
PREAMBLE

There's a time for all things.

W. Shakespeare, *Comedy of Errors* (Act II, Scene 2)

Computers are immobile robots that compute, perform tasks in a logical manner, and collate information furnished to them by scientists. The dawn of a computer age is upon science, and computing devices will play an increasingly larger role in scientific research during the coming decades. Ord focuses attention on the increased computer power available today that enables progress to be made in the area of point pattern analysis, by rendering numerically intensive problems soluble, speculating that new computer advances will spur on subsequent developments, too. The purpose of this paper is to review the main stochastic models used in univariate point pattern analysis, their accompanying traditional statistical tests and inference problems, and their extensions to multivariate point pattern analysis. Ripley's commentary complements the contents of this paper by outlining other important developments in the areas of theory, point processes of objects, statistical inference, and mosaic models. He makes three interesting conjectures, namely that (1) point patterns per se are becoming less important, their few golden years having passed, with their future literature becoming less prolific, (2) fundamental problems of computational complexity do not exist for point pattern analysis, with computational intractability being a misnomer applied to situations of ignorance, and (3) point pattern analysis now is being eclipsed by statistical image analysis. Indeed, then, is now the time to begin finalizing the history of point pattern analysis? Ord thinks not; Ripley thinks so!

The Editor



Statistical Methods for Point Pattern Data

J. Keith Ord *

Departments of Management Science and Statistics, The Pennsylvania State University, University Park, PA 16802.

Overview: Recent developments in spatial point processes are reviewed. After presenting an outline of the major univariate stochastic models for such processes, a Poisson-Gaussian model is described which can be made operational for both areal and distance-based sampling. We then summarize work on tests of randomness using distance, quadrat and line transect methods. Robust estimation procedures for the intensity of the process are examined along with the descriptive tools provided by second order processes. Increased computer power has enabled some progress to be made in estimating the parameters of stochastic models from the likelihood function and further activity is likely in this area.

Finally, work on multitype processes is reviewed and future directions for research are outlined.

1. Introduction

The analysis of spatial data has become a major preoccupation of statisticians only rather recently and most of the advances in the study of spatial point patterns have happened within the past ten years. In part, this reflects a neglect of such problems by statisticians since heuristic statistical methods have been employed in other disciplines such as ecology and forestry for over fifty years. However, the other side of the coin is the level of computational difficulty facing statistical model builders. It is only with the computing power available in the 1980's that such difficulties are being overcome.

After outlining the different approaches to data collection in Section 2, we turn in Section 3 to the main stochastic models which have been developed for spatial point patterns. To avoid overworking the term "point" we shall reserve it for points in the study region and prefer the term "individual" for realized events.

After developing the basic models, Sections 4 and 5 cover the traditional areas of tests of "randomness" (*i. e.* whether the pattern is formed by a Poisson process) and statistical inference. In addition to the usual issues of intensity estimation, Section 5 covers the newer topics of second order methods and parametric model-building. Section 6 is devoted to multitype (or multivariate) point processes and the paper concludes with a brief section on the future directions of the subject.

1.1. A bibliography

Apart from the major pioneering effort by Matern (1960), the first theoretical text on spatial point patterns was that of Bartlett (1975). Since then the literature has expanded considerably. Chapter 4 of Cliff and Ord (1981) considers the analysis of spatial point patterns with an emphasis on geographical applications. Getis and Boots (1978) discuss spatial point, line and areal processes, again with an emphasis on problems in geography. Diggle (1983) gives

* I am grateful to several colleagues, especially Brian Ripley, for comments on an earlier version of this paper.

an excellent overview with many ecological examples. The major book of a more theoretical nature is that of Ripley (1981) which also includes a wide variety of applications. Kinderman and Snell (1980) provide a useful introduction to the theory of Markov random fields. Most recently, Upton and Fingleton (1985) provide a lively presentation of the methods of spatial analysis, drawing on examples from many different disciplines. The most recent and complete discussion of inference problems for spatial processes appears in Ripley (1988).

In addition to these texts, the review paper by Ripley (1984a) was a valuable source of information in the preparation of this paper. Further useful sources are the bibliography compiled by Naus (1979) and the volume of papers edited by Cormack and Ord (1979).

2. Data collection

The manner in which data on spatial point processes are collected and recorded has a major impact upon subsequent methods of analysis. When a single study area is selected and all the individuals within that study area are recorded, we say the data are *mapped*. Alternatively, when a series of sites is selected at random and individuals are recorded only in the neighborhood of those sites, we say that the process has been *sparsely sampled*. The definitions and the terminology follow Diggle (1983).

We may select sampling units (known as quadrats, whatever their shape!) of a prespecified size and record the number of individuals present in that sub-area: this is known as the method of quadrat counts. The other principal option is to select individuals, or points, at random and to measure either point-to-individual (PI) or individual-to-individual (II) distances. These distance methods are often known as *nearest-neighbor* methods, although this term should be taken to include all immediate neighbors, not just the nearest one. The more general label of *distance methods* will be used in this paper.

At one time, sparse sampling methods were widely used in ecology, although they are now less popular. Naturally, geographers prefer mapped data analyses. When complete mapping is undertaken, other approaches become feasible, such as "empty space" methods whereby we examine the probability that a region (disk) of given area is devoid of individuals. Subject matter and research objectives should and do play a major role in the types of analyses undertaken.

3. Modeling spatial point processes

The benchmark model for spatial patterns is the homogeneous Poisson process, often known simply as the Poisson process. In order to describe this we consider a planar region C (sample space) and (sub-)regions $A, B \subset C$. The area of such regions is denoted by $|A|$, $|B|$ and so on. Let $N(A)$ be the random variable denoting the number of individuals in A . The expected number of individuals in A is then

$$E\{N(A)\} = \int_A \lambda(\mathbf{x}) d\mathbf{x}, \quad (3.1)$$

where $\lambda(\mathbf{x})$ is the *intensity function* defined at all points $\mathbf{x} \in C$.

3.1. The Poisson Process

The assumptions underlying the Poisson process are

PP1: $\lambda(\mathbf{x}) = \lambda$ for all $\mathbf{x} \in \mathbf{C}$.

PP2: $N(A)$ follows a Poisson distribution with mean $\lambda|A|$.

PP3: Given that $N(A) = n$, the n events in A are independent and form a random sample from the uniform distribution on A .

It then follows that

PP4: If A and B are disjoint regions, $N(A)$ and $N(B)$ are independent.

In fact, it may be shown that **PP1** and **PP2** imply **PP3** (Ripley, 1976b), although the proof is difficult.

Property **PP3** motivates the description of the Poisson process as being “purely random.” Many tests of the Poisson model (see Section 4) are known as tests of “randomness.” Another central property of the Poisson process relates to distance sampling:

PP5: If P is a randomly selected point in the plane, C and I_1 is the individual nearest to P , the distribution of $U_1 = (PI_1)^2$ is exponential with parameter $\pi\lambda$. Further if I_k is the k^{th} nearest individual to P ,

$$U_k = (PI_k)^2 \text{ is gamma } (k, \pi\lambda). \quad (3.2)$$

It seems a rather minor comment to note that, if \mathbf{P} denotes a randomly selected individual rather than a point in the plane, property **PP5** is unchanged. However, this additional result proves very valuable in estimation and testing problems.

3.2. Departures from “randomness”

Given the specialized nature of the assumptions underlying the Poisson process, a variety of departures from **P1** and **P2** may be considered. The two simplest schemes allow (a) heterogeneous intensity (**HI**) and (b) clusters of individuals respectively. Following Diggle (1983, pp 52-55) the heterogeneous intensity model may be described as follows:

HIP1 $N(A)$ has a Poisson distribution with mean (3.1), $\lambda(\mathbf{x}) \neq \lambda$ for all \mathbf{x}

HIP2 Given $N(A) = n$, the individuals in A form a random sample from the distribution on A with pdf proportional to $\lambda(\mathbf{x})$.

Kooijman (1979) discusses the analysis of mapped data when $\lambda(\mathbf{x})$ is represented by a low order polynomial.

An extension of the heterogeneous scheme is to assume that the intensity $\lambda(\mathbf{x})$ is itself determined by a random process, $\Lambda(\mathbf{x})$. This yields the doubly stochastic process or Cox process, first developed in the time domain by Cox (1955); see also Bartlett (1964, 1975) and Matern (1971).

CP1 $\{\Lambda(\mathbf{x}) : \mathbf{x} \in \mathfrak{R}^2\}$ is a non-negative-valued stochastic process.

CP2 Conditional on $\{\Lambda(\mathbf{x}) = \lambda(\mathbf{x}) : \mathbf{x} \in \mathfrak{R}^2\}$ the number of individuals is described by a heterogeneous intensity Poisson process with intensity function $\lambda(\mathbf{x})$.

When the process in CP1 is stationary, it follows that the intensity is

$$\lambda = E[\Lambda(\mathbf{x})].$$

For two points \mathbf{x} and \mathbf{y} the conditional intensity given $\{\Lambda(\mathbf{x})\}$ is $\Lambda(\mathbf{x})\Lambda(\mathbf{y})$, and we define the second order intensity to be

$$\lambda_2(\mathbf{x}, \mathbf{y}) = E\{\Lambda(\mathbf{x})\Lambda(\mathbf{y})\}.$$

When the process is both stationary and isotropic, this becomes

$$\lambda_2(\mathbf{x}, \mathbf{y}) = \lambda^2 + \gamma(t), \quad (3.3)$$

where $\gamma(t) = \text{cov}[\Lambda(\mathbf{x})\Lambda(\mathbf{y})]$ and $t = \|\mathbf{x} - \mathbf{y}\|^{1/2}$ is the distance between \mathbf{x} and \mathbf{y} .

3.2.1. Poisson cluster process

PCP1 The distribution of parent individuals in A follows a Poisson distribution with intensity ρ .

PCP2 Each parent produces a random number, S , of offspring, where S is independent and identically distributed for each parent with probabilities $\{\pi_s, s = 0, 1, \dots\}$.

PCP3 The positions of the offspring relative to their parents are independently and identically distributed according to a bivariate distribution with pdf $f(\cdot)$.

When $f(\cdot)$ represents a degenerate distribution **PCP1** and **PCP2** yield the class of contagious Poisson distributions (cf. Ord, 1972, pp. 126-7).

When S is described by a logarithmic series distribution, it is well-known that the resulting **PCP** is described by the negative binomial distribution. However, when $f(\cdot)$ is non-degenerate, Diggle and Milne (1983a) found that no plausible process exists which could produce counts following the negative binomial law.

Quite generally, the clustering processes may be described by their characteristic functionals (Bartlett, 1964). Bartlett went on to show that the Poisson cluster processes are Cox processes so that the generating mechanisms cannot be distinguished by any data analytic method.

Another existence result of interest is due to Kingman (1977), who showed that for reproducing populations formulated as Cox processes there is no process with independent displacements of individuals which leads to the Poisson process as an equilibrium. However, Kingman went on to demonstrate that when dependent displacements are allowed, a variety of processes lead to a Poisson process in the equilibrium state. These results help to justify the often implicit assumption that purely spatial models may be viewed as equilibrium processes for an (unobserved) spatio-temporal process.

3.3. Gibbs processes

Now consider a process defined on the planar region C such that the joint pdf that there are exactly n individuals in C located at $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is

$$ag_n(\mathbf{x}_1, \dots, \mathbf{x}_n)\rho^n e^{-\rho|C|}/n!,$$

where a is a normalizing constant, $g_n(\cdot)$ is symmetric in its n arguments and ρ represents intensity as before. It follows that

$$P\{N(C) = n\} = [a\rho^n e^{-\rho|C|}/n!] \int_{C^n} g_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \dots d\mathbf{x}_n,$$

so that the joint pdf for the $\{x_i\}$ given $N(C) = n$, is

$$g(\mathbf{x}_1, \dots, \mathbf{x}_n) / \int_{C^n} g(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mathbf{x}_1 \dots d\mathbf{x}_n. \quad (3.4)$$

Writing $g_n = \exp(-\phi_n)$, we may consider ϕ_n to be a potential function and the joint pdf then defines a Gibbs process. In particular, the special case

$$\phi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i < j} \beta_2(\mathbf{x}_i - \mathbf{x}_j) \quad (3.5)$$

has provided the basis for most empirical work in statistics; $\beta_2 \equiv 0$ clearly reduces to the Poisson process. Ripley and Kelly (1977) gave a seminal development of Markov random fields and show that, for a general class of potential functions including (3.5), the Gibbs processes and Markov random fields are equivalent. Further, Ripley and Kelly (1977) demonstrate that the Gibbs processes provide a particularly useful framework for models of inhibitory processes. For example, setting

$$\beta_2(\mathbf{x}_i - \mathbf{x}_j) = 0, \quad \|\mathbf{x}_i - \mathbf{x}_j\| < d_0$$

provides a *hard-core* model whereby two individuals cannot co-exist within d_0 of each other. Following earlier work by Strauss (1975), Kelly and Ripley (1976) developed a clustering model using (3.5). Sanders *et al.* (1982) have extended this scheme to include local hard-core inhibition and clustering. Taking $\phi_n > 0$ ensures that the normalizing factor in (3.4) is finite; otherwise, the process may not be well defined. If $\phi_n < 0$ for some \mathbf{x} , extreme caution is advised.

3.4. Related processes

Several other processes have been developed and many of these are reviewed by Diggle (1983, Chapter 4). Most of the processes described thus far assume stationarity and isotropy. However, this is more for convenience than necessity. For example, the second order properties discussed in Section 5 below also hold for anisotropic processes.

Byth (1981, 1982a) developed a class of processes which are isotropic with respect to a particular "origin," but non-stationary. The intensity is a function of the distance from the "origin." Byth's motivating example was the pattern of fungi around a tree, but her process has potential value for many diffusion-type processes which emanate from a known origin.

A link between lattice processes and point processes has been provided by Besag *et al.* (1982) who show that the auto-Poisson process (of Besag, 1974) on a regular lattice approaches a limiting inhibitory (pairwise interaction) spatial point process.

Another class of models are the thinned processes discussed by Brown and Holgate (1974) and Brown (1979) among others. The method of thinning may be either random or position dependent.

3.5. Nearest-neighbor results

As suggested by property **PP5** in Section 3.1, distances from randomly selected points (or individuals) to first, second, ..., nearest individuals, often known as nearest-neighbor distances, are widely used in the analysis of spatial point patterns. Further recent developments in this area include Cox (1983) on the probability that m^{th} nearest neighbors to individual I_1 , are also n^{th} nearest neighbors to I_2 ; Cox also develops higher order results of this type. Newman *et al.* (1983) consider the probability that an individual is the nearest neighbor of exactly k other individuals; it is found that the probabilities are of Poisson form with $\lambda = 1$, for a variety of processes. Pickard (1982) gives a general treatment of the isolated nearest neighbor (or reflexive pairs) problem. Warren (1971) and Warren and Batcheler (1979) give nearest neighbor distributions for several specific non-Poisson schemes.

3.6. The heterogeneous Poisson-Gaussian process

The lack of any plausible spatial process underlying the negative binomial distribution for quadrat counts has been demonstrated by Diggle and Milne (1983a). This makes it difficult to interrelate analyses based on areal and distance sampling procedures yet, especially for mapped processes, some interconnection between these approaches is essential.

A review of the assumptions underlying Cox processes (Section 3.2) indicates that the intensity process must be non-negative but also infinitely divisible so that spatial aggregates may be defined for any region. A natural choice would be the Gaussian scheme except that, of course, it may take on negative values. We could condition upon $\Lambda(\mathbf{x}) \geq 0$ for all \mathbf{x} ; done rigorously, however, this imposes considerable additional complexity upon the analysis. In practice, provided the probability of negative values is sufficiently small we need apply this conditioning only notionally and, algebraically at least, may consider the complete Gaussian process as a reasonable approximation. This approach has been employed in models for discrete data where the Poisson mixture with the normal gives rise to the Hermite distribution; Kemp and Papageorgiou (1982) describe the bivariate case and give references to earlier work. However, it should be noted that the correlation function must be non-negative for all $\mathbf{x} - \mathbf{y}$ to ensure a well-defined process (see Section 3.3). We now develop these ideas for a particular Cox process.

Following Bartlett (1975, p. 7) we may write the characteristic functional (c.fl.) for all points \mathbf{u} in some region Q as

$$C(\boldsymbol{\theta}) = E_{\Lambda}[\exp \int \Lambda(\mathbf{u})\{z(\mathbf{u}) - 1\} d\mathbf{u}], \quad z = e^{i\boldsymbol{\theta}},$$

where the integral is taken over all points $\mathbf{u} \in Q$. This expression is more readily understood if we first consider Q to contain a finite number of points when the c.fl. becomes a multivariate probability generating function (p.g.f.), viewed as an argument in \mathbf{z} rather than $\boldsymbol{\theta}$.

When $\Lambda(\mathbf{x})$ is a Gaussian process with mean function $\mu(\mathbf{x}) = \mu$, variance function $\omega(\mathbf{x}) = \omega$ and autocorrelation function

$$\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{x} - \mathbf{y}), \text{ for all } \mathbf{x} \text{ and } \mathbf{y},$$

with $\rho(\mathbf{0}) = 1$, it follows that

$$C(\boldsymbol{\theta}) = \exp[\mu \int \{z(\mathbf{u}) - 1\} d\mathbf{u} + 0.5\omega \iint \rho(\mathbf{u} - \mathbf{v})\{z(\mathbf{u}) - 1\}\{z(\mathbf{v}) - 1\} d\mathbf{u} d\mathbf{v}]. \quad (3.6)$$

If sampling is by quadrats, the p.g.f. for the quadrat count is obtained by setting $z(\mathbf{u}) = z$ for all $\mathbf{u} \in Q$. If Q has area A , (3.6) yields the p.g.f.

$$C_Q(z) = \exp[A\mu(z-1) + 0.5\omega(z-1)^2 \iint \rho(\mathbf{u}-\mathbf{v}) d\mathbf{u} d\mathbf{v}], \quad (3.7)$$

where the double integral is taken over $(\mathbf{u}, \mathbf{v}) \in Q$. Thus, (3.7) defines an Hermite distribution with p.g.f. of the general form

$$\exp[\lambda_1(z-1) + \lambda_2(z-1)^2], \quad (3.8)$$

reducing to the Poisson when $\lambda_2 = 0$. It follows from (3.8) that the quadrat count, R , has

$$E(R) = \lambda_1 < \text{var}(R) = \lambda_1 + 2\lambda_2, \text{ for } \lambda_2 > 0.$$

Expression (3.6) simplifies further in those cases where an explicit functional form is available for the correlation. In particular, suppose that the process is isotropic and

$$\rho(\mathbf{w}) = \exp\{-\alpha^2(w_1^2 + w_2^2)/2\}. \quad (3.9)$$

Further, we assume that the quadrats are rectangular, say h_1 by h_2 , with $h_1 \cdot h_2 = A$. The integral in (3.7) partitions into two integrals of the form

$$\begin{aligned} G(h) &= \int_0^h \int_0^h \exp\{-\alpha^2(u-v)^2/2\} du dv \\ &= [4(2\pi)^{1/2}/\alpha^2][\alpha h F(\alpha h) - \alpha h/2 + f(\alpha h) - f(0)], \end{aligned}$$

where f and F denote the density and distribution functions, respectively, for the standard normal. Thus, the p.g.f. for a single quadrat becomes

$$C_Q(z) = \exp[A\mu(z-1) + \omega(z-1)^2 G(h_1)G(h_2)/2]. \quad (3.10)$$

By extending the argument to two (or more) quadrats, we obtain bivariate (multivariate) Hermite distributions. Comparison of (3.8) and (3.10) indicates that

$$\lambda_1 = A\mu \quad \text{and} \quad \lambda_2 = \lambda_2(\omega, \alpha, h_1, h_2),$$

so that there are three unknown parameters (μ, ω, α) relating to the mean, variance and correlation structures respectively. This is in a one-to-one correspondence with the bivariate Hermite with identical marginal distributions. The simplest procedure inferentially is probably to estimate the parameters of the bivariate Hermite using a pseudo-likelihood approach and then relate these values to the process parameters using the moments derived from (3.10) and its bivariate extension. Detailed development of the method is necessary, but it appears to have some potential.

Finally, we note that the process also corresponds to a **PCP** when the cluster size, S , is binomially distributed with index $n = 2$.

4. Tests of randomness

The Poisson process forms a natural null hypothesis for testing whether structure exists within a spatial data set. Such tests are often known as "tests of randomness" and may be designed with particular alternatives (clustering, inhibition) in mind or may use the completely general alternative of a non-Poisson process.

As has been remarked by Ripley (1981) and others, tests of randomness should represent only the first step in a spatial modeling paradigm, yet it has to be admitted that such tests have often been presented as the end result of an analysis. Either way, such tests clearly have value and we now summarize the main options for distance and area-based sampling respectively.

4.1. Distance methods

The majority of the tests proposed are based upon distance sampling. A summary list is presented in Table 1. When the region is sparsely sampled, the distributional properties listed hold. For mapped data, these results hold up less well (Byth and Ripley, 1980). An alternative procedure is to use Monte Carlo testing (cf. Cliff and Ord, 1981, pp 63-65; Diggle, 1979a). Diggle and Gratton (1984) present the first systematic development of inferential procedures based upon Monte Carlo methods.

Several comparisons of the relative power of these tests have been performed recently; their conclusions may be summarized as follows:

- (1) In general, tests which include squared distances perform better than those which do not.
- (2) The Hopkins (D) test is most powerful across a range of alternatives but is non-feasible for sparse data since randomly selected individuals cannot be selected (Besag and Gleaves, 1973). For mapped data the Hopkins test is often a good choice and the test may be performed by Monte Carlo methods (cf. Diggle, 1979a). Byth and Ripley (1980) have developed a semi-systematic sampling scheme for choosing individuals "at random," which allows use of the Hopkins test. When the semi-systematic sampling is feasible, this approach would appear to give the best power.
- (3) The Besag-Gleaves (G, H) and Cox-Lewis (L) tests perform similarly across a variety of alternatives. Cox-Lewis has a slight edge but has a somewhat more difficult distribution theory. The Eberhardt (J) and Holgate (E, F) statistics are less powerful (Hines and Hines, 1979). The Hines-Hines (K) procedure is comparable in power to the Hopkins (D) test for clustered alternatives, but weaker for regular alternatives (Hines and Hines, 1979).
- (4) Tests (R), (S), (T) and (A) perform better than tests (P) and (Q) and several other alternatives (Ripley, 1979a). Tests (R), (S) and (T) have the advantage that they may suggest an alternative model, but (S) and (T) are clearly dependent on the values of the constants $C1$ and $C2$.

Clayton (1984) suggests a procedure based upon the use of an exclusion angle, but its relative performance characteristics are as yet unknown.

Ripley and Silverman (1978) demonstrate that, when the alternative hypothesis is a Poisson hard-disk process, the uniformly most powerful test is based upon the smallest order

TABLE 1
DISTANCE TESTS OF RANDOMNESS

Code Letter	Statistic	Distribution under H_0	Reference
A	$2\sqrt{\lambda}\Sigma r_i/n$	approx $N[1, (4 - \pi)/(\pi n)]$	Clark & Evans (1954)
B	$\pi\lambda\Sigma u_i/n$	approx $N[1, (\lambda A + n + 1)/(n\lambda A)]$	Pielou (1959), Mountford (1961)
C	$2\pi\lambda\Sigma u_i$	approx χ_{2n}^2	Skellam (1951)
D	$\Sigma u_i/\Sigma(u_i + u_i^*)$	$B(n, n)$	Hopkins (1954)
E	$\Sigma u_i/\Sigma u_{i2}$	$B(n, n)$	Holgate (1965)
F	$\Sigma(u_i/u_{i2})/n$	approx $N(1/2, 1/12n)$	Holgate (1965)
G	$\Sigma u_i/\Sigma(u_i + .5v_{iT})$	$B(n, n)$	Besag and Gleaves (1973)
H	$\Sigma\{u_i/(u_i + .5v_{iT})\}/n$	approx $N(1/2, 1/12n)$	Besag and Gleaves (1973)
I	$-.5\Sigma\ell n\{u_i/(u_i + .5v_{iT})\}$	χ_{2n}^2	Besag and Gleaves (1973)
J	$n\Sigma u_i/(\Sigma r_i)^2$	see Hines & Hines (1979)	Eberhardt (1967)
K	$\Sigma(u_i + .5v_{iT})/\Sigma(r_i + .5\sqrt{v_{iT}})$	see reference for tables	Hines and Hines (1979)
L	$4/3\Sigma(1 - \pi w_i)$	approx $N(1/2, 1/12n)$	Cox and Lewis (1976), Cormack (1977)
M	$\min_i(w_i)$	beta (1, n)	Cox and Lewis (1976)
N	$n\ell n[\Sigma(u_i + .5v_{iT})/n] - \Sigma\ell n(u_i + v_{iT})$	$\chi^2(n - 1)$	Diggle (1977b)
P	$\Sigma(v_i - D)^2/D^2(n - 1)$	by simulation	Brown and Rothery (1978)
Q	$(\pi v_i)^{1/n}/D$ ($D = \Sigma v_i/n$)	by simulation	Brown and Rothery (1978)
R	$\sup_{t \leq t_0} L(t) - t $	by simulation	Ripley (1979a), p. 369.
S	$\Sigma\phi(\mathbf{x}, \mathbf{y})$	asympt. normal	Liebetrau (1977)
T	$\Sigma\phi(\mathbf{x}, \mathbf{y})$	asympt. normal	Jolivet (1978)

Notation for Table 1

r_i = distance from random point to the nearest individual

$$u_i = r_i^2$$

r_{ip} = distance from random point to p-th nearest individual

$$u_{ip} = r_{ip}^2$$

r_i^* = distance from random individual to its nearest neighbor

$$u_i^* = (r_i^*)^2$$

v_i = squared distance from nearest individual to a random point to its nearest neighbor

v_{iT} = v_i , but with nearest neighbor restricted to T-square sampling

$$w_i^{-1} = 2\pi + \sin\theta_i - (\pi + \theta_i)\sin\theta_i; \sin(0.5\theta_i) = (v_i/u_i)^{1/2}$$

$L(t) = [\hat{K}(t)/\pi]^{1/2}$, see (5.3) for definition of $\hat{K}(t)$

$$\phi(\mathbf{x}, \mathbf{y}) = [c_1 - |x_1 - y_1|][c_2 - |x_2 - y_2|] \text{ if } |x_i - y_i| > c_i, i = 1, 2; = 0, \text{ otherwise.}$$

statistic, $r_{(1)}$, of a sample of first nearest-neighbor distances. Concerns about measurement and rounding errors may lead to the use of higher order statistics, but this would occasion loss of power unless the errors are substantial. Saunders and Funk (1977) demonstrate that

for a disk of radius r_0 , $r_{(1)}^2 - r_0^2$ is approximately exponentially distributed; for further developments, see Silverman and Brown (1978).

As noted earlier, the Poisson cluster process may be represented as a Cox process so that clustering and heterogeneity cannot be distinguished. More pragmatically, it may be argued that clustering is essentially a local phenomenon whereas heterogeneity is manifested at a larger scale; that is, $\Lambda(\mathbf{x})$ changes slowly with \mathbf{x} . From this perspective, Diggle (1977b) proposed test (N) as a test of heterogeneity. Diggle recommends that one of the local distance tests be used to detect local departures from randomness and then test (N) be used to detect larger scale patterns. As such it represents an alternative procedure to the transect methods described in Section 4.3. The Diggle procedure has the advantages of being somewhat simpler to apply and being applicable to both sparsely sampled and mapped data. Its power, relative to the methods in Section 4.3, is unknown.

4.2. Areal sampling methods

When data are collected from a Poisson process, the counts distribution is Poisson whether those data are mapped or sparsely sampled. In principle, any goodness-of-fit test could be used to test the null hypothesis that the distribution is Poisson. However, tests such as chi-square have been found to have relatively low power since it is the upper tail observations which serve to distinguish non-Poisson alternatives. The most popular test is the index of dispersion

$$D = \text{Sample variance/sample mean.}$$

Perry and Mead (1979) show that D has good power properties in sparse sampling. Heltshe and Ritchey (1984) show that Stevens' test, defined as

$$Z = \text{number of empty quadrats} \mid S \text{ individuals in } N \text{ quadrats}$$

has power comparable to D for aggregated alternatives, but does not perform as well for regular patterns when large quadrats are used. This is understandable since

$$E(Z) \doteq Ne^{-\lambda}, \text{ where } S = N\lambda,$$

which becomes small for large λ and fixed N .

For mapped data, quadrat methods necessarily lose power as they ignore the spatial dependence between quadrats; the distance-based methods of Section 5.2 are preferable. When only quadrat counts are available, the D test can be complemented by a test for spatial autocorrelation between neighboring quadrats (Cliff and Ord, 1981, pp. 97-99).

4.3. Tests using transect data

There is a considerable literature on line transects which we shall not review in detail; see Gates (1979) and De Vries (1979). One topic relevant to our present discussion is the use of line transect data to test for different scales of spatial pattern. These methods are associated with the name of Greig-Smith (1952) who has used the approach extensively in his ecological work; see Greig-Smith (1979).

Briefly, the computational details of the method may be described as a hierarchical analysis of variance. As noted by Professor Bartlett and others, formal ANOVA tests are

unreliable since detection of one scale of spatial pattern may obscure higher levels. Mead (1974) developed a randomization procedure which avoids the difficulties noted by Bartlett.

In general, the Greig-Smith and Mead procedures have single starting points along the transect (or two-dimensional array). Hill's (1973) method gave improved power by considering all possible starting points and Upton's (1984) procedure does the same for Mead's test. Zahl (1977) developed a Scheffé-type procedure which considers all overlapping blocks and controls the overall probability of Type I error; simulation results indicate that this technique outperforms the Greig-Smith method, and also the random quadrat method of Goodall (1974). Moellering and Tobler (1972) and Cliff and Ord (1981, pp. 123-6) provide geographical applications of this approach.

A different approach is the use of spectral analysis, illustrated by Ripley (1978). This analysis appears to give at least as clear results as spatial domain methods and may indeed be a better vehicle for separating out different scales of pattern; see also Ripley (1981, pp. 112-129), Renshaw and Ford (1983) and Renshaw (1984).

5. Statistical inference

Inference for spatial point processes may focus upon either the first and second order properties required for weak stationarity or upon strict stationarity when the more ambitious goal of fitting a complete parametric model is attempted. We shall examine these in turn.

5.1. Intensity of a process

The simplest and most traditional problem is that of estimating process intensity, or the average number of individuals per unit area. When the process is stationary and isotropic, this is only a problem for sparsely sampled data; for non-stationary processes with mapped data the smoothing methods used in probability density estimation are particularly useful, see Cox (1979), Diggle (1981a) and, more generally, Silverman (1981).

For stationary and isotropic processes, the obvious real estimate

$$\hat{\lambda} = \text{Total number of individuals} / (\text{Number of quadrats}) \cdot (\text{Quadrat area})$$

is unbiased for any spatial process. However, it is often the case that quadrat sampling is either too expensive or impractical which has led to an extensive search for distance-based estimators. Many of the early estimators were based upon precisely the same statistics used to test for randomness (see Table 1); this includes the maximum likelihood estimators for the Poisson process (cf. Pollard, 1971). To the extent that the tests are successful, the lack of robustness of the estimators is perhaps not surprising. Persson (1971) made this point forcibly and suggested several more robust alternatives; see Table 2. Robustness in this context is a somewhat elusive concept. It is very easy to consider a particular departure from a Poisson point process and to produce an estimator that is robust to that change, yet not robust to others. For example, if a pattern includes very tight clusters of variable size, any distance-based estimator is likely to fail for some configurations. Thus, robustness is a quality to be assessed in the eye of the beholder and the researcher should determine whether the method to be used will be robust for the spatial processes likely to be encountered. We now summarize the evidence on different estimators, bearing in mind that published studies have tended to focus upon a rather limited variety of non-Poisson processes.

TABLE 2
ROBUST ESTIMATORS FOR INTENSITY

Code Letter	Estimator	Reference
A	$c/(\sum r_{ip})^2$	Persson (1971)
B	$c/(\sum u_{ip})$	Persson (1971)
C	$c\sum u_{ip}^{-1}$	Persson (1971), Cox (1976)
D	$c/\text{median}(u_{ip})$	Persson (1971)
E	$c/(\sum u_i \sum u_i^*)^{1/2}$	Diggle (1975)
F	$c/(\sum u_i \sum v_{iT})^{1/2}$	Diggle (1975)
G	$c_1(\sum u_i)^{-1} + c_2(\sum u_{i2})^{-1}$	Lewis (1975)
H	$c_1 \sum (u_i/w_i) + c_2(\sum u_{i2})^{-1}$	Cox (1976)
I	$c_1/(\sum u_i)c_2^{-K}$	Batcheler and Hodder (1975), Warren and Batcheler (1979)
J	$c k(n)u[k(n)]/n$	Patil <i>et al.</i> (1979, 1982)
K	$c/(\sum r_i)(\sum v_{iT}^{1/2})$	Byth (1982b)
L	$c/(\sum r_i \sum r_i^*)$	Clayton and Cox (1986)

Notation for Table 2

$c, c_1, c_2 = \text{constants}$

$$K = c\sum(r_i - \bar{r})^2/(\sum r_i)(\sum v_i^{1/2})$$

$k(n) = \text{function of } n \text{ such that } k(n) \rightarrow \infty \text{ and } k(n)/n \rightarrow 0 \text{ as } n \rightarrow \infty,$

$$[e. g., k(n) = n^{1/2}]$$

$u[k(n)] = [k(n)]\text{th order statistic from } u_1, \dots, u_n$

Other notation as for Table 1.

Diggle (1977a) shows that the estimators (E) and (F) are more robust than the earlier suggestions (A-D); (E) performs better than (F) when the clusters are rather diffuse but, of course, (E) may not be usable in practice. Cox (1976) shows (H) to be better than (C). An extensive Monte Carlo study by Byth (1982b) indicates that (E) and (F) perform better than (A-D) or (I), but suggests that (K) may be best of all unless there is a very high degree of clustering. The evidence on other methods is incomplete although method (J) appears to have a high standard error despite its low degree of bias; the several possible estimators based upon (L) seem to perform quite well.

Aherne and Diggle (1978) recommend the use of (E) or (F) in conjunction with a preliminary test of randomness. This opens up the possibility of using different estimators according to whether the spatial pattern is judged regular, random or clustered. For forestry applications, Ord (1978) gives an unbiased estimator based upon the angle-exclusion method.

5.2. Second moment methods

The second order intensity function may be defined more generally than in (3.3) as

$$\lambda_2(\mathbf{x}, \mathbf{y}) = \lim_{\substack{|d\mathbf{x}| \rightarrow 0 \\ |d\mathbf{y}| \rightarrow 0}} E[N(d\mathbf{x})N(d\mathbf{y})]/(|d\mathbf{x}| \cdot |d\mathbf{y}|). \quad (5.1)$$

For stationary processes $\lambda_2(\mathbf{x}, \mathbf{y}) = \lambda_2(\mathbf{x} - \mathbf{y})$ and when the process is also isotropic

$$\lambda_2(\mathbf{x}, \mathbf{y}) = \lambda_2(t),$$

where $t^2 = \|\mathbf{x} - \mathbf{y}\|$. For the Poisson process, $\lambda_2(\mathbf{x}, \mathbf{y}) = \lambda^2$ for all \mathbf{x} and \mathbf{y} .

An alternative representation of second order properties is by means of the function (Ripley, 1976a):

$$K(t) = \lambda^{-1} E[N(t)] \quad (5.2)$$

where $N(t)$ denotes the number of further individuals within distance t of a given individual. For the Poisson process, $K(t) = \pi t^2$. The function $K(t)$ is a natural representation for mapped data since an unbiased estimator is then available as

$$\hat{K}(t) = |A| \sum k(\mathbf{x}, \mathbf{y}) / N^2, \quad (5.3)$$

where

$$\begin{aligned} |A| &= \text{area of the study region } A, \\ N &= \text{number of points in } A, \text{ and} \\ [k(\mathbf{x}, \mathbf{y})]^{-1} &= \text{proportion within } A \text{ of a disk of radius } t \text{ centered on } \mathbf{x} \\ &\quad \text{(and passing through } \mathbf{y}), \end{aligned}$$

and the summation is taken over all pairs $(\mathbf{x}, \mathbf{y}) \in A$ less than distance t apart. Other edge corrections are possible; see Ohser and Stoyan (1981). \hat{K} was introduced by Ripley (1976a); Ripley (1979) shows that the distribution of \hat{K} is approximately Poisson for distances small relative to the size of the study area. A normal approximation is reasonable for $N \geq 50$ provided the average number of points per unit area is not too small.

For non-Poisson processes, the distribution theory is intractable, but the sample $\hat{K}(t)$ may be contrasted with the envelope of a set of random simulations to determine agreement with a specified model (cf. Ripley, 1977; Diggle, 1979a,b; Diggle and Gratton, 1984). Silverman (1978) and Besag (in the discussion on Ripley, 1977) note that $L(t) = [\hat{K}(t)/\pi]^{1/2}$ has a linear plot against t and also $\text{var}[L(t)]$ is approximately constant. This provides the basis for test (R) in Table 1.

As defined, $K(t)$ requires both stationarity and isotropy; however, isotropy is not necessary (Ripley, 1976a). Ohser and Stoyan (1981) suggest plotting a "rose" of directions with $\lambda K(t, \gamma)$ defined as the number of individuals within t of the given individual, with an orientation angle $\theta \leq \gamma$. Hanisch (1983) considers higher order moments, both for isotropic and anisotropic processes.

Another form of analysis useful for mapped patterns is the estimation of

$$P(t) = P[\text{circle of radius } t \text{ contains no individuals}] \quad (5.4)$$

sometimes known as “empty space” methods. Again, edge corrections are necessary to produce an unbiased estimator (Ripley, 1977, 1984b,c). The information content of (5.2) and (5.4) is different, so both analyses should be performed (Ripley, 1977). Cox (1979) describes a simple approach for identifying “sparse” and “dense” regions which may then be incorporated into computer mapping of the study region. For further discussion of second order methods see Getis and Franklin (1987).

5.3. Model-based inference

The fitting of parametric process models to spatial point patterns has proved extremely difficult once the Poisson scheme is rejected as inadequate. Primarily this is due to the very heavy computational burdens which must be borne. These problems arise in evaluating normalization constants for Gibbs processes as in (3.3) or generally in evaluating joint likelihood functions. Ogata and Tanemura (1981, 1984) developed the likelihood for a Gibbs process based on a variety of local pair-potential functions satisfying (3.4). Even so, several heroic approximations were necessary before the computational effort proved feasible. Gates and Westcott (1986) showed that some of the Ogata-Tanemura simulated examples were unrealistic and that the potential functions under consideration violated a stability condition, at least for some parameter combinations. They conclude that data analysis based on models with unstable potential functions must be performed with great care and, in general, recommend the use of more traditional clustering methods. Alternative approximations for hard-core models are given by Westcott (1982).

Shapiro, Schein and De Monasterio (1985) give an interesting example of modeling a spatial point process. Their paper should be read in conjunction with the ensuing discussion by Diggle and Gates.

Ogata and Tanemura (1984) introduce a “very soft core” (VSC) model with potential function

$$\beta(r) = -\log[1 - \exp(-\alpha r^2)],$$

where $r^2 = \|\mathbf{x} - \mathbf{y}\|$. The VSC process represents a weaker form of local inhibition than the hard-core models where rounding error is present.

Sager (1982) develops a non-parametric maximum likelihood estimation procedure using smoothing methods. Diggle and Gratton (1984) further develop the use of kernel methods for estimating the likelihood of an implicitly defined process. They give suitable numerical procedures for obtaining the maximum likelihood estimators from the estimated likelihood function. The central idea, developed from Diggle (1978), is that the sample realization, described by some function such as $\hat{K}(t)$, is fitted as closely as possible to the theoretical function $\hat{K}(t, \theta)$ estimated from S simulated realizations of the process for different values of the parameters θ . The intractability of the likelihood functions led Diggle and Gratton to determine $\hat{\theta}$ by minimizing

$$d_s(\theta) = \int_0^{t_0} \left[\{\hat{K}(t)\}^{1/2} - \left\{ \frac{1}{s} \sum_{j=1}^s \hat{K}_j(t, \theta) \right\}^{1/2} \right]^2 dt$$

so that their estimates are “sensible” rather than “optimal.” Diggle, Gates and Stibbard (1987) give an improved non-parametric estimator for interaction processes based on an

approximation drawn from statistical physics. This works well in simulations and probably represents the best approach to model estimation at the present time, although it is not immune from some numerical problems.

5.4. Simulation of processes

The simulation of spatial point processes represents another computer-intensive task. When the spatial model can be represented as the equilibrium form of a spatio-temporal process, the scheme can be run over sufficient time periods until an equilibrium may be assumed to have been achieved. Kelly and Ripley (1976) give a rejection method for hard-core processes; see also Lotwick (1981). The Dirichlet tessellation algorithm of Green and Sibson (1978) is used to identify empty areas and to speed up the assignment of individuals to available spaces. Lewis and Shedler (1979) simulate a non-homogeneous Poisson process by thinning. For further discussion, see Ripley (1987).

6. Multitype processes

Point processes involving several types of individual are variously known as *multivariate*, *marked* or *multitype* processes; we shall use the last of these terms. A useful discussion of the multitype process is given in Diggle (1983, Chapter 6).

6.1. Bivariate Cox processes

A convenient way to approach such schemes is to allow the individuals to be generated by a mechanism similar to the univariate scheme, and then to superimpose a "marking" scheme. The specification of the basic bivariate Cox process (Diggle, 1983, pp. 96-98) illustrates this:

BCP1 $\Lambda(\mathbf{x}) = \{\Lambda_1(\mathbf{x}), \Lambda_2(\mathbf{x})\}$ is defined for all $\mathbf{x} \in \mathbb{R}^2$ and represents a pair of non-negative valued stochastic processes.

BCP2 Conditional on $\Lambda_j(\mathbf{x}) = \lambda_j(\mathbf{x})$, $j = 1, 2$ and all $\mathbf{x} \in \mathbb{R}^2$, type 1 and type 2 events form a pair of independent non-homogeneous Poisson processes with intensity functions $\lambda_j(\mathbf{x})$.

Dependence between the processes may then be achieved by modifying **BCP2**. For example,

$\Lambda_2(\mathbf{x}) = c\Lambda_1(\mathbf{x})$ represents extreme positive correlation

$\Lambda_1(\mathbf{x}) + a\Lambda_2(\mathbf{x}) = b$ represents extreme negative correlation, and

$\Lambda_j(\mathbf{x}) = \Lambda_0(\mathbf{x}) + \Lambda_j^+(\mathbf{x})$ represents a variety of processes with positive correlation.

The class of **BCP** schemes is studied in detail by Diggle and Milne (1983b), who also consider thinned bivariate processes. Brown *et al.* (1982) develop a class of multitype processes where each marginal process is Poisson but there is negative correlation between the processes.

Isham (1984) has developed a bivariate Markov point process extending the earlier work of Ripley and Kelly (1977). Ogata and Tanemura (1985) have extended their potential function models to the bivariate case.

6.2. Tests of dependence

Let R_{ij} denote the distance from an arbitrary individual of type i to the nearest individual of type j and the subscript $i = 0$ denotes a randomly selected point in the study region. Let $F_{ij}(r) = P(R_{ij} \leq r)$.

If the two point processes are independent, stationary and isotropic it follows that

$$F_{01}(r) = F_{21}(r) \quad \text{and} \quad F_{02}(r) = F_{12}(r). \quad (6.1)$$

Goodall (1965) tested the equality of the distributions in (6.1) by applying a two sample t-test to the square roots of the distances. Diggle and Cox (1981, 1983) show that Goodall's test is fairly robust against departures from the assumptions, but recommend the use of the corresponding Mann-Whitney test. A second pair of tests follows from testing the correlation between R_{01} and R_{02} for a pair of individuals near randomly selected points. The test based upon Kendall's tau appears to perform best but the form of the alternative hypothesis is critical. An advantage of the correlation test is that it does not require random sampling of individuals.

6.3. Second order properties

The appropriate second order function is

$$K_{ij}(t) = \lambda_j^{-1} E[N_{ij}(t)], \quad (6.2)$$

where $N_{ij}(t)$ denotes the number of type j individuals within distance t of an arbitrary type i individual, which follows directly from the univariate case (Ripley, 1976). Two estimators, $\hat{K}_{12}(t)$ and $\hat{K}_{21}(t)$, are available and since $K_{12}(t) = K_{21}(t)$, a simple average of the estimates is usually taken for each value of t (Lotwick and Silverman, 1982). Examples of data analyses are given in Diggle and Milne (1983b), Lotwick and Silverman (1982) and Harkness and Isham (1983). Lotwick (1984) shows that stationary ergodic processes exist for which the interactions cannot be detected using K , so that empty space methods should also be examined; the extension from (5.4) is relatively straightforward.

Byth (1981) develops methods for isotropic but non-stationary multitype processes.

7. Related topics

It is inevitable that any review must draw boundaries around the subject matter that is considered "most relevant" and omit discussion of other topics. Invariably, some topics are caught on the boundary. In this case, perhaps the two major omissions are mosaics and sampling. Comprehensive reviews of earlier work on mosaics is provided by Pielou (1977, pp. 181-199) and by Getis and Boots (1978, Chapters 6 and 7). Roach (1968) should also be consulted. Diggle (1981b) suggests a stochastic model for binary mosaics as a union of overlapping disks and uses the model to describe the spatial pattern of heather. Hall (1985) describes an alternative modeling approach; also see Ripley (1986) for discussion of this model in the context of pattern recognition methods.

A general view of recent developments in sampling is provided by the volume of papers edited by Cormack *et al.* (1979). With regard to sampling point patterns, the early work of Matern (1960) has been followed up by Diggle and Matern (1981) among others; Byth and Ripley (1980) also discuss sampling procedures for distance measurements.

For other recent developments in spatial processes, the review paper of Ripley (1984a) should be consulted. This paper has restricted attention to purely spatial, or static, models. For a recent survey of dynamic spatial models including birth-death and epidemic processes, see Renshaw (1986).

8. Future directions

The field of spatial point processes has undergone very major changes in the past fifteen years and it is a research area where the computer revolution will continue to have a strong impact on future model-building efforts.

It would seem that the major classes of stochastic model have been developed, at least for the univariate case, although further work is merited on interesting special cases, such as the Poisson-Gaussian. Much more research is needed to develop effective inferential methods. The combined use of likelihood approximations, Monte Carlo methods and second order methods seems to offer a fruitful road ahead for both univariate and multitype processes.

The development of tests for randomness would seem to have reached the point of diminishing returns, although some useful work remains to be done on power comparisons; perhaps similar comments are true for the robust estimation of intensity, although the use of preliminary tests in estimation seems worthy of further investigation.

An area with considerable potential for development is the use of more complex sampling schemes such as the semi-systematic scheme of Byth and Ripley (1980) or the divided quadrats scheme of Ord (1970). Close collaboration with researchers in the field will be necessary to ensure that such methods are cost-effective as well as statistically sound.

Finally, it should be noted that work on dependence among multitype processes has only just started and much remains to be done both to describe patterns of dependence and to develop the necessary inferential tools.

9. References

- Aherne, W., and P. Diggle. (1978) The estimation of removal population density by a robust method. *Journal of Microscopy*, **114**, 285-293.
- Bartlett, M. (1964) The spectral analysis of two-dimensional point processes. *Biometrika*, **51**, 299-311.
- Bartlett, M. (1975) *The Statistical Analysis of Spatial Pattern*. London: Chapman and Hall.
- Batcheler, C., and R. Hodder. (1975) Tests of a distance technique for inventory of pine populations. *New Zealand Journal of Forestry Science*, **5**, 3-17.
- Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, **36B**, 192-236.
- Besag, J., and J. Gleaves. (1973) On the detection of spatial pattern in plant communities. *Bulletin of the International Statistical Institute*, **45**, 153-158.
- Besag, J., R. Milne, and S. Zachary. (1982) Point process limits of lattice processes. *Journal of Applied Probability*, **19**, 210-216.
- Brown, D., and P. Rothery (1978) Randomness and local regularity of points in a plane. *Biometrika*, **65**, 115-122.
- Brown, S., and P. Holgate. (1974) The thinned plantation. *Biometrika*, **61**, 253-262.

- Brown, T. (1979) Position dependent and stochastic thinning of point processes. *Stochastic Processes and Their Applications*, **9**, 189-193.
- Brown, T., B. Silverman, and R. Milne. (1981) A class of two-type point processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **58**, 299-308.
- Byth, K. (1981) θ -stationary point processes and their second-order analysis. *Journal of Applied Probability*, **18**, 864-878.
- Byth, K. (1982a) On kernel methods of estimating marginal, radial and angular probability density functions. *Biometrical Journal*, **24**, 49-58.
- Byth, K. (1982b) On robust distance-based density estimators. *Biometrics*, **38**, 127-135.
- Byth, K., and B. Ripley. (1980) On sampling spatial patterns by distance methods. *Biometrics*, **36**, 279-284.
- Clark, P., and F. Evans. (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, **35**, 445-453.
- Clayton, G. (1984) On the random-pairs method of density estimation. *Biometrics*, **40**, 199-202.
- Clayton, G., and T. Cox. (1986) Some robust density estimators for spatial point processes. *Biometrics*, **42**, 753-67.
- Cliff, A., and J. Ord. (1981) *Spatial Processes: Models and Applications*. London: Pion.
- Cormack, R. (1977) The invariance of Cox and Lewis' statistic for the analysis of spatial patterns. *Biometrika*, **64**, 143-144.
- Cormack, R. (1979) Spatial aspects of competition between individuals, in *Spatial and Temporal Analysis in Ecology*, edited by R. Cormack and J. Ord. Fairland, Md.: International Cooperative Publishing House, pp. 151-211.
- Cormack, R., and J. Ord (eds.) (1979) *Spatial and Temporal Analysis in Ecology*. Fairland, Md.: International Cooperative Publishing House.
- Cormack, R., G. Patil, and D. Robson (eds.) (1979) *Sampling Biological Populations*. Fairland, Md.: International Cooperative Publishing House.
- Cox, D. (1955) Some statistical methods related with series of events. *Journal of the Royal Statistical Society*, **17B**, 129-164.
- Cox, T. (1976) The robust estimation of the density of a forest stand using a new conditioned distance method. *Biometrika*, **63**, 493-500.
- Cox, T. (1979) A method for mapping the dense and sparse regions of a forest stand. *Applied Statistics*, **28**, 14-19.
- Cox, T. (1983) Nearest neighbors to nearest neighbors. *Statistics & Probability Letters*, **1**, 161-166.
- Cox, T., and T. Lewis. (1976) A conditioned distance ratio method for analyzing spatial patterns. *Biometrika*, **63**, 483-492.
- De Vries, P. (1979) Line intersect sampling—statistical theory, applications and suggestions for extended use in ecological inventory, in *Sampling Biological Populations*, edited by R. Cormack, G. Patil and D. Robson. Fairland, Md.: International Cooperative Publishing House, pp. 1-70.

- Diggle, P. (1975) Robust density estimation using distance methods. *Biometrika*, **62**, 39-48.
- Diggle, P. (1977a) A note on robust density estimation for spatial point patterns. *Biometrika*, **64**, 91-95.
- Diggle, P. (1977b) The detection of random heterogeneity in plant populations. *Biometrics*, **33**, 390-394.
- Diggle, P. (1978) On parameter estimation for spatial point processes. *Journal of the Royal Statistical Society*, **40B**, 178-181.
- Diggle, P. (1979a) On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrics*, **35**, 87-101.
- Diggle, P. (1979b) Statistical methods for spatial point patterns in ecology, in *Spatial and Temporal Analysis in Ecology*, edited by R. Cormack and J. Ord. Fairland, Md.: International Cooperative Publishing House, pp. 95-150.
- Diggle, P. (1981a) Some graphical methods in the analysis of spatial point patterns, in *Interpreting Multivariate Data*, edited by V. Barnett. Chichester: Wiley, pp. 55-73.
- Diggle, P. (1981b) Binary mosaics and the spatial pattern of heather. *Biometrics*, **37**, 531-539.
- Diggle, P. (1983) *Statistical Analysis of Spatial Pattern*. New York: Academic Press.
- Diggle, P., and T. Cox. (1981) On sparse sampling methods and tests of independence for multivariate spatial point patterns. *Bulletin of the International Statistical Institute*, **49**, 211-229.
- Diggle, P., and T. Cox. (1983) Some distance based tests of independence for sparsely-sampled multivariate spatial point patterns. *International Statistical Review*, **51**, 11-23.
- Diggle, P., D. Gates, and A. Stibbard. (1987) A non-parametric estimator for pairwise interaction point processes. *Biometrika*, **74**, 763-770.
- Diggle, P., and R. Gratton. (1984) Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society*, **46B**, 193-227.
- Diggle, P., and B. Matern. (1981) On sampling designs for the estimation of point-event nearest neighbor distributions. *Scandinavian Journal of Statistics*, **7**, 80-84.
- Diggle, P., and R. Milne. (1983a) Negative binomial quadrat counts and point processes. *Scandinavian Journal of Statistics*, **10**, 257-267.
- Diggle, P., and R. Milne. (1983b) Bivariate Cox processes: some models for bivariate spatial point patterns. *Journal of the Royal Statistical Society*, **45B**, 11-21.
- Eberhardt, L. (1967) Some developments in "distance" sampling. *Biometrics*, **23**, 207-216.
- Gates, C. (1979) Line transect and related issues, in *Sampling Biological Populations*, edited by R. Cormack, G. Patil and D. Robson. Fairland, Md.: International Cooperative Publishing House, pp. 71-154.
- Gates, D., and M. Westcott. (1986) Clustering estimates for spatial point distributions with unstable potentials. *Annals of the Institute of Statistical Mathematics*, **38**, 123-135.
- Getis, A., and B. Boots. (1978) *Models of Spatial Processes*. Cambridge: Cambridge University Press.
- Getis, A., and J. Franklin. (1987) Second order neighborhood analysis of mapped point

- patterns. *Ecology*, **68**, 473-477.
- Goodall, D. (1965) Plot-less tests of interspecific association. *Journal of Ecology*, **53**, 197-210.
- Goodall, D. (1974) A new method for the analysis of spatial pattern by random pairing of quadrats. *Vegetatio*, **29**, 135-146.
- Green, P., and R. Sibson. (1978) Computing Dirichlet tessellations in the plane. *Computing Journal*, **21**, 168-173.
- Greig-Smith, P. (1952) The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany*, **NS16**, 293-316.
- Greig-Smith, P. (1979) Pattern in vegetation. *Journal of Ecology*, **67**, 755-779.
- Hall, P. (1985) Counting methods for inference in binary mosaics. *Biometrics*, **41**, 1049-1052.
- Hanisch, K. (1983) Reduction of the n th moment measure and the special case of the third moment measure of stationary and isotropic point-processes. *Mathematische Operationsforschung und Statistik*, **14**, 421-435.
- Harkness, R., and V. Isham. (1983) A bivariate spatial point pattern of ants' nests. *Applied Statistics*, **32**, 293-303.
- Heltsh, J., and T. Richey. (1984) Spatial pattern detection using quadrat samples. *Biometrics*, **40**, 877-885.
- Hill, M. (1973) The intensity of spatial pattern in plant communities. *Journal of Ecology*, **61**, 225-235.
- Hines, W., and R. Hines. (1979) The Eberhardt index and the detection of non-randomness of spatial point distributions. *Biometrika*, **66**, 73-79.
- Holgate, P. (1965) Tests of randomness based on distance methods. *Biometrika*, **52**, 345-353.
- Hopkins, B. (1954) A new method of determining the type of distribution of plant individuals. *Annals of Botany*, **18**, 213-226.
- Isham, V. (1984) Bivariate Markov point processes—some approximations. *Proceedings of the Royal Society*, **391A**, 39-53.
- Jolivet, E. (1978) Caractérisation et test du caractère agrégatif des processus ponctuels stationnaires sur R^2 , in *Lecture Notes in Mathematics*, #636. Berlin: Springer-Verlag, pp. 1-25.
- Kelly, F., and B. Ripley. (1976) A note on Strauss' model for clustering. *Biometrika*, **63**, 357-360.
- Kemp, C., and H. Papageorgiou. (1982) Bivariate Hermite distributions. *Sankhya*, **44A**, 269-280.
- Kinderman, R., and J. Snell. (1980) *Markov Random Fields and Their Applications*. Providence: American Mathematical Society.
- Kingman, J. (1977) Remarks on the spatial distribution of a reproducing population. *Journal of Applied Probability*, **14**, 577-583.
- Kooijman, S. (1979) The description of point patterns, in *Spatial and Temporal Analysis in Ecology*, edited by R. Cormack and J. Ord. Fairland, Md.: International Cooperative Publishing House, pp. 305-332.

- Lewis, P., and G. Shedler. (1979) Simulation of non-homogeneous Poisson processes by thinning. *Naval Research Logistics Quarterly*, **26**, 403-413.
- Lewis, S. (1975) Robust estimation of density for a two-dimensional point process. *Biometrika*, **62**, 519-521.
- Liebetrau, A. (1977) Tests of randomness in two dimensions. *Communications in Statistics: Theory and Methods*, **6A**, 1367-1383.
- Lotwick, H. (1981) Simulation of some spatial hard core models, and the complete packing problem. *Journal of Statistical Computation and Simulation*, **15**, 293-314.
- Lotwick, H. (1984) Some models for multitype spatial point processes, with remarks on analyzing multitype patterns. *Journal of Applied Probability*, **21**, 575-582.
- Lotwick, H., and B. Silverman. (1982) Methods for analyzing spatial processes of several types of points. *Journal of the Royal Statistical Society*, **44B**, 406-413.
- Matern, B. (1960) *Spatial Variation*. Stockholm: Meddelanden Statens fran Skogsforskningsinstitut, Vol. 49 pp. 1-144.
- Matern, B. (1971) Doubly stochastic Poisson processes in the plane, in *Statistical Ecology*, Vol. 1, edited by G. Patil, E. Pielou and W. Waters. University Park, Pa.: The Pennsylvania State University Press, pp. 195-213.
- Mead, R. (1974) A test for spatial pattern at several scales using data from a grid of contiguous quadrats. *Biometrics*, **30**, 295-307.
- Moellering, H., and W. Tobler. (1972) Geographical variances. *Geographical Analysis*, **4**, 34-50.
- Mountford, M. (1961) On E. C. Pielou's index of non-randomness. *Journal of Ecology*, **49**, 271-275.
- Naus, J. (1979) An indexed bibliography of clusters, clumps and coincidences. *Review of the International Statistical Institute*, **47**, 47-78.
- Newman, C., Y. Rinott, and A. Tversky. (1983) Nearest neighbors and Voronoi regions in certain point processes. *Advances in Applied Probability*, **15**, 726-751.
- Ogata, Y., and M. Tanemura. (1981) Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, **33B**, 315-338.
- Ogata, Y., and M. Tanemura. (1984) Likelihood analysis of spatial point patterns. *Journal of the Royal Statistical Society*, **46B**, 496-518.
- Ogata, Y., and M. Tanemura. (1985) Estimation of interaction potentials of marked spatial patterns through the maximum likelihood method. *Biometrics*, **41**, 421-433.
- Ohser, J., and D. Stoyan. (1981) On the second-order and orientation analysis of planar stationary processes. *Biometrical Journal*, **23**, 523-533.
- Ord, J. (1970) The negative binomial model and quadrat sampling, in *Random Counts in Models and Structures*, edited by G. Patil. University Park, Pa.: The Pennsylvania State University Press, pp. 151-163.
- Ord, J. (1972) *Families of Frequency Distributions*. London: Griffin.
- Ord, J. (1978) How many trees in a forest? *Mathematical Scientist*, **3**, 23-33.

- Patil, S., K. Burnham, and J. Kovner. (1979) Non-parametric estimation of plant density by the distance method. *Biometrics*, **35**, 597-604.
- Patil, S., J. Kovner, and K. Burnham. (1982) Optimum nonparametric estimation of population density based on ordered distances. *Biometrics*, **38**, 243-248.
- Perry, J., and R. Mead. (1979) On the power of the index of dispersion test to detect spatial pattern. *Biometrics*, **35**, 613-622.
- Persson, O. (1971) The robustness of estimating density by distance measurements, in *Statistical Ecology*, Vol. 2, edited by G. Patil, E. Pielou and W. Waters. University Park, Pa.: The Pennsylvania State University Press, pp. 175-190.
- Pickard, D. (1982) Isolated nearest neighbors. *Journal of Applied Probability*, **19**, 444-449.
- Pielou, E. (1959) The use of point-to-plant distances in the study of the pattern of plant populations. *Journal of Ecology*, **47**, 607-613.
- Pielou, E. (1977) *Mathematical Ecology*. New York: Wiley.
- Pollard, J. (1971) On distance estimators of density in randomly distributed forests. *Biometrics*, **27**, 991-1002.
- Renshaw, E. (1984) Competition experiments for light in a plant monoculture: an analysis based on two-dimensional spectra. *Biometrics*, **40**, 717-728.
- Renshaw, E. (1986) A survey of stepping stone models in population dynamics. *Advances in Applied Probability*, **18**, 581-627.
- Renshaw, E., and E. Ford. (1983) The interpretation of process from pattern using two dimensional spectral analysis: methods and problems of interpretation. *Applied Statistics*, **32**, 51-63.
- Ripley, B. (1976a) The second order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255-266.
- Ripley, B. (1976b) Locally finite random sets: Foundations for point process theory. *Annals of Probability*, **4**, 983-994.
- Ripley, B. (1977) Modeling spatial patterns. *Journal of the Royal Statistical Society*, **39B**, 172-212.
- Ripley, B. (1978) Spectral analysis and the analysis of pattern in plant communities. *Journal of Ecology*, **66**, 965-981.
- Ripley, B. (1979) Tests of "randomness" for spatial point patterns. *Journal of the Royal Statistical Society*, **41B**, 368-374.
- Ripley, B. (1981) *Spatial Statistics*. Chichester: Wiley.
- Ripley, B. (1984a) Spatial Statistics: Developments 1980-3. *International Statistical Review*, **52**, 141-150.
- Ripley, B. (1984b) Analyses of nest spacing, in *Statistics in Ornithology*, edited by B. Morgan and P. North. Berlin: Springer-Verlag, Lecture Notes in Statistics, pp. 151-158.
- Ripley, B. (1984c) Edge corrections for spatial processes, in *Stochastic Geometry, Geometrical Statistics, Stereology*, edited by R. Ambartzumian and W. Weil. Leipzig: Teubner Texte, pp. 144-153.
- Ripley, B. (1986) Statistics, images and pattern recognition. *Canadian Journal of Statistics*,

14, 83-102.

- Ripley, B. (1987) *Stochastic Simulation*. New York: Wiley.
- Ripley, B. (1988) *Statistical Inference for Spatial Processes*. New York: Cambridge University Press.
- Ripley, B., and F. Kelly. (1977) Markov point processes. *Journal of the London Mathematical Society*, **15**, 188-192.
- Ripley, B., and B. Silverman. (1978) Quick tests for spatial regularity. *Biometrika*, **65**, 641-642.
- Roach, S. (1968) *The Theory of Random Clumping*. London: Methuen.
- Sager, T. (1982) Nonparametric maximum likelihood estimation of spatial patterns. *Annals of Statistics*, **10**, 1125-1136.
- Saunders, R., and G. Funk. (1977) Poisson limits for a clustering model of Strauss. *Journal of Applied Probability*, **14**, 776-84.
- Saunders, R., R. Kryscio, and G. Funk. (1982) Poisson limits for a hardcore clustering model. *Stochastic Processes and Their Applications*, **12**, 97-106.
- Shapiro, M., J. Schein, and F. De Monasterio. (1985) Regularity and structure of the spatial pattern of blue cones of macaque retina. *Journal of the American Statistical Association*, **80**, 803-812 (with discussion).
- Silverman, B. (1978) Distances on circles, toruses and spheres. *Journal of Applied Probability*, **15**, 136-143.
- Silverman, B. (1981) Density estimation for univariate and bivariate data, in *Interpreting Multivariate Data*, edited by V. Barnett. Chichester: Wiley, pp. 37-53.
- Silverman, B., and T. Brown. (1978) Short distances, flat triangles and Poisson limits. *Journal of Applied Probability*, **15**, 815-825.
- Skellam, J. (1951) Random dispersal in theoretical populations. *Biometrika*, **38**, 196-218.
- Strauss, D. (1975) A model for clustering. *Biometrika*, **62**, 467-475.
- Upton, G. (1984) On Mead's test for pattern. *Biometrics*, **40**, 759-766.
- Upton, G., and B. Fingleton. (1985) *Spatial Data Analysis By Example*, Vol. 1. Chichester: Wiley.
- Warren, W. (1971) The center-satellite concept as a basis for ecological sampling, in *Statistical Ecology*, Vol. 2, edited by G. Patil, E. Pielou and W. Waters. University Park, Pa: The Pennsylvania State University Press, pp. 87-116.
- Warren, W., and C. Batcheler. (1979) The density of spatial patterns: robust estimation through distance methods, in *Spatial and Temporal Analysis in Ecology*, edited by R. Cormack and J. Ord. Fairland, Md.: International Cooperative Publishing House, pp. 247-270.
- Westcott, M. (1982) Approximations to hard-core models and their application to statistical analysis. *Journal of Applied Probability*, **19**, 281-292.
- Zahl, S. (1977) A comparison of three methods for the analysis of spatial pattern. *Biometrics*, **33**, 681-692.

DISCUSSION

"Statistical methods for point pattern data"

by J. Keith Ord

The field of spatial point processes encompasses work in both probability and statistics as well as in several applied fields, and readers should note that Ord's review is much more restricted in scope than its title might suggest. Rather than comment on the details of the paper, I will attempt to complement it with other developments that I see as important. Almost all the issues addressed are covered in some detail in texts such as Diggle (1983), Ripley (1981) and Upton and Fingleton (1985), but many of the topics below are not.

Theory.

It is easy to overlook the importance of theory: major practical developments such as second-moment methods (\hat{K}) and Gibbs processes arose from theoretical ideas. Major theoretical contributions have been made by Kallenberg (1975, 1983), the East German school (Matthes *et al.*, 1978; Stoyan *et al.*, 1987), and Papangelou for conditional intensities (see Kallenberg, 1983). Apart from these monographs there are also expositions by Karr (1985) and Daley and Vere-Jones (1988). The area of Gibbs processes is on the interface with the axiomatization of statistical physics by Preston, Georgii and colleagues.

One theoretical contribution has been to find minimal conditions that define a Poisson process. Rather surprisingly, it suffices to know

$$\begin{aligned} P(N(A) = 0) & \text{ for a sufficiently wide class of sets } A \subset C \\ P(N(\{\mathbf{x}\}) > 1) & \text{ for any } \mathbf{x} = 0 \end{aligned}$$

to specify any point process that is *simple* (the second condition). In particular, a process with

$$N(A) \text{ Poisson mean } \Lambda(A), \quad \Lambda(\{\mathbf{x}\}) = 0 \text{ for all } \mathbf{x}$$

is a Poisson process (a combination of results of Renyi, Kallenberg and myself in the period 1971-6).

Point processes of objects.

One of the major goals of spatial statistics has been the ability to handle more complicated image data than points. *Stochastic geometry* reduces objects to points in other spaces, and much work has been done on 'fibre processes' and their cousins (Stoyan *et al.*, 1987). These can represent lines and curves such as rivers, and Stoyan and Ohser (1982) have studied the interactions between points (trees), fibres (rivers) and areas (soil types) on Dresden Heath.

I would argue that point patterns *per se* are becoming less important. They are almost always approximations to the truth, representing objects with size and shape, and as we become able to cope with the image data we actually observe, the underlying point pattern descriptions will become important rather than direct inference from point patterns.

Statistical inference.

Much more has been done on inference from point patterns recently than appears in this review. Many of the issues are discussed in my own contribution to this volume and in Ripley (1988), but let me underline the importance of pseudo-likelihood and simulation methods, as well as Palm probabilities. The latter correspond to comparing the views of a point pattern from an arbitrary point with those from an 'individual' (in Ord's notation), and this has been exploited by Takacs (1986) and others subsequently.

I do not share Ord's pessimism (first paragraph) on computational intractability. The phrase seems to reflect our ignorance of what to do rather than fundamental problems of computational complexity. Over the last decade the picture has become much more promising due to all of

- (a) developments in computational geometry (Preparata and Shamos, 1985),
- (b) new ideas on approximations, and
- (c) cheap desktop computing power, equal to that of a mainframe of a decade ago.

Unfortunately these subjects are not part of the traditional education of either statisticians or geographers. Indeed, the computing skills needed for modern methods may exclude many potential users, and one day we may see the MINITAB or SAS of spatial statistics.

Mosaic models.

Diggle (1981b) used distance methods to fit a point process model of objects to his heather data, and very successfully demonstrated how methods of low power allow one to fit inappropriate models! The inadequacies of his model are obvious to most people immediately on comparing images of the data with those of simulations of the model. Finding summary measures that are as good as 'eye-balling' took some time, and depended on measuring notions of shape as well as size (Ripley, 1986, 1988). Hall's (1985) methods suffer from similar difficulties.

Hall (1988) gives a comprehensive introduction to a family of *coverage processes* for mosaics. Other models are described by Ahuja and Schachter (1983).

Spatial point processes have had a golden few years, and the methods developed then are beginning to be widely used. It is the case that point patterns are now usually analysed as maps, and the K-function methods (specifically, plotting $L(t) = \sqrt{\hat{K}(t)/t}$ vs t) have become standard in a number of biological fields. Nevertheless, currently point processes are being eclipsed by statistical image analysis, and I suspect developments in point processes in the near future will be less prolific.

References

- Ahuja, N., and B. J. Schachter. (1983) *Pattern Models*. New York: Wiley.
- Daley, D., and D. Vere-Jones. (1988) *An Introduction to the Theory of Point Processes*. New York: Springer-Verlag.
- Hall, P. (1988) *An Introduction to the Theory of Coverage Processes*. New York: Wiley.
- Kallenberg, O. (1975) *Random Measures*. Berlin: Akademie/Academic.
- Kallenberg, O. (1983) *Random Measures*, 3rd ed. London: Academic Press.

- Karr, A. (1985) *Point Processes and their Statistical Inference*. Marcel-Dekker, New York.
- Matthes, K., J. Kerstan, and J. Mecke. (1978) *Infinitely Divisible Point Processes*. Chichester: Wiley.
- Preparata, F. D. and M. L. Shamos. (1985) *Computational Geometry*. Springer, New York.
- Ripley, B. D. (1986) Statistics, images and pattern recognition. *Canadian Journal of Statistics*, 14, 83-112.
- Ripley, B. D. (1988) *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.
- Stoyan, D., and J. Ohser. (1982) Correlations between planar random structures, with an ecological application. *Biometrical Journal*, 24, 631-647.
- Stoyan, D., W. S. Kendall, and J. Mecke. (1987) *Stochastic Geometry and its Applications*. Chichester: Wiley.
- Takacs, R. (1986) Estimator for the pair potential of a Gibbsian point process. *Statistics*, 17, 429-433.

Brian D. Ripley, University of Strathclyde.

A REJOINDER TO RIPLEY'S DISCUSSION

by J. Keith Ord

In some ways my reply must include a defense of Ripley—past from the comments of Ripley—present. But, overall, I do not believe our viewpoints are so very different.

It certainly was not my intention to underrate the role of theory. Rather, my review focuses more on methodology, as the title implies. I am grateful to Dr. Ripley for including the additional references to more theoretical literature.

I would agree entirely, for the case of image processing, that shapes are more important than points. However, there are many areas of application, especially in human geography, where point processes are, and will remain, of central interest.

My opening paragraph was not intended to be pessimistic. To the contrary, it is a comment on the history of spatial modeling. I agree wholeheartedly that the present need is for accessible software.

In terms of stochastic modeling, I suspect it is true that the golden age has passed for spatial point processes. However, the further development of inferential tools and diagnostic procedures remains a substantive challenge for the future. Only by such progress can innovative applications be assisted.

