Statistica Sinica **5**(1995), 221-234

AN APPROACH TO ASYMPTOTIC INFERENCE FOR SPATIAL POINT PROCESSES

Michael L. Stein

University of Chicago

Abstract. The standard asymptotic regime for studying the properties of estimators of the reduced second moment function $K(t)$ of a spatial point process is to fix the law of the process and t and let the observation region grow. No results comparing the asymptotic variability of these estimators are available for processes other than Poisson. This paper obtains asymptotic results for a large class of stationary point processes by using a different asymptotic regime. Specifically, the distance t at which K is to be estimated is allowed to grow with the observation region. Using this asymptotic setup, it is shown that using the notion of projection of a U-statistic leads to estimators that converge to the truth at a faster rate than standard estimators not using this projection idea.

Key words and phrases: Edge effects, interpoint distance distribution, reduced second moment function, U-statistics.

1. Introduction

Correcting for edge effects is a fundamental problem in inference for spatial point processes (Ripley (1988), Chapters 3 and 4). The impact of edge effects on estimating the reduced second moment function has received particular attention: see Ripley (1988, Chapter 3) and the references therein and Stein (1991, 1993). For a stationary point process N in $\mathbb{R}^d$ with intensity λ, the reduced second moment measure $\mathcal{K}(\cdot)$ is defined for Borel sets B as $\mathcal{K}(B) = \lambda^{-1}E\{N(B\backslash\{0\})|0 \in N\}$, where $N(B) = N \cap B$. If the process is isotropic, then $\mathcal{K}(\cdot)$ can be recovered from the reduced second moment function $K(t) = \lambda^{-1}E\{N(b(0,t)\backslash\{0\})|0 \in N\}$, where $b(0,t)$ is the ball of radius t centered at 0. All second moments of N are determined by λ and $\mathcal{K}(\cdot)$.

Consider estimating $K(t)$ based on observing a stationary isotropic point process N in a bounded measurable set $A \subset \mathbb{R}^d$ with $\nu_d(A) = a$, where ν_d indicates Lebesgue measure in d dimensions. The basic problem is that for a point x in $N \cap A$ within t of a boundary of A, we only have partial information on the number of points in N within t of x. Two commonly used methods for accounting for this effect are the isotropic (Ripley (1976)) and rigid motion corrections (Miles (1974), Ohser and Stoyan (1981)). The idea behind both of these procedures is first to estimate $\lambda^2 a^2 K(t)$ unbiasedly using an estimator of the form $\sum' \phi(x,y)$, where $\sum'$ means summation over all distinct ordered pairs of points of N in A. Specifically, for all stationary isotropic N, ϕ is required to satisfy $E\{\sum' \phi(x,y)\} = \lambda^2 a^2 K(t)$. Without loss of generality, we can assume ϕ is symmetric in its arguments. Then $K(t)$ is estimated by dividing $\sum' \phi(x,y)$ by an estimator of $\lambda^2 a^2$, typically $N(A)^2$ or $N(A)\{N(A)-1\}$. The resulting estimator for $K(t)$ is, in general, biased. Stein (1993) argued that any estimator of $\lambda^2 a^2 K(t)$ can be improved upon by viewing $\sum' \phi(x,y)$ as a U-statistic and applying the notion of projecting U-statistics (Serfling (1980)). Specifically, consider N Poisson, in which case, conditional on $N(A)=n$, the n points in A are independent and uniformly distributed in A. Next, for each n, find the function g_n satisfying $\int_A g_n(x)dx = 0$ that minimizes

$$\operatorname{Var}\Big(\sum{}' \phi(x,y) - \sum_{x\in A} g_n(x)\Big)$$

under the uniform distribution for the n points in A. The minimizing $g_n(x)$ is $2(n-1)\bar{\phi}(x)$, where

$$\bar{\phi}(x) = a^{-1}\int_A \phi(x,y)dy - a^{-2}\int_{A^2} \phi(y,z)dydz,$$

leading to the estimator of $\lambda^2 a^2 K(t)$ given by

$$\sum{}' \phi(x,y) - 2\{N(A)-1\}\sum_{x\in N\cap A} \bar{\phi}(x) = \sum{}' \phi^P(x,y), \tag{1.1}$$

where $\phi^P(x,y) = \phi(x,y) - \bar{\phi}(x) - \bar{\phi}(y)$. I will call an estimator whose numerator is of this form and whose denominator is $N(A)\{N(A)-1\}$ a projected estimator. The idea is that replacing ϕ by ϕ^P should induce very little bias in the estimation of $\lambda^2 a^2 K(t)$ for any stationary N and at the same time lower the variance when N is Poisson. Stein (1993) provides both theory and simulations to support these contentions.

To study the asymptotic variability for edge-corrected estimators of $K(t)$, Ripley (1988) uses the approach of taking t and the law of N as fixed and letting the observation region A grow. In two dimensions, assuming the boundary of A can be approximated by a straight edge in an appropriate sense as A grows, he obtains expansions for the variance of various edge-corrected estimators of $K(t)$ of the form

$$\frac{2}{\lambda^2}\Big(\frac{\pi t^2}{a} + c_1\frac{pt^3}{a^2} + c_2\frac{pt^5}{a^2}\Big) + o\Big(\frac{p}{a^2}\Big),$$

where p is the perimeter of A. Stein (1993) shows that among a large class of estimators the minimum possible value for c_1 is $2/3$ and for c_2 is 0 and that by applying the projection idea to the rigid motion correction due to Miles (1974) and Ohser and Stoyan (1981), these minimum values are both attained. As A grows, p will typically be $o(a)$, for example, if we have a sequence of observation regions $A_1, A_2, \ldots$ that is a convex averaging sequence: each set is convex and the radii of their largest inscribed circles tend to infinity (Daley and Vere-Jones (1988)). Thus, to the highest order term, all reasonable edge-corrected estimators have the same asymptotic variance and it is only in the second term that differences emerge. Since the fraction of points within t of an edge of A decreases as A increases, this result is not surprising.

A clear weakness in these results is that they say nothing about the variability of the estimators when N is not Poisson. For non-Poisson N, one would expect the differences in the variability of the various edge-corrected estimator again to occur only in the second order term of the asymptotic expansion for the variances. It appears to be rather difficult to obtain such a second-order expansion for non-Poisson N. Heinrich (1988) shows that a large class of statistics, including estimators of $\lambda^2 K(t)$, is asymptotically normal as A grows for Poisson cluster processes. Bertram, Wendrock and Stoyan (1993) give a heuristic approximation to the variance of estimators of the pair correlation function for non-Poisson processes. However, the results in these works do not distinguish between the behavior of different edge-corrected estimators (Heinrich (1988)) and thus do not provide any guidance as to which edge corrections work best for non-Poisson processes.

The approach taken here to distinguish the asymptotic behavior of different edge-corrected estimators for non-Poisson processes is to fix the ratio between t and the size of the observation region A. Specifically, for a stationary process N, consider estimating $K(\beta t)$ based on observing N in βA as $\beta \to \infty$, where $\beta A = \{x : x/\beta \in A\}$. Equivalently, we can take A and t to be fixed and consider the class of processes N_β defined by $N_\beta(B) = N(\beta B)$ and then consider what happens as $\beta \to \infty$. The two viewpoints are formally identical; the second is more convenient mathematically and is the one used here. Since, under this asymptotic paradigm, the fraction of points near an edge of A stays roughly constant as β increases, one might expect that the variances of the various edge-corrected estimators will differ in the constant multiplying their leading term. However, for a process in $\mathbb{R}^d$ whose first four factorial moment measures satisfy certain technical conditions, the main result of this work shows that for an unprojected estimator, the error is $O_p(\beta^{-d/2})$, and for a projected estimator, the error is $O_p(\beta^{-d})$. That is, projecting actually increases the rate of convergence of the estimator. For N Poisson, this result follows immediately from the properties of U-statistics (Serfling (1980)); the faster rate of convergence for projected estimators corresponds to the fact that projected estimators are essentially degenerate U-statistics.

Baddeley and Gill (1993) use another asymptotic approach in which edge effects are kept constant to study estimates of the empty space function F, the distribution of the distance from a fixed point in space to the nearest point in the process, and the nearest neighbor distribution G, the distribution of the distance from a point in the process to its nearest neighbor in the process. Specifically, they consider what happens if many independent replicates of a Poisson process with low intensity are observed through a fixed, bounded and convex window. These sparse Poisson limit results do not tell us anything about how the estimators behave for non-Poisson processes. On the other hand, the approach taken here may not provide much insight for studying estimates of F or G, even if it could be carried out. In particular, suppose N is Poisson in, say, two dimensions, with parameter λ and define $F_\beta(t) = F(\beta t)$ as the empty space function for N_β, and define G_β similarly. Then $F_\beta(t) = G_\beta(t) = 1 - \exp\{-\pi\lambda\beta^2 t^2\}$, which tends to 1 exponentially fast in the intensity $\lambda\beta^2$ of N_β. It would be of greater interest to study F at distances where it is not so near to 0 or 1. The Baddeley and Gill approach considers estimates of F and G at distances where the truth tends to 0, but the rate of convergence is proportional to the intensity.

Section 2 states the main results comparing projected and unprojected estimators. Section 3 discusses their relevance to problems of practical interest. Section 4 provides proofs.

2. Main Results

Consider a stationary point process N on $\mathbb{R}^d$ and define for each $\beta > 0$ the process N_β by $N_\beta(B) = N(\beta B)$ for all Borel sets B. Suppose A is the observation region with $\nu_d(A) = a$ and consider a statistic of the form

$$T_\beta = \frac{\sum'_\beta \psi(x,y)}{\sum'_\beta \tau(x,y)}, \tag{2.1}$$

where ψ and τ are symmetric in their arguments and $\sum'_\beta$ means summation over all distinct pairs of points x and y that are in $N_\beta \cap A$. If the denominator in (2.1) is 0, define $T_\beta = 0$. When $\tau(x,y) \equiv 1$, which is commonly the case for existing estimators, the denominator is $N_\beta(A)\{N_\beta(A) - 1\}$. Ripley (1988) and Stein (1993) use the denominator $N_\beta(A)^2$; to the order of approximation studied here, the difference between the two denominators does not affect the variability of T_β. The form of the denominator in (2.1) will be slightly more convenient here.

Suppose ψ and τ are absolutely integrable over A^2 and let $c_\psi = \int_{A^2} \psi(x,y)dxdy$

and $c_\tau = \int_{A^2} \tau(x,y)dxdy$ where $c_\tau > 0$. Then

$$T_\beta = \frac{c_\psi}{c_\tau} + \frac{\sum_\beta' f(x,y)}{\sum_\beta' \tau(x,y)},$$

where

$$f(x,y) = \psi(x,y) - \frac{c_\psi}{c_\tau}\tau(x,y). \tag{2.2}$$

By construction,

$$\int_{A^2} f(x,y)dxdy = 0. \tag{2.3}$$

The asymptotic variance of T_β depends critically on whether or not

$$\int_A \Big\{ \int_A f(x,y)dy \Big\}^2 dx = 0, \tag{2.4}$$

which is equivalent to $\int_A f(x,y)dy = 0$ for almost every $x \in A$. For example, (2.4) holds if $\tau(x,y) \equiv 1$, f is as given in (2.2) and ψ in (2.2) is of the form ϕ^P given in (1.1). Thus, (2.4) holds for the estimator of $K(t)$ recommended in Stein (1993) but does not hold for any of the standard estimators of $K(t)$ discussed by Ripley (1988). Suppose that N has finite fourth moments and that the first four factorial moment measures have densities with respect to Lebesgue measure. Let $m_{[k]}$ denote the density of the kth factorial moment measure of N, which implies that the density of the kth factorial moment measure of N_β at $(x_1, \ldots, x_k)$ is $\beta^{dk} m_{[k]}(\beta x_1, \ldots, \beta x_k)$. Then (Ripley (1988))

$$\begin{aligned}
&\mathrm{Var}_\beta\Big\{ \sum\nolimits_\beta' f(x,y) \Big\} \\
&= \beta^{4d} \int_{A^4} \{ m_{[4]}(\beta \boldsymbol{x}) - m_{[2]}(\beta x_1, \beta x_2) m_{[2]}(\beta x_3, \beta x_4) \} f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} \\
&\quad + 4\beta^{3d} \int_{A^3} m_{[3]}(\beta x_1, \beta x_2, \beta x_3) f(x_1,x_2) f(x_1,x_3) dx_1 dx_2 dx_3 \\
&\quad + 2\beta^{2d} \int_{A^2} m_{[2]}(\beta x_1, \beta x_2) f(x_1,x_2)^2 dx_1 dx_2,
\end{aligned} \tag{2.5}$$

where $\boldsymbol{x} = (x_1, x_2, x_3, x_4), d\boldsymbol{x} = dx_1 dx_2 dx_3 dx_4$ and Var_β and E_β indicate expectations taken under the law of N_β. To approximate this variance for large β, we need some sort of mixing condition on the first four factorial moment densities. To this end, for two finite sets of points U and V in $\mathbb{R}^d$, define the distance between U and V as

$$d(U,V) = \min_{x \in U,\ y \in V} |x - y|.$$

The main results here assume that for all U and V containing i and j points, respectively, where $i \geq 1, j \geq 1$ and $i+j \leq 4$, there exists a nonnegative function h such that $h(t) = o(t^{-3d})$ as $t \to \infty$ and

$$|m_{[i+j]}(U,V) - m_{[i]}(U)m_{[j]}(V)| \leq h(d(U,V)). \tag{2.6}$$

If $m_{[4]}$ is bounded, this condition implies that the fourth reduced cumulant measure is integrable, a condition used by Mase (1982), Jolivet (1981) and Karr (1987) to obtain asymptotic results for point processes using the standard asymptotic approach of fixing the characteristic of the process to be estimated and letting the observation region grow. These works consider only first order asymptotics and so do not provide a basis for studying edge effects. Whether the integrability of the fourth reduced cumulant measure would be sufficient here is unclear, although (2.6) does appear to be stronger than necessary.

Define the covariance density function $\alpha(y) = m_{[2]}(0,y) - \lambda^2$ (Daley and Vere-Jones (1988)). If $m_{[2]}$ is bounded, then (2.6) with $i = j = 1$ implies $\int_{\mathbb{R}^d} |\alpha(y)|dy < \infty$. We have the following results:

Theorem 1. *Suppose N is a stationary point process whose first four factorial moment densities are bounded and satisfy* (2.6) *and A is a finite union of bounded convex sets and $\nu_d(A) = a$. Furthermore, assume f is a symmetric, bounded measurable function from A^2 to $\mathbb{R}$ satisfying* (2.3) *and that for almost every $x \in A, f(x,y)$ is continuous in y for almost every $y \in A$. Then as $\beta \to \infty$*

$$\beta^{-3d}\mathrm{Var}_\beta(\textstyle\sum'_\beta f(x,y)) \longrightarrow 4\lambda^2\Big(\lambda + \displaystyle\int_{\mathbb{R}^d} \alpha(y)dy\Big) \int_A \Big\{\int_A f(x,y)dy\Big\}^2 dx.$$

Theorem 2. *Assume the same conditions on N, A and f as in Theorem* 1. *If, in addition, f satisfies* (2.4), *then as $\beta \to \infty$*

$$\beta^{-2d}\mathrm{Var}_\beta\Big(\textstyle\sum'_\beta f(x,y)\Big) \longrightarrow 2\Big(\lambda + \displaystyle\int_{\mathbb{R}^d} \alpha(y)dy\Big)^2 \int_{A^2} f(x,y)^2 dxdy.$$

The proof of Theorem 2 is given in Section 4. The proof of Theorem 1 is similar and is omitted.

Since

$$\beta^{-2d}\textstyle\sum'_\beta \tau(x,y) \xrightarrow{L^2} c_\tau\lambda^2,$$

it is tempting to conclude, for example, that under the conditions of Theorem 2,

$$\beta^{2d}\mathrm{Var}_\beta\Big(\frac{\sum'_\beta f(x,y)}{\sum'_\beta \tau(x,y)}\Big) \longrightarrow \frac{2}{c_\tau^2\lambda^4}\Big(\lambda + \int_{\mathbb{R}^d} \alpha(y)dy\Big)^2 \int_{A^2} f(x,y)^2 dxdy. \tag{2.7}$$

However, this result does not immediately follow from Theorem 2, because of the difficulty in computing the moments of ratios.

When N is Poisson and $\tau(x,y) \equiv 1$, it is not difficult to prove (2.7) holds with $\alpha(y) \equiv 0$. Moreover, using the asymptotic theory of U-statistics, the asymptotic distribution of T_β, defined in (2.1), is easily obtained. Specifically, for N Poisson, $\tau(x,y) \equiv 1$, f defined by (2.2) and A and f satisfying the conditions of Theorem 1,

$$\beta^{d/2}(T_\beta - \frac{c_\psi}{a^2}) \xrightarrow{\mathcal{L}} N\left(0, \frac{4}{\lambda a^4}\int_A \left\{\int_A f(x,y)dy\right\}^2 dx\right), \tag{2.8}$$

which follows by Theorem A, page 192 of Serfling (1980) by first conditioning on $N_\beta(A)$ and then unconditioning. Similarly, if f satisfies the conditions of Theorem 2, then using the fact that conditional on the value of $N_\beta(A)$, $\sum_\beta' f(x,y)$ is a degenerate U-statistic (Serfling (1980)),

$$\beta^d\Big(T_\beta - \frac{c_\psi}{a^2}\Big) \xrightarrow{\mathcal{L}} (2\lambda a)^{-1}\sum \lambda_j(\chi^2_{1j} - 1), \tag{2.9}$$

where $\chi^2_{11}, \chi^2_{12}, \ldots$ are independent χ^2_1 random variables and $\lambda_1, \lambda_2, \ldots$ are eigenvalues associated with f: corresponding to λ_j there exists a function h_j not identically 0 on A such that $a^{-1}\int_A f(x,y)h_j(y)dy = \lambda_j h_j(x)$ for almost every $x \in A$.

Silverman (1978) studied the asymptotic behavior of the natural estimator of the interpoint distance distribution, given by (2.1) with $\psi(x,y) = I\{|x-y| \le t\}$ and τ identically 1, where $I\{\cdot\}$ is an indicator function. He showed that for independent observations uniformly distributed on a surface with no edges, such as a sphere or torus, as the number of observations tends to infinity, this estimator posesses an asymptotic distribution of the form (2.9), and that otherwise, (2.8) gives the limiting distribution. The results here provide a clearer understanding as to why regions without edges yield such a radically different result. The point is that for regions without edges, the estimator considered by Silverman (1978) is already a projected estimator. However, for other regions we can easily apply the projection technique described in Section 1 to obtain an estimator of the interpoint distance distribution satisfying (2.4). For projected estimators, there is no sharp distinction in their asymptotic behavior depending on the presence or absence of edges: for all observation regions satisfying the conditions in Theorem 1, (2.9) holds as the number of independent uniformly distributed observations tends to infinity.

3. Discussion

Let us now consider the implications of the results of the previous section for the estimation of the reduced second moment function of an isotropic process. In this case, $\alpha(y)$ depends on y only through $|y|$, so we define $\bar{\alpha}(|y|) = \alpha(y)$. Then for N_β as defined in the previous section,

$$K_\beta(t) = \mu_d t^d + \frac{\omega_d}{\lambda^2 \beta^d} \int_0^{\beta t} \bar{\alpha}(r) r^{d-1} dr,$$

where μ_d and ω_d are the volume and surface area, respectively, of the unit sphere in d dimensions. Under (2.6), we have the approximation

$$K_\beta(t) = \mu_d t^d + \frac{\omega_d}{\lambda^2 \beta^d} \int_0^{\infty} \bar{\alpha}(r) r^{d-1} dr + o(\beta^{-2d}). \tag{3.1}$$

First, consider the bias characteristics of estimators of $K_\beta(t)$. For both the rigid motion (Miles (1974), Ohser and Stoyan (1981)) and isotropic corrections (Ripley (1976)) we can take $\tau(x,y) \equiv 1$ in (2.1), in which case, ϕ satisfies $\beta^{-2d} E_\beta\{\sum'_\beta \phi(x,y)\} = \lambda^2 a^2 K_\beta(t)$ for all stationary isotropic N. Thus, for the projected versions of ϕ (see (1.1)) we have, under (2.6),

$$\begin{aligned}\beta^{-2d} E_\beta\{{\textstyle\sum'_\beta} \phi^P(x,y)\} &= \lambda^2 a^2 K_\beta(t) - 2\beta^{-2d} \int_{A^2} \bar{\phi}(x) \alpha(\beta(x-y)) dx dy \\ &= \lambda^2 a^2 K_\beta(t) + O(\beta^{-d-1}),\end{aligned}$$

using $\int_A \bar{\phi}(x) dx = 0$. Thus, in this asymptotic regime, $\beta^{-2d} \sum'_\beta \phi^P(x,y)$ has a bias that is asymptotically negligible relative to $K_\beta(t) - \mu_d t^d$, the deviation of $K_\beta(t)$ from the reduced second moment function for a Poisson process. On the other hand, under (2.6) it can be shown that

$$\beta^{-2d} E_\beta[N_\beta(A)\{N_\beta(A) - 1\}] = \lambda^2 a^2 + \beta^{-d} a \int_0^\infty \bar{\alpha}(r) r^{d-1} dr + o(\beta^{-d})$$

and

$$\beta^{-2d} E_\beta[N_\beta(A)^2] = \lambda^2 a^2 + \beta^{-d} a \Big\{\lambda + \int_0^\infty \bar{\alpha}(r) r^{d-1} dr\Big\} + o(\beta^{-d}).$$

Thus, no matter which denominator we use, we see that the bias is $O(\beta^{-d})$ and in general not $o(\beta^{-d})$, so it is of a higher order than the bias in the numerator caused by projecting. This result is in agreement with simulations reported in Stein (1993), in which both projected and unprojected estimators sometimes exhibited substantial bias but that projecting did not noticeably change the bias of an estimator.

Next, consider the variation of T_β about $K_\beta(t)$. For both the projected and unprojected rigid motion and isotropic corrections,

$$T_\beta = \mu_d t^d + \frac{\sum'_\beta f(x,y)}{N_\beta(A)\{N_\beta(A) - 1\}}, \tag{3.2}$$

where f satisfies (2.3). Thus, from Theorem 1 and (3.1), $T_\beta - K_\beta(t) = O_p(\beta^{-d/2})$. For the projected estimators, f in (3.2) satisfies (2.4), so that $T_\beta - K_\beta(t) = O_p(\beta^{-d})$.

The specific forms for the asymptotic variance of $\sum'_\beta f(x,y)$ in Theorems 1 and 2 provide some insight into the behavior of the various estimators of $K_\beta(t)$ for N not Poisson. In both theorems, the answer factors into a term depending on $\lambda + \int_{\mathbb{R}^d} \alpha(y)dy$ and a term depending on f. An expected result is that for clustered processes, for which α tends to be positive, the limiting variance is increased relative to the Poisson in both theorems. Less obviously, since the term $\lambda + \int_{\mathbb{R}^d} \alpha(y)dy$ is squared in Theorem 2 and not in Theorem 1, we should expect the relative advantage due to projecting to be somewhat less for a clustered process than for a Poisson process. Conversely, for a process that has points more evenly spaced than the Poisson, α tends to be negative so that the improvement due to projecting should be particularly large in such cases. Note that these conclusions do not depend on f since in both theorems the asymptotic variance factors into a term depending on λ and α and a term depending on f. This factorization further implies that among some specified class of functions, the one that asymptotically minimizes $\mathrm{Var}_\beta(\sum'_\beta f(x,y))$ does not depend on the process N as long as (2.6) is satisfied.

The asymptotic approach used here essentially assumes that the distance at which we are estimating the reduced second moment function is large relative to the distances over which there are nonnegligible dependencies in the process that can be captured by the first four moments. In particular, it is important to note that the bias and spread of even a projected estimator of $K_\beta(t)$ are of order β^{-d}, which is the same order as $K_\beta(t) - \mu_d t^d$, which by (3.1), is the deviation of $K_\beta(t)$ from the Poisson result. Thus, we should not expect to be able to detect, reliably, variations from the Poisson model over short distances by looking at estimates of K at distances much longer than the scale on which nontrivial dependencies exist. To detect deviations from the Poisson models over small distances, we need to look at estimates of K at these same small distances.

The asymptotic approach used by Ripley (1988) of fixing the law of the process and t and letting the observation region grow is better suited for investigating this problem of detecting deviations from the Poisson model over small distances. However, for non-Poisson processes, it appears quite difficult to carry out the second-order variance calculations necessary to assess the effect of various edge corrections. Furthermore, there may be circumstances where deviations from randomness at larger distances are of interest even when the process is obviously not Poisson at shorter distances. For example, when studying the locations of trees in a forest (Diggle (1983)), there will often be a short-range repulsion between trees just due to the physical extent of the branches or roots. In such cases, it may still be of interest to detect departures from randomness over larger scales due, for example, to heterogeneity in soil conditions or constraints on the dispersal of seeds. The results obtained here show that even if the process is obviously not Poisson over short distances, there may still be an enormous benefit to projecting if we want to find departures from randomness over longer distances. Finally, it is interesting to note that the mathematical formulation of a problem can affect our perception of what asymptotic approach is natural. For example, as noted earlier, Silverman (1978) studied the behavior of statistics as the number of independent uniformly distributed observations on a fixed region increases. Because we just have the standard asymptotic setting of a sequence of independent identically distributed observations, it does not seem unusual in any way. However, this scenario is also essentially the special case of the one used in Section 2 where N is Poisson. Thus, if one is willing to accept the setting in Silverman (1978) as an appropriate asymptotic formulation, it follows that the approach used here is also appropriate in at least some circumstances.

4. Proof of Theorem 2

Define

$$J(\beta) = \int_{A^4} \{m_{[4]}(\beta \boldsymbol{x}) - m_{[2]}(\beta x_1, \beta x_2) m_{[2]}(\beta x_3, \beta x_4)\} f(x_1, x_2) f(x_3, x_4) d\boldsymbol{x}.$$

The main step in proving Theorem 2 is to show that as $\beta \to \infty$,

$$\beta^{2d} J(\beta) \longrightarrow 2\Big\{\int_{\mathbb{R}^d} \alpha(y) dy\Big\}^2 \int_{A^2} f(x,y)^2 dx dy. \tag{4.1}$$

To obtain (4.1), use the symmetry properties of $m_{[4]}$ and f to write

$$J(\beta) = \int_{A^4} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} + 2\int_{A^4} \alpha(\beta(x_1 - x_3))\alpha(\beta(x_2 - x_4)) f(x_1, x_2) f(x_3, x_4) d\boldsymbol{x},$$

where

$$\begin{aligned}\gamma_\beta(\boldsymbol{x}) =& \{m_{[4]}(\beta \boldsymbol{x}) - \lambda^4 - \alpha(\beta(x_1 - x_2))\alpha(\beta(x_3 - x_4)) - \alpha(\beta(x_1 - x_3))\alpha(\beta(x_2 - x_4)) \\ & - \alpha(\beta(x_1 - x_4))\alpha(\beta(x_2 - x_3))\} f(x_1, x_2) f(x_3, x_4).\end{aligned}$$

Then (4.1) follows by showing

$$\begin{aligned}&\beta^{2d} \int_{A^4} \alpha(\beta(x_1 - x_3))\alpha(\beta(x_2 - x_4)) f(x_1, x_2) f(x_3, x_4) d\boldsymbol{x} \\ &\longrightarrow \Big\{\int_{\mathbb{R}^d} \alpha(y) dy\Big\}^2 \int_{A^2} f(x,y)^2 dx dy \end{aligned} \tag{4.2}$$

and

$$\beta^{2d} \int_{A^4} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} \;\longrightarrow\; 0. \tag{4.3}$$

To obtain (4.2), make the substitutions $y = \beta(x_3 - x_1)$ and $z = \beta(x_4 - x_2)$ in (4.2), yielding

$$\begin{aligned}
&\beta^{2d} \int_{A^4} \alpha(\beta(x_1 - x_3))\alpha(\beta(x_2 - x_4)) f(x_1, x_2) f(x_3, x_4) d\boldsymbol{x} \\
&= \int_{A^2} \int_{\beta A_{-x_1}} \int_{\beta A_{-x_2}} \alpha(y)\alpha(z) f(x_1, x_2) f\Big(x_1 + \frac{y}{\beta}, x_2 + \frac{z}{\beta}\Big) dz dy dx_1 dx_2,
\end{aligned}$$

where $A_x = \{y : y - x \in A\}$ indicates set translation and set translation is done before scalar multiplication, so that $\beta A_{-x_1} = \{y : \frac{y}{\beta} + x_1 \in A\}$. Now, for almost every $(x_1, x_2, y, z) \in A^2 \times \mathbb{R}^{2d}$,

$$\begin{aligned}
&f(x_1, x_2) f\Big(x_1 + \frac{y}{\beta}, x_2 + \frac{z}{\beta}\Big)\alpha(y)\alpha(z) I\{y \in \beta A_{-x_1}, z \in \beta A_{-x_2}\} \\
&\quad \longrightarrow f(x_1, x_2)^2 \alpha(y)\alpha(z).
\end{aligned} \tag{4.4}$$

Using f bounded, the left-hand side of (4.4) is dominated by some constant times $|\alpha(y)\alpha(z)|$, which is integrable over $A^2 \times \mathbb{R}^{2d}$, so (4.2) follows by dominated convergence.

To obtain (4.3), we need to introduce some notation. Let $Q = \{\{i, j\} : 1 \leq i < j \leq 4\}$ and for any $S \subset Q$, define

$$B(S, r) = \{\boldsymbol{x} : \boldsymbol{x} \in A^4, |x_i - x_j| \leq r \text{ for } \{i, j\} \in S \text{ and } |x_i - x_j| > r \text{ for } \{i, j\} \in S^c\}.$$

That is, $B(S, r)$ is the subset of A^4 such that pairs of points that are within r of each other are exactly those whose indices are the elements of S. Furthermore define

$$C(S, r) = \{\boldsymbol{x} : \boldsymbol{x} \in A^4 \text{ and } |x_i - x_j| \leq r \;\text{ for }\; \{i, j\} \in S\},$$

the subset of A^4 for which pairs of points whose indices are in S are within distance r of each other and no explicit restrictions are placed on pairs whose indices are not in S. It will be useful to think of S as the edges of a graph and to use terminology from graph theory occasionally, see, for example, Harary (1969).

Using these definitions we have $A^4 = \bigcup_{S \subset Q} B(S; r)$ for any r. By (2.6), there exists a function $g(\beta)$ such that $g(\beta) = o(\beta^{-2/3})$ and $h(g(\beta)) = o(\beta^{-2d})$. For convenience, write $B_\beta(S) = B(S; g(\beta))$ and $C_\beta(S) = C(S; g(\beta))$. On $B_\beta(\emptyset)$, where $\emptyset$ is the null set, $\gamma_\beta(\boldsymbol{x})$ is uniformly $O(h(g(\beta)))$, so

$$\int_{B_\beta(\emptyset)} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} = o(\beta^{-2d}). \tag{4.5}$$

For any S defining a connected graph, $\nu_{4d}(B_\beta(S)) = O(g(\beta)^{-3d}) = o(\beta^{-2d})$. Since $\gamma_\beta(\boldsymbol{x})$ is uniformly bounded in $\boldsymbol{x}$ and β,

$$\int_{B_\beta(S)} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} = o(\beta^{-2d}) \tag{4.6}$$

for all S defining a connected graph.

Next, consider S defining a graph with a connected component of three points, such as $\{12,13\}$ or $\{12,13,23\}$, where we use the shorthand $ij = \{i,j\}$. For example, consider $S = \{12,13\}$; cases such as $\{12,13,23\}$ can be handled similarly. On $B_\beta(S), \gamma_B(\boldsymbol{x}) = \lambda m_{[3]}(\beta x_1, \beta x_2, \beta x_3) - \lambda^4 + O(h(g(\beta)))$ uniformly in $\boldsymbol{x}$, so

$$\begin{aligned}
&\int_{B_\beta(S)} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} \\
&= \lambda \int_{B_\beta(S)} \{m_{[3]}(\beta x_1, \beta x_2, \beta x_3) - \lambda^3\} f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} + O\Big(h(g(\beta)) g(\beta)^{2d}\Big) \\
&= -\lambda \int_{C_\beta(S)\backslash B_\beta(S)} \{m_{[3]}(\beta x_1, \beta x_2, \beta x_3) - \lambda^3\} f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} + o(\beta^{-2d}) \\
&= o(\beta^{-2d}),
\end{aligned} \tag{4.7}$$

using (2.6) and $\nu_{4d}(C_\beta(S)\backslash B_\beta(S)) = O(g(\beta)^{3d})$.

For S containing a single edge, we demonstrate, for example,

$$\begin{aligned}
&\int_{B_\beta(\{12\})} \gamma_\beta(\boldsymbol{x}) d\boldsymbol{x} \\
&= \lambda^2 \int_{B_\beta(\{12\})} \alpha(\beta(x_1 - x_2)) f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} + o(\beta^{-2d}) \\
&= -\lambda^2 \int_{C_\beta(\{12\})\backslash B_\beta(\{12\})} \alpha(\beta(x_1 - x_2)) f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} + o(\beta^{-2d}) \\
&= -\lambda^2 \int_{B_\beta(\{12,34\})} \alpha(\beta(x_1 - x_2)) f(x_1,x_2) f(x_3,x_4) d\boldsymbol{x} + o(\beta^{-2d}),
\end{aligned} \tag{4.8}$$

where the second equality follows from (2.4). To obtain the third equality, first write

$$C_\beta(\{12\})\backslash B_\beta(\{12\}) = \bigcup_{\substack{S \subset Q\backslash\{12\}, \\ S \neq \emptyset}} B_\beta(\{12\} \cup S).$$

If S contains a single edge including x_1 or x_2 as a vertex, say $S = \{13\}$, then the integral of $\gamma_\beta(\boldsymbol{x})$ over $B_\beta(\{12\} \cup S)$ will be $o(\beta^{-2d})$ as in (4.7). Similarly, all $S \subset Q\backslash\{12\}$ containing at least two edges contribute a term that is $o(\beta^{-2d})$

to the integral. Thus, only $S=\{34\}$ can contribute a term that is not $o(\beta^{-2d})$, yielding the third equality in (4.8). In addition,

$$\begin{aligned}
&\int_{B_\beta(\{12,34\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}\\
&=\lambda^2\int_{B_\beta(\{12,34\})}\{\alpha(\beta(x_1-x_2))+\alpha(\beta(x_3-x_4))\}f(x_1,x_2)f(x_3,x_4)d\boldsymbol{x}+o(\beta^{-2d})\\
&=2\lambda^2\int_{B_\beta(\{12,34\})}\alpha(\beta(x_1,x_2))f(x_1,x_2)f(x_3,x_4)d\boldsymbol{x}+o(\beta^{-2d}),
\end{aligned}$$

so that

$$\int_{B_\beta(\{12,34\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}+\int_{B_\beta(\{12\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}+\int_{B_\beta(\{34\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}=o(\beta^{-2d}).$$

Similarly,

$$\int_{B_\beta(\{13,24\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}+\int_{B_\beta(\{13\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}+\int_{B_\beta(\{24\})}\gamma_\beta(\boldsymbol{x})d\boldsymbol{x}=o(\beta^{-2d}).$$

Combining these results with (4.5)-(4.7) yields (4.3) and hence (4.1).

Finally, using simpler versions of the argument leading to (4.1) yields

$$\int_{A^2}f(x_1,x_2)^2m_{[2]}(\beta x_1,\beta x_2)dx_1dx_2\longrightarrow\lambda^2\int_{A^2}f(x,y)^2dxdy \tag{4.9}$$

and

$$\begin{aligned}
&\beta^d\int_{A^3}f(x_1,x_2)f(x_1,x_3)m_{[3]}(\beta x_1,\beta x_2,\beta x_3)dx_1dx_2dx_3\\
&\longrightarrow\lambda\int_{\mathbb{R}^d}\alpha(y)dy\int_{A^2}f(x,y)^2dxdy,
\end{aligned} \tag{4.10}$$

Theorem 2 follows from (2.5), (4.1), (4.9) and (4.10).

Acknowledgement

This research was supported in part by National Science Foundation Grant no. DMS 92-04504. This manuscript was prepared using computer facilities supported in part by The National Science Foundation Grants DMS 89-05292, DMS 87-03942, and DMS 86-01732 awarded to the Department of Statistics at The University of Chicago, and by The University of Chicago Block Fund.

References

Baddeley, A. J. and Gill, R. D. (1993). Kaplan-Meier estimators for interpoint distance distributions for spatial point processes. Report BS-R9315, Centrum voor Wiskunde en Informatica, Amsterdam.

Bertram, U., Wendrock, H. and Stoyan, D. (1993). Estimation variances for estimators of product densities and pair correlation functions of planar point processes. Ann. Inst. Statist. Math. **45**, 211-221.

Daley, D. J. and Vere-Jones, D. (1988). An Introduction to the Theory of Point Processes. Springer-Verlag, New York.

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

Harary, F. (1969). Graph Theory. Addison-Wesley, Reading, MA.

Heinrich, L. (1988). Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes. Statistics **19**, 87-106.

Jolivet, E. (1981). Central limit theorem and convergence of empirical processes for stationary point processes. In Point Processes and Queuing Problems (Edited by P. Bártfai and J. Tomkó), 117-161. North-Holland, Amsterdam.

Karr, A. F. (1987). Estimation of palm measures of stationary point processes. Probab. Theory Related Fields **74**, 55-69.

Mase, S. (1982). Properties of fourth-order strong mixing rates and its application to random set theory. J. Multivariate Anal. **12**, 549-561.

Miles, R. E. (1974). On the elimination of edge effects in planar sampling. In Stochastic Geometry (Edited by E. F. Harding and D. G. Kendall), 228-247. John Wiley, Chichester.

Ohser, J. and Stoyan, D. (1981). On the second-order and orientation analysis of planar stationary point processes. Biometrical J. **23**, 523-533.

Ripley, B. D. (1976). The second-order analysis of stationary point processes. J. Appl. Probab. **13**, 255-266.

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

Serfling, R. J. (1980). Approximation Theorems of Mathematical Statistics. John Wiley, New York.

Silverman, B. W. (1978). Distances on circles, toruses and spheres. *J▷Appl▷Probab▷* **15**, 136-143.

Stein, M. L. (1991). A new class of estimators for the reduced second moment measure of point processes. Biometrika **78**, 281-286.

Stein, M. L. (1993). Asymptotically optimal estimation for the reduced second moment measure of point processes. Biometrika **80**, 443-449.

Department of Statistics, The University of Chicago, Chicago, IL 60637, U.S.A.

(Received August 1993; accepted July 1994)