Open access • Journal Article • DOI:10.1111/J.1467-9574.1996.TB01501.X

A nonparametric measure of spatial interaction in point patterns — Source link

van Marie-Colette Lieshout, Adrian Baddeley

Institutions: University of Warwick, University of Western Australia

Published on: 01 Nov 1996 - Statistica Neerlandica (Wiley-Blackwell)

Topics: Nearest neighbour distribution, Complete spatial randomness, Point process, Function (mathematics) and Estimator

A nonparametric measure of spatial interaction in point patterns

A J Baddeley[1,2,4]

M N M van Lieshout[3,4]

[1] *Department of Mathematics, University of Western Australia, Nedlands WA 6009, Australia*

[2] *Department of Mathematics and Computer Science, University of Leiden, Netherlands*

[3] *Department of Mathematics and Computer Science, Free University De Boelelaan 1081 a, 1081 HV Amsterdam, Netherlands*

[4] *CWI, P.O. Box 94079, 1090 GB Amsterdam, Netherlands*

In memory of Philip Holgate

Abstract

The strength and range of interpoint interactions in a spatial point process can be quantified by the function $J = (1 - G)/(1 - F)$, where G is the nearest–neighbour distance distribution function and F the empty space function of the process. $J(r)$ is identically equal to 1 for a Poisson process; values of $J(r)$ smaller or larger than 1 indicate clustering or regularity, respectively. We show that, for a large class of point processes, $J(r)$ is *constant* for distances r greater than the range of spatial interaction. Hence both the range and type of interaction can be inferred from J without parametric model assumptions. It is also possible to evaluate $J(r)$ explicitly for many point process models, so that J is also useful for parameter estimation. Various properties are derived, including the fact that the J function of the superposition of independent point processes is a weighted mean of the J functions of the individual processes. Estimators of J can be constructed from standard estimators of F and G. We compute estimates of J for several standard point pattern datasets and conclude that it is a useful indicator of spatial interaction.

1. Introduction

The statistical analysis of a point pattern usually begins with the computation of estimates of the summary functions F (empty space function), G (nearest–neighbour distance distribution function) and K (reduced second moment function), defined e.g. in [11, 14, 34, 35]. While these are useful descriptions of spatial pattern, and can easily be estimated from data, there are very few stochastic models for which F, G or K is known analytically, so that parameter estimation and inference based on F, G, K are difficult.

Recall that, for a stationary point process, F is the distribution function of the distance from an arbitrary fixed point to the nearest random point of the process, and G of the distance from a point *of the process* to the nearest other point of the process. This paper advocates the use of

$$J(r) = \frac{1 - G(r)}{1 - F(r)}.$$

This is a nonparametric measure of the type of spatial interaction: the value 1 can be interpreted as indicating complete randomness or lack of interaction, while values less than 1 imply 'clustered' pattern and values greater than 1 imply 'ordered' or 'inhibitory' pattern.

We show that, for a very large class of point processes, the function J is *constant* for values of r larger than the effective range of spatial interaction. Hence J can be used to infer both the range and type of spatial interaction. Furthermore we are able to evaluate J explicitly for several stochastic models, so that it could be used directly for parameter estimation.

An appealing interpretation of J is that it compares the environment of a typical random point of the process with the environment of a fixed arbitrary point. $J(r)$ is the ratio of the probabilities, under these two situations, of the event that there are no points within a distance r of the given point. In terms of survival analysis, J is the ratio of the survival functions of the distance-to-nearest-point under these two situations; and our main result states that the hazard measures [25] of F and G are *equal* beyond the effective range of interaction r.

Special cases of these results are implicit in the literature. The forms of F and G for a Neyman-Scott cluster process were derived by Bartlett [9]; see [1, 31, 32], [10, pp. 8–9], and for detailed derivations [12, §8.3, p. 243 ff.], [41, p. 143]. For a general Poisson cluster process (Poisson parent points, i.i.d. offspring)

$$1 - G(r) = (1 - F(r))\, E(r)$$

where $E(r)$ is the probability that a randomly-chosen point in a typical cluster is more than r units distant from any other point belonging to the same cluster. Hence in particular if all offspring lie within a radius t of the parent point, we have $J(r) = 1$ for all $r > 2t$. Again, for a stationary, pairwise-interaction Gibbs process, Stoyan et al. [41, p. 159] exhibit a relationship between $1 - F(r)$ and $1 - G(r)$ when r is exactly equal to the interaction distance R. In this paper we extend the relationship to all $r \geq R$.

Statistical inference based on comparisons between F and G has occasionally been suggested. Diggle [13, (5.7)] proposed the statistic $D = \sup_r \mid \widehat{F}(r) - \widehat{G}(r) \mid$ as a measure of deviation from the Poisson process.

This paper is organised as follows. In Section 2 we review the main techniques from spatial statistics that are used in the sequel. Section 3 introduces the J-function; the main theorem states that $J(r)$ is constant beyond the effective range of interaction. We also examine the behaviour of J under the basic operations of superposition and thinning and show that the J-function of a superposition of independent processes is a convex combination of the J-functions of the components. The relationship between the J-function of a thinned process and that of the original process appears to be rather complex; in particular, in contrast to Ripley's K-function, the J-function is not invariant under thinning.

In Section 4 we show that the J-function can be computed explicitly for a large class of point process models, including Poisson processes, Markov point porcesses, Neyman-Scott and Cox processes. For these examples at least, the classification of patterns as 'clustered' or 'regular' on the basisof their J-function values agrees with similar classifications based on F, G and K.

In Section 5 we discuss briefly how the J-statistic can be used for parameter estimation, while Section 6 is a simple illustration on three standard data sets [14], representing regular, random and clustered patterns.

2. Background

Throughout this paper we consider a stationary point process X in $\mathbb{R}^k$, regarded as a random set of points. For details of the theory of point processes see [12] or [11, 41].

Define the empty space function F of X to be the distribution function

$$F(r) = \mathbb{P}\{\rho(y, X) \leq r\}$$

of

$$\rho(y, X) = \min\{||y - x|| : x \in X\},$$

the distance from an arbitrary fixed point $y \in \mathbb{R}^k$ to the nearest point of the process. By stationarity, the definition of F does not depend on y.

Write $B(y, r) = \{x \in \mathbb{R}^k : \rho(x, y) \leq r\}$ for the closed ball of radius $r > 0$ centred at y in $\mathbb{R}^k$. Then $1 - F(r)$ is the probability that X puts no points in $B(y, r)$:

$$1 - F(r) = \mathbb{P}\{X \cap B(y, r) = \emptyset\}.$$

For example, for a Poisson process of intensity λ in $\mathbb{R}^2$ we obtain $F(r) = 1 - \exp\{-\lambda\pi r^2\}$. F has been variously dubbed the 'empty space, 'point-event distance' and 'spherical contact' distribution function.

To define G we need the *Palm distribution* $\mathbb{P}^y$ of X at $y \in \mathbb{R}^k$, which can be regarded as the conditional distribution of the entire process given that there is a point of X at y [12, chap. 12], [11, pp. 630-631], [41, p. 110 ff.]. Then define

$$G(r) = \mathbb{P}^y\{\rho(y, X \setminus \{y\}) \leq r\};$$

again this does not depend on y, by stationarity. Thus G is the distribution function of the distance from a point of the process to the nearest other point, and is known variously as the 'nearest–neighbour' or 'event–event' distribution function.

It is convenient to use the *reduced Palm distribution* $\mathbb{P}^!_y$ defined as the distribution of $X \setminus \{y\}$ under $\mathbb{P}^y$, i.e. the conditional distribution of the *rest* of the process given that there is a point at y. Then the definition of G reads

$$G(r) = \mathbb{P}^!_y\{\rho(y, X) \leq r\}$$

in harmony with the definition of F. For example, for a stationary Poisson process of intensity λ, the reduced Palm distribution $\mathbb{P}^!_y$ is identical to $\mathbb{P}$, and $G \equiv F$.

Our main tool will be the *Takacs-Fiksel formula* which relates the reduced Palm distribution of X to its (ordinary) distribution:

$$\lambda \, \mathbb{E}_y^! f(X) = \mathbb{E}[\lambda(y; X) \, f(X)] \tag{2.1}$$

holding (under suitable conditions on X) for any bounded nonnegative measurable function on the space of realizations of X [22, 24, 42, 43] (see also [37], [29, 30], [35, p. 54–55], [15, §2.4]). Here λ is the intensity of X and $\lambda(y; X)$ is the *Papangelou conditional intensity* of X at y. In other words, (2.1) states that $\mathbb{P}_y^!$ is equivalent to the $\lambda(y; X)$-weighted distribution of X. In particular

$$\lambda = \mathbb{E}\lambda(0; X). \tag{2.2}$$

A necessary and sufficient condition (in the stationary case) for validity of (2.1) is that $\mathbb{P}_y^!$ be absolutely continuous with respect to $\mathbb{P}$, whereupon $\lambda(y; X)$ is uniquely defined by (2.1). The Takacs-Fiksel formula holds in particular for all stationary Gibbs point processes [33, 35] and for Poisson cluster processes when the cluster distribution is absolutely continuous. The corresponding expressions for $\lambda(y; X)$ are given in Section 4. Examples of processes which fail to satisfy (2.1) are randomly translated grids, and cluster processes consisting of pairs of points separated by a fixed distance.

Kallenberg [29, 30] gives a detailed explanation of the duality between the Palm distribution and Papangelou conditional intensity. The reduced Palm distribution is concerned with the remainder of the pattern given that a point falls at a particular location ('internal conditioning'), while the conditional intensity describes the behaviour of the process at a single point in space given the realisation everywhere else ('external conditioning').

3. The J-function

Definition 1 *For a stationary point process X define*

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \tag{3.3}$$

for all $r \geq 0$ such that $F(r) < 1$.

For example, if X is a Poisson process then $F \equiv G$, so we obtain $J(r) \equiv 1$. Note that, even in a completely nonparametric context, the function J has an interpretation as the ratio of the survival functions of the distance to the nearest (other) point of X from (a) a point of the process, (b) a fixed arbitrary point. Values $J(r) < 1$ indicate that the survival function in situation (a) is smaller than that for (b), which may be interpreted as indicating 'clustered' pattern; values $J(r) > 1$ indicate 'ordered' pattern. In the examples in Section 4 we will reconcile this with other definitions of 'clustering' and 'ordering'.

Note that $J(0) = 1$ always. The denominator $1 - F$ is always absolutely continuous [3] but the numerator $1 - G$ need not be, so the discontinuity points of J are those of G. In general $1 - G(r)$ might be nonzero when $1 - F(r)$ is zero (e.g. for a randomly-translated unit square grid when $r = 1/\sqrt{2}$) but this does not occur for point processes of real interest.

Theorem 1 *Let X be a stationary point process with intensity λ whose Papangelou conditional intensity $\lambda(y; X)$ exists. Then $G(r) < 1$ implies $F(r) < 1$ and*

$$J(r) = \left(\mathbb{E}_0^! \left[\frac{\lambda}{\lambda(0;X)} \mid X \cap B(0,r) = \emptyset \right] \right)^{-1} \tag{3.4}$$

In particular, suppose X has 'interactions of finite range R' in the sense that $\lambda(0;X)$ is constant (and thus equal to $\lambda(0;\emptyset)$) for all point patterns X which contain no points in $B(0,R)$. Then

$$J(r) = \frac{\lambda(0;\emptyset)}{\lambda} \qquad \text{for } r \geq R. \tag{3.5}$$

Proof: Let A be the event $\{X \cap B(0,r) = \emptyset\}$, so that $1-F(r) = \mathbb{P}(A)$ and $1-G(r) = \mathbb{P}_0^!(A)$. Apply the Takacs-Fiksel formula (2.1) to

$$f(X) = \frac{\mathbb{1}_A}{\lambda(0;X)}, \quad r > 0$$

(cf. [41, (5.5.18), p. 159]). The right hand side of (2.1) is $\mathbb{E}[\lambda(0;X)\, f(X)] = 1 - F(r)$ giving

$$1 - F(r) = \lambda \mathbb{E}_0^! \left[\frac{\mathbb{1}_A}{\lambda(0;X)} \right]$$

Dividing this by $1 - G(r) = \mathbb{P}_0^!(A)$ gives the reciprocal of (3.4).

In the second case, if $\lambda(0;X) \equiv \lambda(0;\emptyset)$ on A then

$$f(X) \equiv \frac{\mathbb{1}_A}{\lambda(0;\emptyset)}$$

so that the left side of (2.1) is

$$\lambda \frac{1}{\lambda(0;\emptyset)} \mathbb{P}_0^!(A) = \frac{\lambda}{\lambda(0;\emptyset)}(1 - G(r))$$

yielding (3.5). □

Next we examine the behaviour of J under the basic point process operations of superposition and thinning.

Theorem 2 *Let X_1, X_2 be independent, stationary point processes with intensities λ_1, λ_2 and J-functions J_1, J_2 respectively. Then the J-function of the superposition $X = X_1 \cup X_2$ is a convex combination of the J-functions of the components:*

$$J(r) = \frac{\lambda_1}{\lambda_1 + \lambda_2} J_1(r) + \frac{\lambda_2}{\lambda_1 + \lambda_2} J_2(r). \tag{3.6}$$

Proof: By independence

$$1 - F(t) = (1 - F_1(t))(1 - F_2(t)).$$

Writing $P_{(i)}^0$ for the Palm distribution (on the entire probability space) with respect to X_i, $i = 1, 2$ and P^0 for the Palm distribution with respect to X, we have (e.g. [41, p. 116])

$$P^0 = \frac{\lambda_1}{\lambda_1+\lambda_2} P^0_{(1)} + \frac{\lambda_2}{\lambda_1+\lambda_2} P^0_{(2)}.$$

The joint distribution of X_1 and X_2 under $P^0_{(1)}$ is independent, with X_1 governed by its Palm distribution (the Palm distribution of its marginal distribution) and X_2 by its (ordinary) marginal distribution. Similarly for $P^0_{(2)}$. Hence

$$1 - G(t) = \frac{\lambda_1}{\lambda_1+\lambda_2}(1-G_1(t))(1-F_2(t)) + \frac{\lambda_2}{\lambda_1+\lambda_2}(1-F_1(t))(1-G_2(t)).$$

Dividing this by the identity for F gives (3.6). □

For comparison, the K-function of the superposition in the same situation is

$$\begin{aligned} K(t) &= \frac{\lambda_1}{\lambda_1+\lambda_2}\left[\frac{1}{\lambda_1+\lambda_2}(\lambda_2\pi t^2 + \lambda_1 K_1(t))\right] + \frac{\lambda_2}{\lambda_1+\lambda_2}\left[\frac{1}{\lambda_1+\lambda_2}(\lambda_1\pi t^2 + \lambda_2 K_2(t))\right] \\ &= \frac{1}{(\lambda_1+\lambda_2)^2}\left[2\lambda_1\lambda_2\pi t^2 + \lambda_1^2 K_1(t) + \lambda_2^2 K_2(t)\right]. \end{aligned}$$

Theorem 3 *Let X_p be the process obtained from a stationary point process X by randomly deleting or retaining each point independently of other points, with retention probability $p > 0$. Then the J-function of X_p is*

$$J_p(r) = \frac{Q^0_r(1-p)}{Q_r(1-p)} \tag{3.7}$$

where Q^0_r, Q_r are the generating functions of $n(X \cap B(0,r))$ under $\mathbb{P}^!_0$ and $\mathbb{P}$ respectively. [The J-function of X itself is the case $p = 1$.]

Proof : Let F_p, G_p be the F and G functions for X_p. Clearly $1 - F_p(r) = Q_r(1-p)$. To prove $1 - G_p(r) = Q^0_r(1-p)$ use the fact that the Palm distribution of X_p coincides with the effect of random p-thinning on the Palm distribution of X. □

Thus while the K-function is invariant under random thinning [14, p. 67], [41, p. 134], in general the J-function is not. There does not appear to be a simple general relationship between J_p and J.

4. Examples

4.1 Poisson process

For a stationary Poisson process of intensity λ we have $F \equiv G$ so that $J \equiv 1$. We could also derive this from Theorem 1 by observing that $\lambda(0; X) = \lambda$ for arbitrary X.

4.2 Pairwise-interaction Markov point process

For a pairwise interaction point process [36], [41, section 5.5] with activity constant β and interaction $\gamma(u,v)$ between points $u, v \in \mathbb{R}^k$,

$$\lambda(y; X) = \beta \prod_{x \in X} \gamma(x,y). \tag{4.8}$$

The process is *Markov* (in the Ripley-Kelly sense [36]) with interaction range R, if $\gamma(u,v) = 1$ when $||u - v|| \geq R$. Examples include the *hard core process* defined by

$$\gamma(u,v) = \begin{cases} 0 & \text{if } ||u-v|| < R \\ 1 & \text{otherwise} \end{cases} \tag{4.9}$$

and the *Strauss process* defined by replacing 0 in (4.9) by a constant $0 < \gamma < 1$.

Theorem 4 *For a Markov pairwise-interaction process with interaction range R,*

(a) $J(r)$ is defined for all r;

(b) $$J(r) = \frac{\beta}{\lambda} \quad \text{for } r \geq R; \tag{4.10}$$

(c) for 'purely inhibitive' interactions ($\gamma(u,v) \leq 1$ for all u,v) we have $J(r) \geq 1$ for all r;

(d) for the hard core process $J(r) = 1/(1-F(r))$ for $r < R$, and in particular J is continuous and monotone increasing for $r < R$. Furthermore $J(r) = 1/(1 - \lambda m(B(0,r)))$ for $r < R/2$.

Thus, the hard-core and Strauss processes yield values (for r outside the interaction range) indicating 'ordered' pattern in the sense defined below Definition 1. Equation (4.10) was implicitly computed in [41, (5.5.18), p. 159] for the value $r = R$ only.

Proof : To prove this we note that the product in (4.8) depends only on points $x \in X$ with $||x - y|| \leq R$, so $\lambda(y; X)$ depends only on $X \cap B(y, R)$. Hence X has finite range interaction in the sense of Theorem 1(b) with $\lambda(0; \emptyset) = \beta$, and we get (4.10).

Note that λ, the intensity of X, is determined by the parameters β and $\gamma(\cdot,\cdot)$ in a generally complex way. However for a purely inhibitive process we have $\lambda(0; X) \leq \beta$ a.s. so that $\lambda \leq \beta$ using (2.2). This gives $J(r) \geq 1$ for $r \geq R$.

For values $r < R$ it is again a complex task to compute $J(r)$, except that for purely inhibitive γ we can again show that $J(r) \geq 1$ for all r. For a hard core process, clearly $G(r) = 0$ for $r < R$, so $J(r) = 1/(1 - F(r))$ for $r < R$. In particular J is monotone nonincreasing. Furthermore since spheres of radius $r < R/2$ centred at the points of a hard core process do not overlap, we have $F(r) = \lambda m(B(0,r))$ for $r < R/2$, and hence $J(r) = 1/(1 - \lambda m(B(0,r)))$, for $r < R/2$. □

4.3 Markov point processes (general)

Many of the arguments in the preceding paragraph carry over to Markov point processes in general. A stationary process X is Markov with finite interaction range R if its conditional intensity $\lambda(0; X)$ at 0 exists and depends only on $X \cap B(0, R)$. It follows that for any X satisfying $X \cap B(0, R) = \emptyset$

$$\lambda(0; X) = \lambda(0; \emptyset)$$

so that Theorem 1(b) applies and $J(r)$ is constant for $r > R$.

An example of interest is the *area-interaction process* [4] for which

$$\lambda(0;X) = \beta\gamma^{-m(B(0,t)\setminus U(X))} \tag{4.11}$$

where m is Lebesgue measure and $U(X) = \bigcup_{x\in X} B(x,t)$ for a fixed $t > 0$. The process is defined for all finite γ, with $\gamma < 1$ generating 'ordered' patterns and $\gamma > 1$ 'clustered' patterns. For any X such that $X \cap B(0,2t) = \emptyset$ we have $U(X) \cap B(0,t) = \emptyset$ so that $\lambda(0,X) = \beta\gamma^{-m(B(0,t))} = \beta\eta$, say. Thus Theorem 1(b) applies with $R = 2t$, and

$$J(r) = \frac{\beta\eta}{\lambda} \qquad \text{for } r > 2t. \tag{4.12}$$

Since $0 \le m(B(0,t) \setminus U(X)) \le m(B(0,t))$ we have for $\gamma < 1$ that $\lambda(0;X) \le \beta\eta$ for all X so that $\lambda \le \beta\eta$ and hence $J(r) \ge 1$ for all r, i.e. this is also 'ordered' in terms of J. Similarly, for $\gamma > 1$ we have $\lambda(0;X) \ge \beta\eta$ a.s. so that $\lambda \ge \beta\eta$ and $J(r) \le 1$ for all r, i.e. this is 'clustered' in terms of J.

4.4 Poisson cluster processes

A stationary Poisson cluster process is constructed by generating a stationary Poisson process Y of 'parent points'; generating i.i.d. finite point processes ('clusters') Z_y for each $y \in Y$; and forming the superposition $X = \bigcup_{y\in Y}(y + Z_y)$ of the translated clusters. *Neyman-Scott* processes are the special case where the typical cluster Z consists of a random number N of i.i.d. points. The *Matérn* cluster process is the further special case of Neyman-Scott processes where N is a Poisson variable and the common distribution of the cluster points is uniform over the ball of radius t centred on the parent point.

Stoyan et al. [41, p. 142 ff.] (and Bartlett [10, p. 8–9]) show that for any stationary Poisson cluster process

$$1 - G(r) = [1 - F(r)]\, C_0\{Z \cap B(0,r) = \{0\}\}, \qquad r \ge 0$$

where C_0 is the Palm distribution of the typical cluster Z. This follows from a fundamental identity for the Palm distribution of a Poisson cluster process [41, (5.3.2), p. 142].

Since Z is a finite point process, C_0 can be interpreted as the $n(Z)$-weighted distribution of $Z - z$ where, given Z, z is one of the points of Z chosen with equal probability. Hence we may interpret $C_0\{Z \cap B(0,r) = \{0\}\}$ as the defective distribution function of the distance from a typical point of Z to the nearest other point of Z, if any [41, p. 143]. Hence we have the following result.

Theorem 5 *For any stationary Poisson cluster process, $J(r)$ is defined for all $r \ge 0$;*

$$J(r) = C_0\{Z \cap B(0,r) = \{0\}\}$$

is a monotone nonincreasing function, with values $J(r) \le 1$, determined only by the distribution of the clusters.

If the typical cluster Z is a.s. contained within the ball of radius t around the parent point, then $J(r)$ is constant for $r > 2t$ where it is equal to $\mathbb{P}\{n(Z) = 1\}\,/\mathbb{E}n(Z)$.

Hence all stationary Poisson cluster processes are 'clustered' with respect to J as defined below Definition 1.

For example, for the Matérn cluster process in $\mathbb{R}^2$ with $\mathbf{E}n(Z) = \mu$ points per cluster we find

$$\begin{aligned} J(r) &= \frac{1}{m(B(0,t))} \int_{B(0,t)} e^{-\mu V(x,r,t)} dx \\ &= e^{-\mu} \qquad \text{for } r > 2t \end{aligned}$$

where $V(x,r,t) = m(B(x,r) \cap B(0,t))/m(B(0,t))$.

Note that Theorem 5 is proved using the cluster formula [41, (5.3.2), p. 142] rather than Theorem 1, and holds even in cases when the Papangelou conditional intensity does not exist. However, if the Palm distribution of the typical cluster Z is absolutely continuous with respect to the distribution of Z, then Theorem 1 applies and yields the conclusions of Theorem 5.

This result is perhaps less surprising in view of the recent proof [5] that Poisson cluster processes with bounded clusters are nearest-neighbour Markov processes in the sense of [6].

4.5 Cox processes

Cox point processes are constructed by generating a random measure Λ and, conditional upon Λ, generating an inhomogeneous Poisson point process X with intensity measure Λ.

Theorem 6 *Let X be a Cox point process with driving random measure Λ which is stationary and a.s. nonatomic. Then the J-function of X is defined for all $r \geq 0$ and equals*

$$J(r) = \frac{\mathbf{E}^0 e^{-\Lambda(B(0,r))}}{\mathbf{E} e^{-\Lambda(B(0,r))}}$$

where $\mathbf{E}^0$ denotes expectation with respect to the Palm distribution of Λ.

This follows from the fact that the reduced Palm distribution of X is the distribution of a Cox process with driving measure distributed as the Palm distribution of Λ, cf. [41, p. 141].

For example, consider a mixed Poisson process, where Λ is a random constant multiple of Lebesgue measure, $\Lambda = \alpha\, m(\cdot)$ where α is any nonnegative random variable not identically equal to zero. Then the Palm distribution of Λ is simply the α-weighted distribution, and

$$J(r) = \frac{\mathbf{E}\left[\alpha e^{-\alpha\pi r^2}\right]}{\mathbf{E}\alpha\ \mathbf{E} e^{-\alpha\pi r^2}}.$$

5. Statistical aspects

5.1 Nonparametric estimation of J

Edge-corrected estimators for F and G based on observations of X within a bounded window $W \subset \mathbb{R}^k$ are reviewed in [35, chap. 3], [41, pp. 122–131], [11, chap. 8]. For recent variations see [3, 7, 8, 16, 17, 18, 19, 20, 21, 23, 39].

We propose estimating J by plugging into (3.3) estimates of F and G obtained by methods that are comparable to one another. For example one may estimate F by the standard 'border correction' estimator [35, chap. 3] and G by Hanisch's border correction estimator $\widehat{G}_4$ [26] (see [41, p. 128] where G is called D). These are Horvitz-Thompson type ratio estimators with comparable denominators, and are pointwise unbiased for F and pointwise approximately

unbiased for G, respectively. Alternatively the Kaplan-Meier style estimators of F and G proposed by Baddeley & Gill [3] could be used. These have the advantage of being proper distribution functions (possibly defective), and correspond to unbiased and approximately unbiased estimators of the hazard measures of F and G, respectively. Furthermore the estimator of F in [3] has the same continuity properties as F itself.

We know little about the sampling properties of either estimator of J. Clearly $\widehat{J}(0) = 1$ always. It seems plausible that the relative error of J will increase with r, since this is true of standard estimators of $F(r)$ and $G(r)$ [3, 18, 20, 21]. Central limit theorems have been proved for $\widehat{F}$ and $\widehat{G}$ of both the Horvitz-Thompson and Kaplan-Meier types under various regimes [2, 3, 27, 28, 40], [11, p. 480]; a joint CLT for $(\widehat{F}, \widehat{G})$, and hence for $\widehat{J}$, seems plausible but has not been established to the authors' knowledge.

Edge effects have a far greater influence on $\widehat{G}$ than on $\widehat{F}$ [35, chap. 3],[3]. The sampling properties of $\widehat{G}$ and therefore of $\widehat{J}$ may be particular cause for concern when the sampling window W is irregular, or in dimensions higher than two [3, 7].

5.2 Estimation and inference based on J

In section 4 we were able to calculate the J function (at least for $r > R$) for a number of parametric or semi-parametric stochastic models. One could use these results to estimate the parameters of a chosen model from values of $\widehat{J}$.

It should be stressed that this approach is merely a special case of the Takacs-Fiksel estimation method [22, 24, 42, 43], [35, p. 54–55], [15, §2.4], [37, 38] since the basic equations (3.4)–(3.5) are special cases of (2.1) with the choice of f given in the proof of Theorem 1.

For a Markov pairwise-interaction process, (4.10) gives the constant value of $J(r)$ for all $r > R$ in terms of the parameter β and the intensity λ. The intensity is determined by β and by the interaction function $\gamma(\cdot, \cdot)$ in a complex way. However λ may be estimated directly from the data, as $\widehat{\lambda} = n(X \cap W)/m(W)$ in the usual way. If R is assumed known then β can be estimated via (4.10). This is semi-parametric estimation, since γ is unknown apart from the constraint that $\gamma(u, v) = 1$ for $||u - v|| > R$.

Similarly, for an area-interaction process, (4.12) allows us to estimate the parameters β and $\eta = \gamma^{m(B(0,t))}$ given the interaction radius $R = 2t$.

Estimation of the interaction distance R, in any of the models studied, amounts to estimating the largest interval $[R, \infty)$ on which J is constant. At present we have only the *ad hoc* suggestion of taking

$$\widehat{R} = \inf \left\{ R : \sup_{r \geq R} \widehat{J}(r) - \inf_{r \geq R} \widehat{J}(r) < \epsilon \right\}$$

where ϵ is of order $n(X \cap W)^{-1/2}$.

6. Examples

We have taken three standard point pattern datasets discussed at length by Diggle [14], entitled `pines` ('Japanese pine saplings'), `redwood` ('Californian redwood seedlings') and `cells` ('biological cells'). These were exhibited as typical examples of random, clustered, and ordered patterns respectively.

Figures 1–3 show the data and corresponding estimates $\widehat{J}$ obtained using the Kaplan-Meier estimators of F and G [3]. For `pines` the value of J is close to 1 for almost the entire range of r

values expect at high r values; for `redwood` it is below 1 and monotonically decreasing except for small fluctuations; and for `cells` it is above 1 for the entire range and is monotonically increasing. These results are consistent with our expectations.

We may conclude provisionally that the J-statistic is a useful indicator of the type of spatial pattern. Further numerical experiments will be described elsewhere.

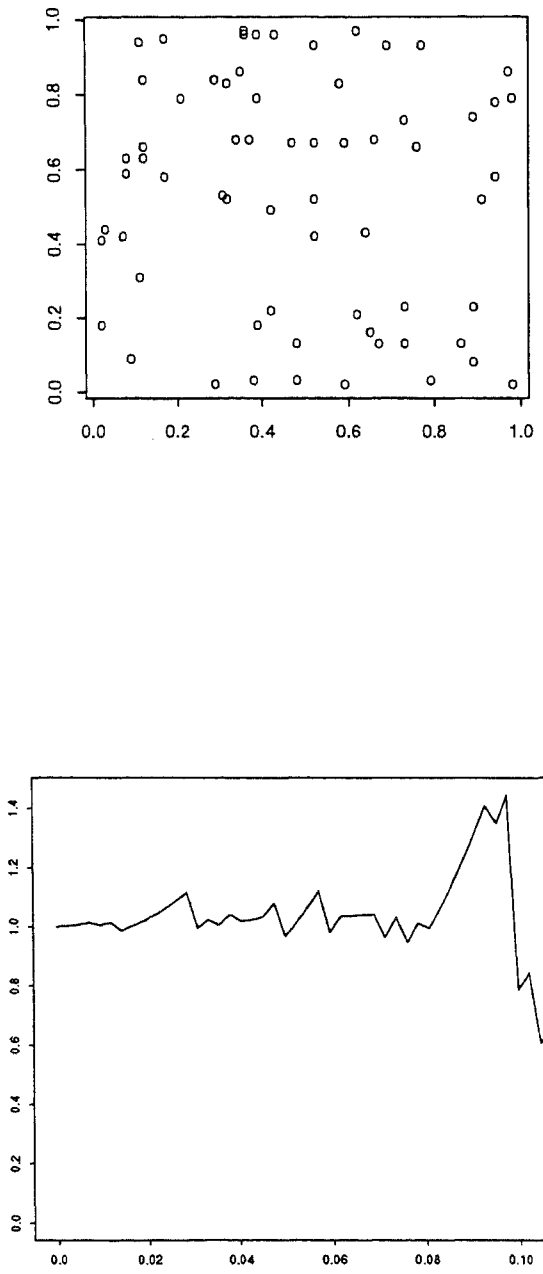
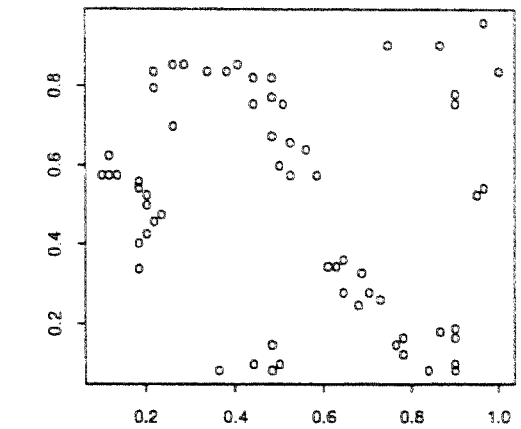


Figure 1: **pines** data (above) and estimate of J (below).



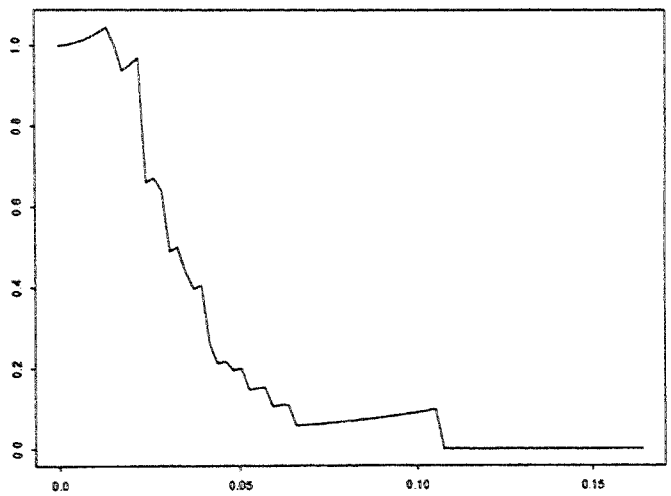
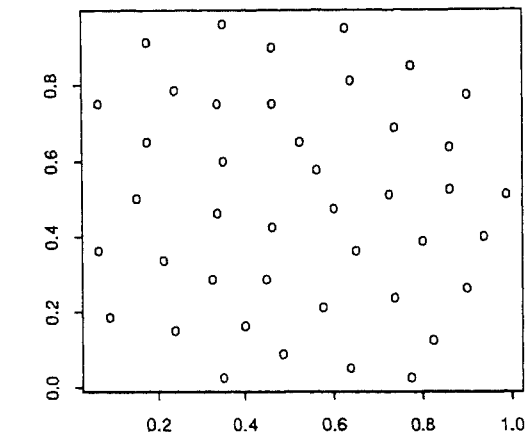


Figure 2: **redwood** data (above) and estimate of J (below).



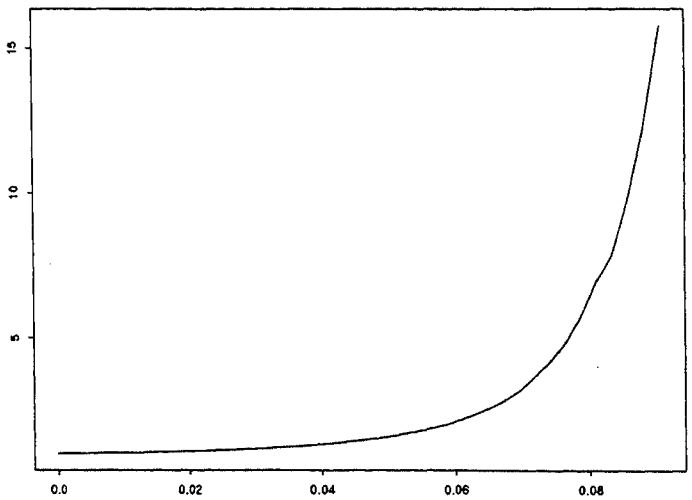


Figure 3: `cells` data (above) and estimate of J (below).

REFERENCES

1. R.V. Ambartzumian. Palm distributions and superpositions of independent point processes in R^n. In P.A.W. Lewis, editor, *Stochastic point processes: statistical analyis, theory, and applications*, pages 626–645. John Wiley and Sons, New York, 1972.
2. A.J. Baddeley. A limit theorem for statistics of spatial data. *Advances in Applied Probability*, 12:447–461, 1980.
3. A.J. Baddeley and R.D. Gill. Kaplan-Meier estimators for interpoint distance distributions of spatial point processes. Research Report BS-R9315, Centrum voor Wiskunde en Informatica, July 1993.
4. A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. Research report BS-R9318, Centrum voor Wiskunde en Informatica, November 1993.
5. A.J. Baddeley, M.N.M. van Lieshout and J. Møller. Markov properties of cluster processes. Manuscript in preparation.
6. A.J. Baddeley and J. Møller. Nearest-neighbour Markov point processes and random sets. *International Statistical Review*, 57:89–121, 1989.
7. A.J. Baddeley, R.A. Moyeed, C.V. Howard and A. Boyde. Analysis of a three-dimensional point pattern with replication. *Applied Statistics*, 42(4):641–668, 1993.
8. L.G. Barendregt and M.J. Rottschäfer. A statistical analysis of spatial point patterns. A case study. *Statistica Neerlandica*, 45:345–363, 1991.
9. M.S. Bartlett. The spectral analysis of two-dimensional point processes. *Biometrika*, 51:299–311, 1964.
10. M.S. Bartlett. *The statistical analysis of spatial pattern.* Chapman and Hall, London, 1975.
11. N.A.C. Cressie. *Statistics for spatial data.* John Wiley and Sons, New York, 1991.
12. D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes.* Springer Verlag, New York, 1988.
13. P.J. Diggle. On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrika*, 35:87–101, 1979.
14. P.J. Diggle. *Statistical analysis of spatial point patterns.* Academic Press, London, 1983.
15. P.J. Diggle, T. Fiksel, Y. Ogata, D. Stoyan and M. Tanemura. On parameter estimation for pairwise interaction processes. *International Statistical Review*, 1993. to appear.
16. S.I. Doguwa. A comparative study of the edge-corrected kernel-based nearest neighbour density estimators for point processes. *Journal of Statistical Computation and Simulation*, 33:83–100, 1989.
17. S.I. Doguwa. On edge-corrected kernel-based pair correlation function estimators for point processes. *Biometrical Journal*, 32:95–106, 1990.
18. S.I. Doguwa. On the estimation of the point-object nearest neighbour distribution $F(y)$ for point processes. *Journal of Statistical Computation and Simulation*, 41:95–107, 1992.
19. S.I. Doguwa and D.N. Choji. On edge-corrected probability density function estimators

for point processes. *Biometrical Journal*, 33:623–637, 1991.

20. S.I. Doguwa and G.J.G. Upton. Edge-corrected estimators for the reduced second moment measure of point processes. *Biometrical Journal*, 31:563–675, 1989.
21. S.I. Doguwa and G.J.G. Upton. On the estimation of the nearest neighbour distribution, $G(t)$, for point processes. *Biometrical Journal*, 32:863–876, 1990.
22. T. Fiksel. Estimation of parametrized pair potentials of marked and non-marked Gibbsian point processes. *Elektronische Informationsverarbeitung und Kybernetika*, 20:270–278, 1984.
23. T. Fiksel. Edge-corrected density estimators for point processes. *Statistics*, 19:67–75, 1988.
24. T. Fiksel. Estimation of interaction potentials of Gibbsian point processes. *Statistics*, 19:77–86, 1988.
25. R.D. Gill. Lectures on survival analysis. In P. Bernard, editor, *22e Ecole d'Eté de Probabilités de Saint-Flour 1992*, Lecture Notes in Statistics. Springer, 1993. to appear.
26. K.-H. Hanisch. Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point patterns. *Mathematische Operationsforschung und Statistik, series Statistics*, 15:409–412, 1984.
27. L. Heinrich. Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes. *Statistics*, 19:87–106, 1988.
28. E. Jolivet. Central limit theorem and convergence of empirical processes for stationary point processes. In P. Bastfai and J. Tomko, editors, *Point processes and queueing problems*, pages 117–161. North-Holland, Amsterdam, 1980.
29. O. Kallenberg. *Random measures*. Akademie Verlag/Academic Press, Berlin/New York, third edition, 1983.
30. O. Kallenberg. An informal guide to the theory of conditioning in point processes. *International Statistical Review*, 52:151–164, 1984.
31. J.E. Paloheimo. Discussion contribution to the paper by Matérn. In G.P. Patil, E.C. Pielou, and W. E. Waters, editors, *Statistical ecology, Volume 1: Spatial patterns and statistical distributions*, pages 210–212. Pennsylvania State University Press, University Park, 1971.
32. J.E. Paloheimo. On a theory of search. *Biometrika*, 58:61–75, 1971.
33. C.J. Preston. *Random fields*. Springer Verlag, Berlin-Heidelberg-New York, 1976.
34. B.D. Ripley. *Spatial statistics*. John Wiley and Sons, New York, 1981.
35. B.D. Ripley. *Statistical inference for spatial processes*. Cambridge University Press, 1988.
36. B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188–192, 1977.
37. A. Särkkä. On parameter estimation of Gibbs point processes through the pseudo-likelihood estimation method. Research report 4, Department of Statistics, University of

Jyväskylä, Finland, 1989.

38. A. Särkkä. Applications of Gibbs point processes: pseudo-likelihood method with comparisons. Research report 10, Department of Statistics, University of Jyväskylä, Finland, 1990.

39. M.L. Stein. A new class of estimators for the reduced second moment measure of point processes. Technical Report 278, Department of Statistics, University of Chicago, Chicago, Illinois, USA, July 1990.

40. M.L. Stein. Asymptotically optimal estimation for the reduced second moment measure of point processes. Technical Report 300, Department of Statistics, University of Chicago, Chicago, Illinois, USA, April 1991.

41. D. Stoyan, W. S. Kendall and J. Mecke. *Stochastic Geometry and its Applications.* John Wiley and Sons, Chichester, 1987.

42. R. Takacs. Estimator for the pair-potential of a Gibbsian point process. Institutsbericht 238, Institut für Mathematik, Johannes Kepler Universität Linz, Austria, 1983.

43. R. Takacs. Estimator for the pair potential of a Gibbsian point process. *Statistics*, 17:429–433, 1986.