over n, which can be justified as in Ripley ((1988), p. 37), finally gives the approximation

$$\operatorname{var} g_O(r) = \frac{c(t)^2}{4\lambda^2}\left(X + \lambda Y + \frac{\lambda}{a}Z\right). \tag{4.7}$$

Table 2. Comparison of empirical and estimated standard deviations of the pair correlation function estimators in the case of a Poisson process for Fiksel's ε.

Simulations made by	t	empirical value	approximation obtained by (4.7)
Doguwa	0.1	0.125	0.205
	0.3	0.11	0.167
Fiksel	1.8	0.07	0.106
	4.5	0.08	0.136
authors	1	0.057	0.075
	3	0.052	0.097

It yields acceptable values, as comparisons with simulated variances have shown. Table 2 gives the figures.

5. A heuristic approximation of the estimation variance of the pair correlation function estimator

In the following we show how a simple Poisson approximation can yield acceptable approximations for the estimation variances of the pair correlation function estimators also for non-Poisson point processes.

For known λ the estimator of the pair correlation function can be written as

$$g_O^*(t) = \frac{c(t)}{2\lambda^2} \sum_{\substack{\text{sample points} \\ x \neq y}} k(t - \|x - y\|), \qquad r \geq 0.$$

If always only one member of the summand pairs $k(t - \|x - y\|)$ and $k(t - \|y - x\|)$ is included, then the form

$$g_O^*(t) = \frac{c(t)}{\lambda^2} \sum_{i=1}^{N_t} X_{t,i}$$

is obtained, where N_t is the number of the point pairs within distance t. We will approximate its variance and use the result also in the case of unknown λ.

Clearly, the summands of this sum are only weakly variable, since only point pairs of an inter-point distance close to t contribute to the sum. The main variability comes from the number N_t of those point pairs. This number is small compared with the number of all pairs of points of the process in the window. Thus, observing a point pair of a distance close to t is a "rare event", and a Poisson approximation is plausible. Consequently, distributional properties of the sum above will be determined assuming that

$$N_t \text{ has a Poisson distribution of mean } \mu_t$$

and

the $X_{t,i}$ are i.i.d.

The $X_{i,t}$ are clearly of the form $X_{t,i} = k(t - D_i)$ with $t - \varepsilon \leq D_i \leq t + \varepsilon$. It seems to be natural to assume that the D_i are "completely random", i.e. uniformly distributed on the interval $[t - \varepsilon, t + \varepsilon]$. This yields the first two moments of $X_{t,i}$ as

$$EX_{t,i} = \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} k(s) ds = \frac{1}{2\varepsilon}$$

and

$$EX_{t,i}^2 = \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} k^2(s) ds = \frac{9}{32\varepsilon^3} \int_{-\varepsilon}^{\varepsilon} \left(1 - \frac{s^2}{\varepsilon^2}\right)^2 ds = \frac{0.3}{\varepsilon^2}.$$

The mean μ_t of N_t can approximately be obtained by using the "unbiasedness" property of $g_O^*(t)$, using

$$g(t) = \frac{c(t)}{\lambda^2} E \sum_{i=1}^{N_t} X_{t,i} = \frac{c(t)}{\lambda^2} \mu_t EX_{t,i}.$$

This yields

$$\mu_t = \frac{2\varepsilon g(t)\lambda^2}{c(t)}.$$

Using this μ_t, the variance $\sigma^2(t)$ of $g_O^*(t)$ can be approximated by

$$\sigma^2(t) = \frac{c(t)^2}{\lambda^4} \left[\mu_t \operatorname{var} X_{t,i} + \operatorname{var} N_t (EX_{t,i})^2\right] = \frac{c(t)^2}{\lambda^4} \mu_t EX_{t,i}^2.$$

The final result is

$$\sigma^2(t) = \frac{0.6c(t)g(t)}{\varepsilon\lambda^2} \tag{5.1}$$

with

$$c(t) = \frac{1}{\pi t \gamma_E(t)}.$$

The qualitative behaviour of this $\sigma(t)$ in dependence on window size, band width ε and $g(t)$ is in good agreement with expectations of statisticians and empirical results. The function $\sigma(t)$ is decreasing for small t, has a minimum for moderate t and is then increasing in t. For $t = 0$ it has a pole, which well corresponds to the behaviour of all estimators (2.3) to (2.5) at $t = 0$. Taken as a function of ε, it is decreasing, and it increases with increasing $g(t)$. In the case of a Poisson process it gives good approximations, in particular for small and moderate t, also for ε-values different from those suggested by Fiksel (1988). Figure 1 shows $\sigma(t)$ in comparison with empirical values resulting from our simulations. The deviations obtained in this example are typical for the approximation.

Also for hard core processes (5.1) gives good approximations, which are better than in the case of a Poisson process, see also Fig. 1. In contrast, (5.1) is not acceptable for cluster processes with a higher degree of clustering, as discussed in all simulations: the empirical standard deviations are much greater than those predicted by (5.1).

References

Doguwa, S. I. (1990). On edge-corrected kernel-based pair-correlation function estimators for point processes, *Biometrical J.*, **32**, 95–106.

Fiksel, T. (1988). Edge-corrected density estimators for point processes, *Statistics*, **19**, 67–75.

Ohser, J. (1991). On estimation of pair correlation functions, *Geometrical Problems of Image Processing* (eds. U. Eckhardt, A. Hübler, W. Nagel, and G. Werner), 147–152, Akademie Verlag, Berlin.

Penttinen, A. K. and Stoyan, D. (1989). Statistical analysis for a class of line segment processes, *Scand. J. Statist.*, **16**, 152–168.

Ripley, B. D. (1977). Modelling spatial patterns (with discussion), *J. Roy. Statist. Soc. Ser. B*, **39**, 172–212.

Ripley, B. D. (1981). *Spatial Statistics*, Wiley, New York.

Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.

Stoyan, D., Kendall, W. S. and J. Mecke (1987). *Stochastic Geometry and Its Applications*, Wiley, Chichester.