

16

Spatial Point Process Theory

Marie-Colette van Lieshout

CONTENTS

16.1	Spatial Point Process Theory	263
16.1.1	Introduction	263
16.1.2	Characterization of Point Process Distributions.....	266
16.1.3	Campbell and Moment Measures.....	267
16.1.4	Reduced and Higher-Order Campbell Measures.....	270
16.1.5	Palm Theory and Conditioning	273
16.1.6	Finite Point Processes	278
16.1.7	Gibbs Measures by Local Specification.....	280
	References.....	281

16.1 Spatial Point Process Theory

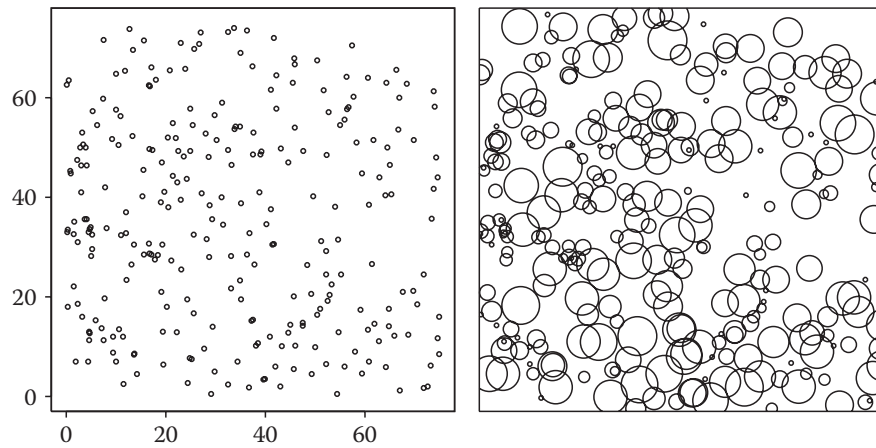
16.1.1 Introduction

A *spatial point process* is a stochastic process each of whose realizations consists of a finite or countably infinite set of points in the plane. This chapter sets out the mathematical theory of such processes. After some motivation, several ways to characterize the distribution of a point process are described in [Section 16.1.2](#). The important notions of Campbell and moment measures are introduced in [Section 16.1.3](#), and generalized to higher orders in [Section 16.1.4](#). Interior and exterior conditioning by means of Palm theory form the topic of [Section 16.1.5](#). The section closes with a discussion of finite point processes and local specifications ([Sections 16.1.6 to 16.1.7](#)).

Tracing the roots of modern point process theory is not an easy task. One may refer to Poisson (1837), or to the pioneering works by Erlang (1909) and Neyman (1939) in the respective contexts of telephone networks and spatial cluster processes. Influential classic textbooks include Janossy (1948), Khinchin (1960), Matheron (1975), Matthes, Kerstan, and Mecke (1978), and Srinivasan (1974). A rich scholarly reference book is the second edition of Daley and Vere-Jones (1988) published in two volumes (2003, 2008). A more accessible, yet rigorous, course on finite point processes is van Lieshout (2000). See also Cox and Isham (1980) or Reiss (1993).

As a motivating example, consider the pattern of 270 tree locations in a 75×75 meter plot in the Kaluzhskie Zaseki Forest in central Russia depicted in the left-hand panel of [Figure 16.1](#). The data were kindly provided by Dr. P. Grabarnik and previously analyzed in Smirnova (1994) and Grabarnik and Chiu (2002).

In a forestry context such as this, competition for nutrients and space may result in patterns in which large trees do not occur close to each other and smaller trees fill up the remaining gaps. On the other hand, seedlings tend to grow up close to mature trees, which

**FIGURE 16.1**

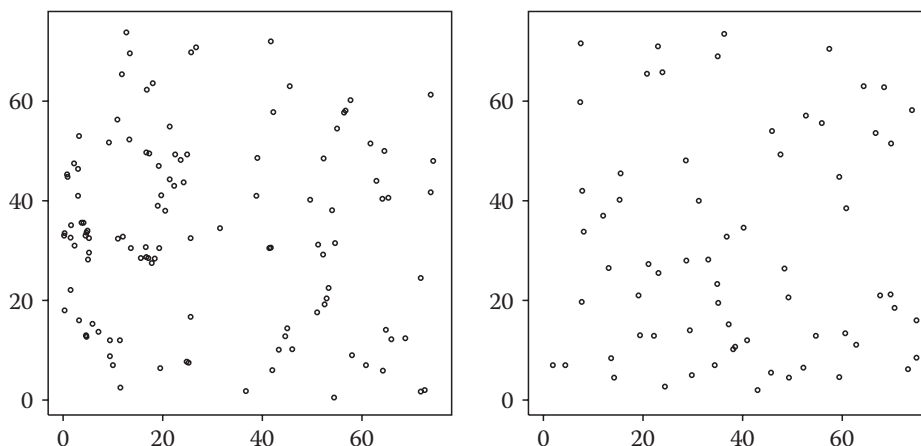
Trees in a broad-leaved, multispecies old-growth forest in the Kaluzhskie Zaseki Forest in central Russia (Smirnova, 1994; Grabarnik and Chiu, 2002) in a square plot of side length 75 m. (Left) Locations of the 270 trees, (right) graphical representation in which each tree is located at the center of a disk of radius equal to 0.103 times the tree diameter.

would result in a cluster of nearby trees. Such apparent clustering can also result from variations in soil fertility, as trees are more likely to flourish in regions of high fertility.

For these data, additional measurements in the form of tree diameters are available, graphically represented by the disk radii in the right-hand side panel of Figure 16.1. Such a pattern with extra information attached to each point (location) is referred to as a “marked point pattern.” Clusters of small trees in the gaps left by larger trees are apparent. This clustering effect can be seen even more clearly if we plot the subpatterns of small and large trees separately (Figure 16.2). The large (old) trees seem, apart from some gaps, to have a preferred distance of about 4 meters from each other, possibly due to a planting design.

Mapped data such as those depicted in Figure 16.1 and Figure 16.2 can be described mathematically as a finite, integer-valued measure

$$\sum_{i=1}^I k_i \delta_{x_i}$$

**FIGURE 16.2**

Locations of trees in a broad-leaved, multispecies old-growth forest (Smirnova, 1994; Grabarnik and Chiu, 2002) in a square plot of side length 75 m. (Left) Trees with diameter less than 15 cm, (right) those with diameter larger than 25 cm.

where $I \in \mathbb{N}_0$, the \mathbf{x}_i are distinct (marked) points in some topological space D , and each weight k_i is a strictly positive integer, the *multiplicity* of \mathbf{x}_i . An equivalent description is in terms of a finite, unordered set of not necessarily distinct (marked) points

$$\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \in D^n / \equiv, \quad n = 0, 1, \dots$$

The \mathbf{x}_i lie in a topological space D , and two or more $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in D^n$ that are identical up to permutation are indistinguishable. We shall refer to \mathcal{X} as a *configuration*.

For example, for the unmarked pattern depicted in the left-hand panel of Figure 16.1, $D = [0, 75]^2$ equipped with the Euclidean distance topology, while marked points represented graphically in the right-hand panel of Figure 16.1 lie in the product space $[0, 75]^2 \times [0, \infty)$ generated by the Euclidean topologies on $[0, 75]^2$ and $[0, \infty)$. More precisely, if U and V are open neighborhoods of $\mathbf{x} \in [0, 75]^2$ and $m \in [0, \infty)$, respectively, then we may define an open neighborhood of (\mathbf{x}, m) by $U \times V$. The class of such product sets defines a topology, the *product topology*. Equivalently, the product topology can be defined by the supremum metric. In order to obtain Figure 16.2, trees were marked by a type label in the set {large, small} equipped with the discrete metric.

If D is not bounded, the requirement that I or n be finite is relaxed to allow configurations that contain a countable number of points, each with finite multiplicity. Such a configuration $\mathcal{X} \subseteq D$ is said to be *locally finite* if it places at most a finite number of points in any bounded Borel set $A \subseteq D$, that is, if

$$N_{\mathcal{X}}(A) = \sum_{i=1}^{\infty} k_i \mathbb{1}_{\partial_{\mathbf{x}_i}}(A) < \infty.$$

The family of all locally finite configurations will be denoted by N^{lf} .

Definition 16.1

Let (D, d) be a complete, separable metric space equipped with its Borel σ -algebra \mathcal{B} . A point process on D is a mapping X from a probability space (Ω, \mathcal{A}, P) into N^{lf} such that for all bounded Borel sets $A \subseteq D$, the number $N(A) = N_X(A)$ of points falling in A is a (finite) random variable.

From a practical point of view, it is convenient to work with metric spaces. Theoretically, the topological structure is needed to ensure the existence of regular versions of Palm distributions and related characteristics that will be discussed later.

Definition 16.1 may be rephrased as follows. A point process X is a random variable with values in the measurable space $(N^{\text{lf}}, \mathcal{N}^{\text{lf}})$, where \mathcal{N}^{lf} is the smallest σ -algebra such that for all bounded Borel sets $A \subseteq D$, the mapping $\mathcal{X} \mapsto N_{\mathcal{X}}(A)$ is measurable. The induced probability measure \mathcal{P} on \mathcal{N}^{lf} is called the *distribution* of the point process.

When the $\mathbf{x}_i \in \mathcal{X}$ have a location component, i.e., when D is of the form $\mathbb{R}^d \times L$ for some mark space L , assumed to be a complete, separable metric space, and the marginal point process of locations is well-defined, X is said to be a *marked point process*. If \mathcal{P} is translation invariant in the sense that it does not change if all marked points $\mathbf{x}_i = (\mathbf{a}_i, m_i) \in X$ are translated over some vector $\mathbf{y} \in \mathbb{R}^d$ into $(\mathbf{a}_i + \mathbf{y}, m_i)$, X is *stationary*. If, additionally, \mathcal{P} is not affected by rotations of the location component, X is *isotropic*.

Example 16.1

Let D be a compact subset of the plane of strictly positive area $|D|$. A *binomial point process* is defined as the union $X = \{X_1, \dots, X_n\}$ of a fixed number $n \in \mathbb{N}$ of independent, uniformly distributed points X_1, \dots, X_n in D . In other words, $\mathbf{P}(X_i \in A) = |A|/|D|$ for all Borel subsets $A \subseteq D$. Now,

$$N(A) = \sum_{i=1}^n \mathbb{1}\{X_i \in A\}$$

is a sum of random variables; hence, a random variable itself, and, as $N(A) \leq N(D) = n$, X takes values in \mathbb{N}^{lf} .

Example 16.2

Further to the previous example, by replacing the fixed number $n \in \mathbb{N}$ by an integer-valued, random variable following a Poisson distribution with parameter $\lambda|D|$ proportional to the area $|D|$ of D , we obtain a *homogeneous Poisson process* on D . Now,

$$\mathbf{P}(N(A) = k) = \sum_{n=k}^{\infty} e^{-\lambda|D|} \frac{(\lambda|D|)^n}{n!} \binom{n}{k} \left(\frac{|A|}{|D|}\right)^k \left(1 - \frac{|A|}{|D|}\right)^{n-k} = e^{-\lambda|A|} \frac{(\lambda|A|)^k}{k!},$$

for all $k \in \mathbb{N}_0$, so that $N(A)$ is Poisson distributed with parameter $\lambda|A|$, and, by similar arguments, for disjoint Borel sets $A, B \subseteq D$, and $k, l \in \mathbb{N}_0$,

$$\mathbf{P}(N(A) = k; N(B) = l) = \mathbf{P}(N(A) = k) \mathbf{P}(N(B) = l).$$

Hence, the random variables $N(A)$ and $N(B)$ are independent. It should be noted that the definition of a Poisson process may be extended to the plane by tiling \mathbb{R}^2 and defining independent Poisson processes on each tile as before.

16.1.2 Characterization of Point Process Distributions

The distribution of a real-valued random variable may be characterized by its distribution function, characteristic function, moment generating function, or Laplace transform. Recall that the distribution \mathcal{P} of a point process is that induced by the integer-valued random variables $N(A)$ counting the number of objects placed in bounded Borel sets A . Thus, it is natural to characterize \mathcal{P} by properties of the distributions of the random variables $N(A)$.

Definition 16.2

The family of finite-dimensional distributions (*fidis*) of a point process X on a complete, separable metric space (D, d) is the collection of the joint distributions of $(N(A_1), \dots, N(A_m))$, where (A_1, \dots, A_m) ranges over the bounded Borel sets $A_i \subseteq D, i = 1, \dots, m$, and $m \in \mathbb{N}$.

The following uniqueness theorem holds.

Theorem 16.1

The distribution of a point process X on a complete, separable metric space (D, d) is completely specified by its finite-dimensional distributions.

In other words, if two point processes have identical fidis, they also share the same distribution. The result follows from the observation that the family of sets

$$\{\omega \in \Omega : N_{X(\omega)}(A_i) \in B_i, \quad i = 1, \dots, m\},$$

where $A_i \subseteq D$ are bounded Borel sets and $B_i \subseteq \mathbb{R}$ are Borel, is a semiring generating \mathcal{N}^{lf} .

A point processes on (D, d) is called *simple* if its realizations contain no points with multiplicity $k_i \geq 2$, so that $N(\{\mathbf{x}\}) \in \{0, 1\}$ almost surely for all $\mathbf{x} \in D$. Surprisingly, as shown by Rényi (1967) for the Poisson process and by Móncch (1971) more generally, if two simple point processes X and Y assign the same *void probabilities*, $v(A) = \mathcal{P}(N(A) = 0)$, to all bounded Borel sets A , their distributions coincide.

Theorem 16.2

The distribution of a simple point process X on a complete, separable metric space (D, d) is uniquely determined by the void probabilities of bounded Borel sets $A \subseteq D$.

In the above theorem, the collection of bounded Borel sets may be replaced by a smaller class, such as the compact sets. For details, we refer to Berg, Christensen, and Ressel (1984), Choquet (1953), Daley and Vere-Jones (1988), Matheron (1975), McMillan (1953), and Norberg (1989).

Example 16.3

The binomial point process introduced in Example 16.1 is simple and has void probabilities $v(A) = (1 - |A|/|D|)^n$, $A \in \mathcal{B}(D)$. For the homogeneous Poisson process considered in Example 16.2, $v(A) = \exp[-\lambda|A|]$.

For any point process X , one may define a simple point process X^s by ignoring the multiplicity. It follows from Theorem 16.2 that if point processes X and Y have identical void probabilities, the distributions of X^s and Y^s must be the same. Thus, X and Y exhibit the same interaction structure except for the multiplicity of their points.

Trivial cases apart, it is not possible to plot the void probabilities as a function of $A \subseteq D$. Nevertheless, a graphical representation may often be achieved. Indeed, suppose that the distribution \mathcal{P} that produced a Euclidean point pattern, such as that depicted in Figure 16.1, is invariant under translations. Then $v(A) = v(T_x A)$ for all $T_x A = \{\mathbf{a} + \mathbf{x} : \mathbf{a} \in A\}$, $\mathbf{x} \in \mathbb{R}^d$. In particular, if $A = \bar{b}(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{R}^d : \|\mathbf{y} - \mathbf{x}\| \leq r\}$ is a closed ball, the stationarity of \mathcal{P} implies that $v(\bar{b}(\mathbf{x}, r))$ is a function of r only. Based on this observation, the *empty space function* is defined as

$$F(r) = \mathcal{P}(d(\mathbf{x}, X) \leq r) = 1 - v(\bar{b}(\mathbf{x}, r)), \quad r \geq 0, \quad (16.1)$$

that is, F is the distribution function of the distance from an arbitrary point $\mathbf{x} \in \mathbb{R}^d$ to the nearest point of X . Although \mathcal{P} is not uniquely specified by the void probability it assigns to balls, plots of the estimated empty space function do provide valuable information on the interaction structure.

Example 16.4

Estimated empty space functions related to the Kaluzhskie Zaseki Forest data of Figure 16.2 are given in Figure 16.3. On biological grounds, it may be expected that the patterns formed by young and established trees are qualitatively different, and this is confirmed by Figure 16.3. Using the observed size mark as a substitute for the nonrecorded age of trees, estimated F -functions for both types of trees are shown and compared to the empty space function of a Poisson process with the same expected number of points (cf. Example 16.3). For small trees, large empty spaces occur with a higher probability than under the Poisson model. Such a behavior is typical for clustered patterns with groups of nearby points separated by gaps. The graph for larger trees is closer to that of its Poisson counterpart.

16.1.3 Campbell and Moment Measures

In the previous section, we saw that the distribution of a simple point process is described fully by its void probabilities and that a graphical representation is available in the empty space function. For point processes that may contain multiple points, the finite-dimensional distributions of Definition 16.2 can be used to specify the distribution, and their moments provide a useful complementary collection of summary statistics (see, e.g., Matthes et al., 1978).

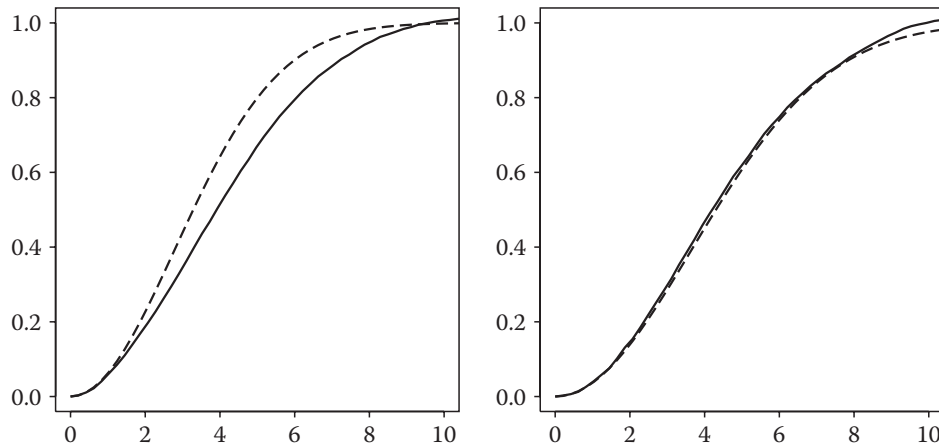


FIGURE 16.3

Estimated empty space functions of Kaluzhskie Zaseki data (solid lines). (Left) Locations of trees with diameter less than 15 cm, (right) those with diameter larger than 25 cm. The dashed lines are the empty space functions for homogeneous Poisson processes with, on average, the same number of points per unit area as the data.

Definition 16.3

Let X be a point process on a complete, separable metric space (D, d) . Define a set function M by $M(A) = \mathbf{E}N(A)$ for all bounded Borel sets $A \subseteq D$.

Example 16.5

Consider the binomial point process of Example 16.1. Then,

$$M(A) = n \mathbf{P}(X_1 \in A) = n |A|/|D|$$

for any Borel set $A \subseteq D$. In particular, $M(D) = n < \infty$. For the homogeneous Poisson process of Example 16.2, $M(A) = \lambda |A|$.

By standard measure theoretic arguments (see, e.g., Daley and Vere-Jones (1988) or van Lieshout (2000)), the set function M can be extended to a Borel measure.

Theorem 16.3

If the function M introduced in Definition 16.3 is finite for all bounded Borel sets, it can be extended uniquely to a σ -finite measure on the Borel sets of D , the (first-order) moment measure.

Let us consider a stationary point process X on the Euclidean space \mathbb{R}^d , and denote by $T_{\mathbf{x}}A = \{\mathbf{a} + \mathbf{x} : \mathbf{a} \in A\}$ the translation of A over the vector $\mathbf{x} \in \mathbb{R}^d$. Then, for any Borel set A ,

$$M(A) = \mathbf{E}N(A) = \mathbf{E}N(T_{\mathbf{x}}A) = M(T_{\mathbf{x}}A).$$

Consequently, provided M is finite for bounded A , a constant $0 \leq \lambda < \infty$ can be found such that $M(A) = \lambda |A|$. The scalar constant λ is called the *intensity* of X . An example of a stationary point process is the homogeneous Poisson process on the plane, (cf. Example 16.2).

More generally, if M is absolutely continuous with respect to Lebesgue measure, then

$$M(A) = \int_A \lambda(\mathbf{x}) d\mathbf{x}$$

for Borel sets A , and $\lambda(\cdot)$ is referred to as an *intensity function* of the point process X .

In integral terms, Theorem 16.3 asserts that

$$\int_A dM(\mathbf{a}) = \mathbf{E} \left[\sum_{\mathbf{x} \in X} 1\{\mathbf{x} \in A\} \right]$$

for all $A \in \mathcal{B}(D)$. Linearity and monotonicity arguments imply that

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}) \right] = \int_D g(\mathbf{x}) dM(\mathbf{x}) \tag{16.2}$$

for any measurable function $g : D \rightarrow \mathbb{R}$ that is either nonnegative or integrable with respect to M , provided the moment measure M exists, i.e., is finite on bounded Borel sets. Identities of the form Equation (16.2) are usually referred to as *Campbell theorems* in honor of Campbell’s (1909) paper. Intuitively, suppose one takes a measurement $g(\mathbf{x})$ at each \mathbf{x} in some data configuration \mathcal{X} . Then the grand total of these measurements is equal in expectation to the spatial integral of g with respect to the moment measure of the stochastic process that generated the data.

Example 16.6

Let us revisit the forestry data discussed in Section 16.1.1, and assume it can be seen as a realization of a stationary marked point process on $D = \mathbb{R}^2 \times \mathbb{R}^+$, observed within the window $W = [0, 75]^2$. Then, for bounded Borel sets $A \subseteq \mathbb{R}^2$, $B \subseteq \mathbb{R}^+$, and $\mathbf{x} \in \mathbb{R}^2$,

$$M(T_{\mathbf{x}}A \times B) = M(A \times B).$$

If we additionally assume that the first-order moment measure of the marginal point process of locations exists, the translation invariance implies that M can be factorized as $\lambda \ell \times \nu_M$ where $\lambda \geq 0$ is the *intensity* of the marked point process, ℓ Lebesgue measure on \mathbb{R}^2 , and ν_M a probability measure on \mathbb{R}^+ , the *mark distribution*. By the Campbell theorem

$$\mathbf{E} \left[\sum_{(\mathbf{a}, m) \in X} m 1\{\mathbf{a} \in W\} \right] = 75^2 \lambda \int_0^\infty m d\nu_M(m).$$

In other words, a ratio-unbiased estimator of the mean tree diameter with respect to ν_M is the average observed diameter. The smoothed empirical mark distribution function ν_M is depicted in Figure 16.4. Note that ν_M has two clear modes.

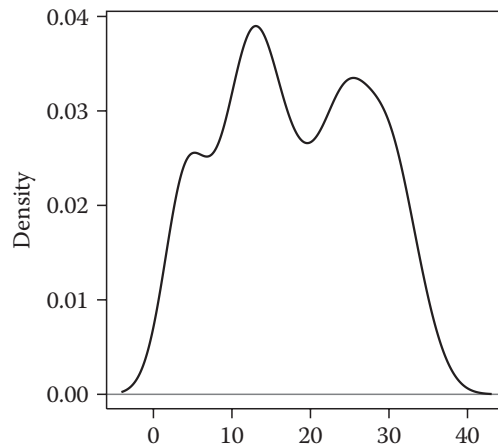


FIGURE 16.4
Estimated mark probability density function of Kaluzhskie Zaseki Forest data.

Moment measures may be refined by restricting attention to configurations with specified properties.

Definition 16.4

Let X be a point process on a complete, separable metric space (D, d) . Define

$$C(A \times F) = \mathbf{E}[N(A) 1\{X \in F\}]$$

for all bounded Borel sets $A \subseteq D$ and all $F \in \mathcal{N}^{\text{lf}}$.

Moment measures may not exist, but C always defines a proper measure, the Campbell measure of Kummer and Matthes (1970).

Theorem 16.4

The function C of Definition 16.4 can be extended uniquely to a σ -finite measure on the product σ -algebra of $\mathcal{B}(D)$ and \mathcal{N}^{lf} , the (first-order) Campbell measure.

Since $C(A \times \mathcal{N}^{\text{lf}}) = \mathbf{E}N(A) = M(A)$ for any Borel set A , it follows that C is finite if and only if M is also finite.

Rephrased in integral terms, Theorem 16.4 states that

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}, X) \right] = \int_D \int_{\mathcal{N}^{\text{lf}}} g(\mathbf{x}, \mathcal{X}) dC(\mathbf{x}, \mathcal{X}) \quad (16.3)$$

for any measurable function $g : D \times \mathcal{N}^{\text{lf}} \rightarrow \mathbb{R}$, which is either nonnegative or integrable with respect to the Campbell measure.

16.1.4 Reduced and Higher-Order Campbell Measures

Second- and higher-order moment measures are defined by considering $N(A_1), N(A_2), \dots$ jointly.

Definition 16.5

Let X be a point process on a complete, separable metric space (D, d) . Define, for $n \in \mathbb{N}$, and bounded Borel subsets A_1, \dots, A_n of D ,

$$M_n(A_1 \times \dots \times A_n) = \mathbf{E}[N(A_1) \cdots N(A_n)].$$

If all $A_i \equiv A$ are equal, $M_n(A \times \dots \times A) = M(A)^n$. Since the Borel rectangles form a semiring generating the Borel product σ -algebra, provided M_n is finite, it can be extended uniquely to a σ -finite measure on the product σ -algebra, the n th order moment measure of X . Its integral representation states that

$$\mathbf{E} \left[\sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in X} g(\mathbf{x}_1, \dots, \mathbf{x}_n) \right] = \int_D \cdots \int_D g(\mathbf{x}_1, \dots, \mathbf{x}_n) dM_n(\mathbf{x}_1, \dots, \mathbf{x}_n)$$

for any measurable function $g : D^n \rightarrow \mathbb{R}$ that is either nonnegative or integrable with respect to M_n , provided that M_n exists as a σ -finite measure.

Example 16.7

Let A and B be bounded Borel sets. The covariance of the random variables counting the number of points in A and B can be written in terms of the second-order moment

measure as

$$\text{Cov}(N(A), N(B)) = M_2(A \times B) - M(A)M(B).$$

In particular,

$$\text{var}(N(A)) = M_2(A \times A) - M(A)^2.$$

Example 16.8

For the binomial point process of Example 16.1, as $N(A_1)N(A_2) = \sum_{i=1}^n 1\{X_i \in A_1 \cap A_2\} + \sum_{i=1}^n \sum_{j \neq i} 1\{X_i \in A_1; X_j \in A_2\}$,

$$M_2(A_1 \times A_2) = n \frac{|A_1 \cap A_2|}{|D|} + n(n-1) \frac{|A_1||A_2|}{|D|^2}.$$

Next, consider the homogeneous Poisson process of Example 16.2. As the total number of points is Poisson distributed with mean $\lambda |D|$,

$$M_2(A_1 \times A_2) = \mathbf{E}[\mathbf{E}[N(A_1) \times N(A_2) \mid N(D)]] = \lambda |A_1 \cap A_2| + \lambda^2 |A_1||A_2|.$$

In the above example, M_2 was computed by distinguishing between pairs of identical and pairs of distinct points. If only the latter type are taken into consideration, we obtain the n th order factorial moment measure μ_n defined by the integral representation

$$\mathbf{E} \left[\sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in X}^{\neq} g(\mathbf{x}_1, \dots, \mathbf{x}_n) \right] = \int_D \cdots \int_D g(\mathbf{x}_1, \dots, \mathbf{x}_n) d\mu_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \tag{16.4}$$

for all measurable functions $g : D^n \rightarrow \mathbb{R}^+$ (see, e.g., Mecke (1976)). Standard measure-theoretic arguments imply Equation (16.4) holds true for any g that is integrable with respect to μ_n . The sum is over all n -tuples of distinct points. For $n = 1$, $\mu_1 = M_1 = M$, the first-order moment measure. We shall say that the n th order factorial moment measure exists if $\mu_n(A)$ is finite for all bounded Borel sets $A \subseteq D$.

Example 16.9

Let A_1, A_2 be Borel subsets of some compact set $D \subseteq \mathbb{R}^2$. Further to Example 16.8, the homogeneous Poisson process with intensity λ has second-order factorial moment measure

$$\mu_2(A_1 \times A_2) = \lambda^2 |A_1||A_2|.$$

For the binomial point process with n points,

$$\mu_2(A_1 \times A_2) = n(n-1) |A_1||A_2|/|D|^2.$$

If μ_n is absolutely continuous with respect to some n -fold product measure ν^n on D^n , then Equation (16.4) can be written as

$$\mathbf{E} \left[\sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in X}^{\neq} g(\mathbf{x}_1, \dots, \mathbf{x}_n) \right] = \int_D \cdots \int_D g(\mathbf{x}_1, \dots, \mathbf{x}_n) \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\nu(\mathbf{x}_1) \dots d\nu(\mathbf{x}_n).$$

The Radon–Nikodym derivative ρ_n of μ_n is referred to as its *product density*, and $\rho_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\nu(\mathbf{x}_1) \dots d\nu(\mathbf{x}_n)$ may be interpreted as the joint probability of a point falling in each of the infinitesimal regions centered at $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Example 16.10

Further to Example 16.9, the homogeneous Poisson process with intensity λ has constant second-order product density $\rho_2(\mathbf{x}_1, \mathbf{x}_2) = \lambda^2$, while for the binomial point process with

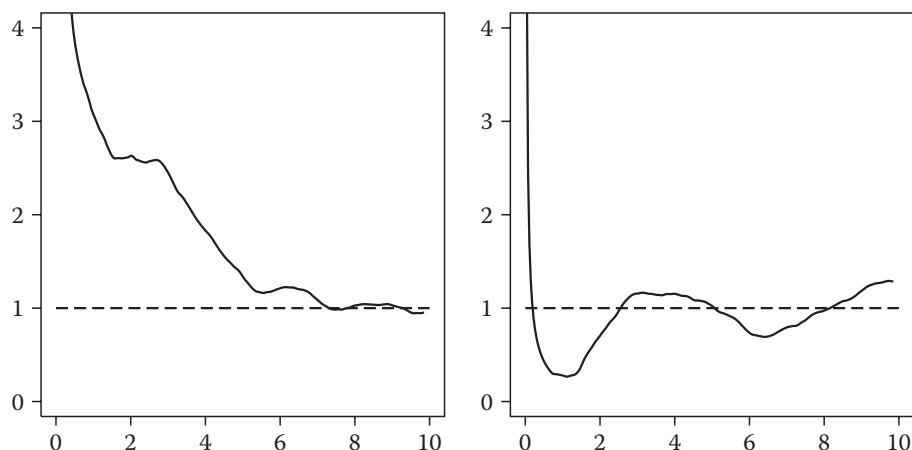


FIGURE 16.5

Estimated pair correlation functions of Kaluzhskie Zaseki Forest data (solid lines). (Left) Trees with diameter less than 15 cm, (right) those with diameter larger than 25 cm. The dashed lines are the pair correlation functions of a homogeneous Poisson process.

n points,

$$\rho_2(\mathbf{x}_1, \mathbf{x}_2) \equiv n(n-1)/|D|^2.$$

Example 16.11

We return to the setup of Example 16.4, and make the further assumptions that the (unmarked) point patterns X_s and X_l of small and large trees, respectively, are isotropic and that first- and second-order product densities exist. Then $\rho_1 \equiv \lambda$, the intensity, and $\rho_2(\mathbf{x}_1, \mathbf{x}_2) = \rho_2(\|\mathbf{x}_1 - \mathbf{x}_2\|)$ is a function only of the distance $r = \|\mathbf{x}_1 - \mathbf{x}_2\|$ between the points. Scaled by intensity, the estimated *pair correlation functions*

$$g(r) = \rho_2(r)/\lambda^2$$

of both patterns are plotted in Figure 16.5. As $g(r)$ greater than 1 indicates a higher probability of finding a pair of r -close points than for a homogeneous Poisson process with the same intensity, the left-hand panel of Figure 16.5 indicates clustering of small (young) trees. The trees having larger diameters avoid each other up to about 2 cm. The mode around 4 cm reflects the typical interpoint distance, which also explains the cyclic behavior of the graph. Note that this interpretation differs somewhat from our earlier interpretation of the empty space function for these data (right-hand panel of Figure 16.3). Later chapters will discuss in more detail the way in which different summary descriptions of point processes can yield different insights when used as tools for data analysis.

Higher-order Campbell measures can be defined similarly (Kallenberg, 1975), for instance

$$C_2(A_1 \times A_2 \times F) = \mathbf{E}[N(A_1)N(A_2)1\{X \in F\}]$$

for bounded Borel sets A_i , $i = 1, 2$, and $F \in \mathcal{N}^{\text{lf}}$.

To conclude this survey of Campbell-type measures, let us consider the concept of reduction. The basic idea is to single out each point in a configuration and consider the remaining pattern from this point's perspective. Therefore, reduced Campbell measures are useful tools to describe in a rigorous way the conditional distribution of a point process given that it places mass on a certain (marked) point and, dually, its behavior at a given location conditioned on the pattern elsewhere. A more detailed development of this idea will be given in the next section.

Definition 16.6

Let X be a simple point process on the complete, separable metric space (D, d) . Define

$$C^1(A \times F) = \int_{N^{lf}} \sum_{\mathbf{x} \in \mathcal{X}} 1\{\mathbf{x} \in A\} 1\{\mathcal{X} \setminus \{\mathbf{x}\} \in F\} d\mathcal{P}(\mathcal{X}) = \mathbf{E} \left[\sum_{\mathbf{x} \in X \cap A} 1\{X \setminus \{\mathbf{x}\} \in F\} \right]$$

for all bounded Borel sets $A \subseteq D$ and all $F \in \mathcal{N}^{lf}$.

As the Campbell measure of Theorem 16.4, C^1 can be extended uniquely to a σ -finite measure on the product σ -algebra of $\mathcal{B}(D)$ and \mathcal{N}^{lf} , the *first-order reduced Campbell measure*. Moreover, the integral representation

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}, X \setminus \{\mathbf{x}\}) \right] = \int_D \int_{N^{lf}} g(\mathbf{x}, \mathcal{X}) dC^1(\mathbf{x}, \mathcal{X}) \tag{16.5}$$

holds for any measurable function $g : D \times N^{lf} \rightarrow \mathbb{R}$ that is either nonnegative or C^1 -integrable. As for higher orders, Mecke (1979) (see also Hanisch (1982)) defined the *nth-order reduced Campbell measure* of a simple point process X by

$$C_n^1(A_1 \times \dots \times A_n \times F) = \mathbf{E} \left[\sum_{\substack{\neq \\ \mathbf{x}_1, \dots, \mathbf{x}_n \in X}} 1_{A_1}(\mathbf{x}_1) \dots 1_{A_n}(\mathbf{x}_n) 1\{X \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \in F\} \right],$$

or in integral terms

$$\begin{aligned} & \mathbf{E} \left[\sum_{\substack{\neq \\ \mathbf{x}_1, \dots, \mathbf{x}_n \in X} g(\mathbf{x}_1, \dots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \right] \\ &= \int_D \dots \int_D \int_{N^{lf}} g(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathcal{X}) dC_n^1(\mathbf{x}_1, \dots, \mathbf{x}_n, \mathcal{X}) \end{aligned}$$

for any measurable function $g : D^n \times N^{lf} \rightarrow \mathbb{R}^+$, cf. (16.4).

16.1.5 Palm Theory and Conditioning

Henceforth, assume the first-order moment measure M exists and is σ -finite. Fix $F \in \mathcal{N}^{lf}$ and note that if $C(A \times N^{lf}) = M(A) = 0$, so is $C(A \times F)$. In other words, the marginal Campbell measure with second argument fixed at F is absolutely continuous with respect to M , so that for all $A \in \mathcal{B}(D)$,

$$C(A \times F) = \int_A \mathcal{P}_{\mathbf{x}}(F) dM(\mathbf{x})$$

for some nonnegative Borel measurable function $\mathbf{x} \mapsto \mathcal{P}_{\mathbf{x}}(F)$ on D defined uniquely up to an M -null set. As the exceptional null set may depend on F , it is not immediately clear that the $\mathcal{P}_{\mathbf{x}}(\cdot)$ are countably additive. However, the topological structure imposed on D implies that a version of the Radon–Nikodym derivatives $\mathcal{P}_{\mathbf{x}}(F)$ can be found such that

- For fixed $\mathbf{x} \in D$, $\mathcal{P}_{\mathbf{x}}(F)$ is a probability distribution on $(N^{lf}, \mathcal{N}^{lf})$.
- For fixed F , $\mathcal{P}_{\mathbf{x}}(F)$ is a Borel measurable function on D .

The probability distributions $\mathcal{P}_{\mathbf{x}}(\cdot)$ thus defined are *Palm distributions* of X at $\mathbf{x} \in D$, named in honor of Palm (1943). The Radon–Nikodym approach discussed above is due to Ryll–Nardzewski (1961).

Palm distributions are especially useful in simplifying Campbell-type formulas. Indeed, if \mathbf{E}_x denotes expectation with respect to \mathcal{P}_x , (16.3) can be rewritten as

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}, X) \right] = \int_D \int_{N^{\text{lf}}} g(\mathbf{x}, \mathcal{X}) d\mathcal{P}_x(\mathcal{X}) dM(\mathbf{x}) = \int_D \mathbf{E}_x [g(\mathbf{x}, X)] dM(\mathbf{x})$$

for any measurable function $g: D \times N^{\text{lf}} \rightarrow \mathbb{R}$ that is either nonnegative or C -integrable.

Heuristically speaking, recall that the distribution of a simple point process is fully determined by its void probabilities. Pick some small $\epsilon > 0$ and consider

$$\mathcal{P}(N(A) = 0 \mid N(b(\mathbf{x}, \epsilon)) > 0) \approx \frac{C(b(\mathbf{x}, \epsilon) \times \{N(A) = 0\})}{\mathcal{P}(N(b(\mathbf{x}, \epsilon)) > 0)} = \frac{\int_{b(\mathbf{x}, \epsilon)} \mathcal{P}_y(N(A) = 0) dM(\mathbf{y})}{\mathcal{P}(N(b(\mathbf{x}, \epsilon)) > 0)}.$$

For small ϵ , the numerator is approximately equal to $\mathcal{P}_x(N(A) = 0) M(b(\mathbf{x}, \epsilon))$, the denominator to $M(b(\mathbf{x}, \epsilon))$. Hence, $\mathcal{P}_x(N(A) = 0)$ may be seen as the conditional void probability of A given that X places mass on \mathbf{x} .

Example 16.12

The lack of dependence between the points of a Poisson process (cf. Example 16.2) implies a particularly simple form for its Palm distribution. Indeed, let X be a homogeneous Poisson process with intensity $\lambda > 0$ on \mathbb{R}^2 and write \mathcal{P} for its distribution. Then, a Palm distribution of X at $\mathbf{x} \in \mathbb{R}^2$ is given by $\mathcal{P} * \delta_x$, the convolution of \mathcal{P} with an atom at \mathbf{x} .

To see this, note that since both \mathcal{P}_x and $\mathcal{P} * \delta_x$ are simple, by Theorem 16.2 it is sufficient to prove that their void probabilities coincide. To do so, let B be a bounded Borel set and write $v_x(B)$ for the probability under $\mathcal{P} * \delta_x$ that there are no points in B , $v(B)$ for $\mathcal{P}(N(B) = 0)$. For any bounded Borel set $A \subseteq \mathbb{R}^2$, by definition

$$C(A \times \{N(B) = 0\}) = \lambda \int_A \mathcal{P}_x(N(B) = 0) d\mathbf{x}.$$

On the other hand,

$$\begin{aligned} C(A \times \{N(B) = 0\}) &= C(A \setminus B \times \{N(B) = 0\}) = \lambda |A \setminus B| v(B) = \lambda \int_{A \setminus B} v(B) d\mathbf{x} \\ &= \lambda \int_A v_x(B) d\mathbf{x}, \end{aligned}$$

using the fact that for a Poisson process the random variables $N(A \setminus B)$ and $N(B)$ are independent. Since A was chosen arbitrarily, $\mathcal{P}_x(N(B) = 0) = v_x(B)$ for almost all \mathbf{x} .

Example 16.13

For a simple, stationary marked point process on $D = \mathbb{R}^2 \times \mathbb{R}^+$, as considered in Example 16.6, for all Borel subsets B of \mathbb{R}^+ and $\mathbf{y} \in \mathbb{R}^2$, let $F_B^y = \{\mathcal{X} : (\mathbf{y}, m) \in \mathcal{X} \text{ for some } m \in B\}$ be the event of finding a point at \mathbf{y} with mark in B . Because of the stationarity, a version of Palm distributions can be found that are translates of a single probability distribution, $\mathcal{P}_{(\mathbf{y}, m)}(\{T_y \mathcal{X} : \mathcal{X} \in F\}) = \mathcal{P}_{(0, m)}(F)$ (almost everywhere). An application of the Campbell theorem for $g((\mathbf{y}, m), \mathcal{X}) = 1\{\mathbf{y} \in A; \mathcal{X} \in F_B^y\}$ yields

$$\lambda \int_A \int_0^\infty \mathcal{P}_{(\mathbf{y}, m)}(F_B^y) dv_M(m) d\mathbf{y} = \lambda \int_A \int_0^\infty \mathcal{P}_{(0, m)}(F_B^0) dv_M(m) d\mathbf{y} = \lambda v_M(B) |A|$$

for all bounded Borel sets A . Thus, the mark distribution v_M may be interpreted as the probability distribution of the mark at an arbitrarily chosen point.

Higher-order and reduced Palm distributions can be defined in a straightforward fashion. For example, suppose that the second-order moment measure M_2 exists. Then, Jagers (1973) defined second-order Palm distributions $\mathcal{P}_{\mathbf{x}_1, \mathbf{x}_2}(F)$ that satisfy

$$C_2(A_1 \times A_2 \times F) = \int_{A_1} \int_{A_2} \mathcal{P}_{\mathbf{x}_1, \mathbf{x}_2}(F) dM_2(\mathbf{x}_1, \mathbf{x}_2)$$

for all Borel sets $A_i, i = 1, 2$, and $F \in \mathcal{N}^{\text{lf}}$. To obtain a reduced Palm distribution, replace M_n by μ_n and C_n by $C_n^!$ (Mecke, 1979; Hanisch, 1982). The equivalent integral representations hold as well. In particular, Equation (16.5) can be rephrased as

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}, X \setminus \{\mathbf{x}\}) \right] = \int_D \int_{\mathcal{N}^{\text{lf}}} g(\mathbf{x}, \mathcal{X}) d\mathcal{P}_{\mathbf{x}}^!(\mathcal{X}) dM(\mathbf{x}) = \int_D \mathbf{E}_{\mathbf{x}}^! [g(\mathbf{x}, X)] dM(\mathbf{x}),$$

where $\mathcal{P}_{\mathbf{x}}^!$ is a reduced Palm distribution at $\mathbf{x} \in D$.

Example 16.14

From Example 16.12, it follows immediately that a reduced Palm distribution of a homogeneous planar Poisson process with intensity λ is its distribution \mathcal{P} . This fundamental property is a consequence of the *Slivnyak–Mecke theorem* stating that the homogeneous Poisson process is characterized by reduced Palm distributions that are translates of a single probability distribution $\mathcal{P}_{\mathbf{x}}^! (\{T_{\mathbf{x}}\mathcal{X} : \mathcal{X} \in F\}) = \mathcal{P}_0^!(F)$ for almost all \mathbf{x} and $\mathcal{P}_0^! = \mathcal{P}$ (Jagers, 1973; Kerstan and Matthes, 1964; Mecke, 1967; Slivnyak, 1962). Hence, the reduced Campbell measure

$$C^!(A \times F) = \lambda \int_A \mathcal{P}_{\mathbf{x}}^!(F) d\mathbf{x} = \lambda \mathcal{P}(F) |A|,$$

where $A \subseteq \mathbb{R}^2$ is a bounded Borel set and $F \in \mathcal{N}^{\text{lf}}$ is a product measure.

The analog of the empty space function (16.1) in a Palm context is the *nearest-neighbor distance distribution function* G . Let X be a simple, stationary point process with locations in \mathbb{R}^d . Then, for any $r \geq 0$,

$$G(r) = \mathcal{P}_{\mathbf{x}}^!(d(\mathbf{x}, X) \leq r)$$

is the probability that X places at least one point within distance r of some arbitrarily chosen $\mathbf{x} \in \mathbb{R}^d$. As translation invariance is inherited by Palm distributions, $G(r)$ is well defined and does not depend on the choice of \mathbf{x} .

Example 16.15

Estimated nearest-neighbor distance distribution functions for young and established trees in the Kaluzhskie Zaseki Forest data are given in Figure 16.6 and are compared to the G -function of a Poisson process with the same expected number of points (cf. Example 16.12). For small (young) trees, small distances to the nearest neighbor occur with a higher probability than under the Poisson model. Such a behavior is typical for clustered patterns with groups of nearby points separated by gaps. In contrast to the empty space distribution, the G -graph for larger trees differs from its Poisson counterpart in that up until about $r = 4$ cm, $G(r)$ is smaller than under the Poisson model, indicating repulsion at such range; beyond this range, the nearest-neighbor distance distribution is larger than what would be expected if there were no spatial interaction with the steep increase beyond $r = 4$ cm reflecting the preferred interpoint separation.

The ratio

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

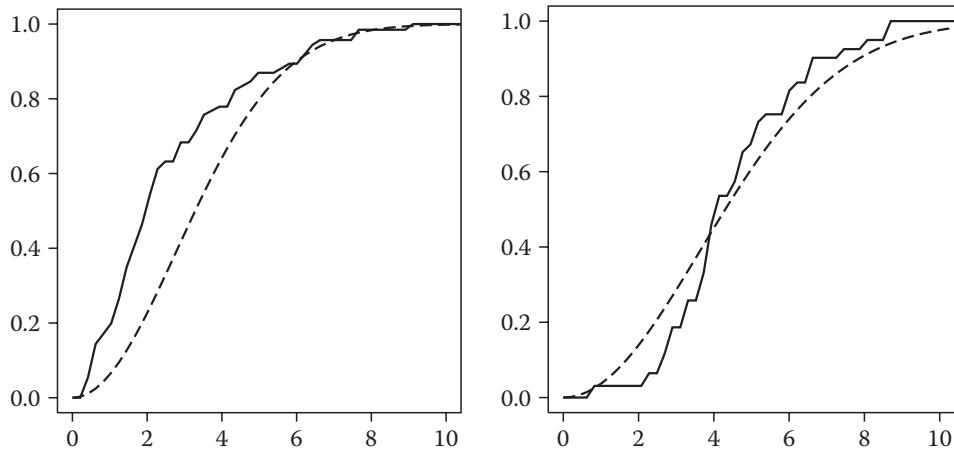


FIGURE 16.6

Estimated nearest-neighbor distance distribution functions of Kaluzhskie Zaseki Forest data (solid lines). (Left) Locations of trees with diameter less than 15 cm, (right) those with diameter larger than 25 cm. The dashed lines are the nearest-neighbor distance distribution functions for homogeneous Poisson processes with, on average, the same number of points-per-unit area as the data.

defined for all $r \geq 0$ such that $F(r) < 1$, compares the void probabilities of closed balls $\bar{b}(\mathbf{x}, r)$ under $\mathcal{P}_x^!$ and \mathcal{P} . The advantage of considering the ratio is that no reference to the intensity of some Poisson process needs to be made. Indeed, values less than 1 indicate that the size of empty spaces tends to be larger than the distance between nearest-neighbor pairs (clustering), whereas values exceeding 1 suggest a more regular pattern (van Lieshout and Baddeley, 1996). The left and center panels of Figure 16.7 confirm the interaction structure implied by the nearest-neighbor distance distribution function. By way of illustration, the J -function of all locations regardless of the mark is shown as well. It indicates clustering at the smallest scales, has a repulsion peak around 2 to 3 cm, then decreases.

Functions, such as F , g , G , or J , provide valuable information and their empirical counterparts are useful tools for data analysis, as discussed in later chapters. Nevertheless, no such low-dimensional function fully characterizes the distribution of the process and it is wise to plot several to grasp different aspects of the underlying distribution.

As we have seen, Palm distributions can be interpreted as conditional distributions given that there is mass at some fixed point. We now turn to describing the dual notion of the

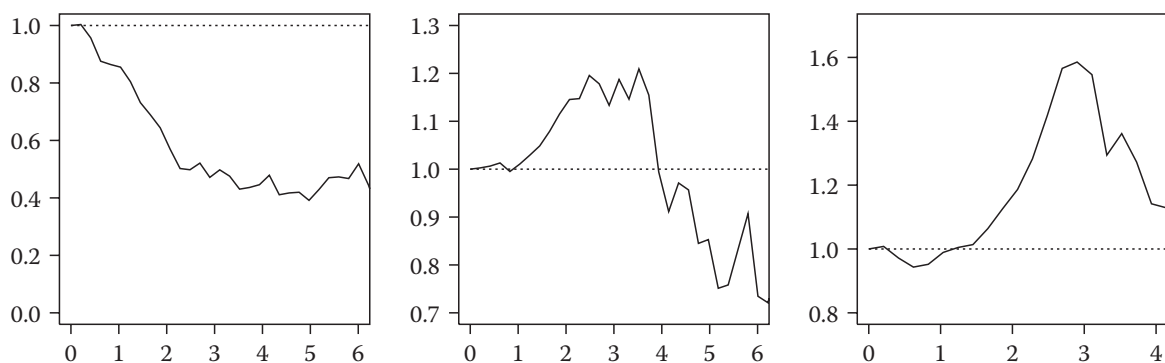


FIGURE 16.7

Estimated J -functions of Kaluzhskie Zaseki Forest data (solid lines). (Left) Locations of trees with diameter less than 15 cm, (center) those with diameter larger than 25 cm, and (right) all tree locations. The dashed lines are the J -functions of a homogeneous Poisson process.

probability mass of some fixed point, conditional on knowing the full realization of the point process elsewhere.

Let X be a simple point process on a complete, separable metric space for which the first-order moment measure $M(\cdot)$ exists as a σ -finite measure. Furthermore, assume that for any fixed bounded Borel set A , the marginal reduced Campbell measure $C^!(A \times \cdot)$ is absolutely continuous with respect to the distribution \mathcal{P} of X . Then

$$C^!(A \times F) = \int_A P_x^!(F) dM(\mathbf{x}) = \int_F \Lambda(A; \mathcal{X}) d\mathcal{P}(\mathcal{X}) \tag{16.6}$$

for some \mathcal{N}^{lf} -measurable function $\Lambda(A; \cdot)$, specified uniquely up to a \mathcal{P} -null set. If $\Lambda(\cdot; \mathcal{X})$ admits a Radon–Nikodym derivative $\lambda(\cdot; \mathcal{X})$ with respect to some Borel measure ν on D ,

$$C^!(A \times F) = \int_F \int_A \lambda(\mathbf{x}; \mathcal{X}) d\nu(\mathbf{x}) d\mathcal{P}(\mathcal{X}) = \int_A \mathbf{E}[1_F(X) \lambda(\mathbf{x}; X)] d\nu(\mathbf{x}).$$

Replacing the indicator function $1_A(\mathbf{x}) 1_F(\mathcal{X})$ by arbitrary nonnegative functions, we obtain the following definition.

Definition 16.7

Let X be a simple point process on a complete, separable metric space (D, d) equipped with Borel measure ν . If for any measurable function $g : D \times \mathcal{N}^{\text{lf}} \rightarrow \mathbb{R}^+$,

$$\mathbf{E} \left[\sum_{\mathbf{x} \in X} g(\mathbf{x}, X \setminus \{\mathbf{x}\}) \right] = \int_D \mathbf{E}[g(\mathbf{x}, X) \lambda(\mathbf{x}; X)] d\nu(\mathbf{x})$$

for some measurable function $\lambda : D \times \mathcal{N}^{\text{lf}} \rightarrow \mathbb{R}^+$, then X is said to have Papangelou conditional intensity λ (Papangelou, 1974).

From a heuristic point of view, by Equation (16.6), $dP_x^!(\mathcal{X}) dM(\mathbf{x}) = \lambda(\mathbf{x}; \mathcal{X}) d\mathcal{P}(\mathcal{X}) d\nu(\mathbf{x})$; hence, $\lambda(\mathbf{x}; \mathcal{X}) d\nu(\mathbf{x})$ can be interpreted as the infinitesimal probability of finding a point of X at the infinitesimal region centered at \mathbf{x} conditional on the event that X and \mathcal{X} agree on the complement of this region. As we shall see in the next section, conditional intensities are especially useful for finite point processes.

Example 16.16

Further to Example 16.14, note that for a homogeneous planar Poisson process with intensity λ ,

$$C^!(A \times F) = \lambda \mathcal{P}(F) |A| = \int_F [\lambda |A|] d\mathcal{P}(\mathcal{X})$$

so that $\Lambda(A; \mathcal{X}) \equiv \lambda |A|$ is a Papangelou kernel. This kernel has density λ with respect to Lebesgue measure and does not depend on either \mathbf{x} or \mathcal{X} , as indeed one would have anticipated because of the strong independence properties of the Poisson process.

Example 16.17

Let X be a stationary, simple point process on \mathbb{R}^d with intensity $\lambda > 0$. Suppose X admits a conditional intensity and let $g : \mathcal{N}^{\text{lf}} \rightarrow \mathbb{R}^+$ be a measurable function. Then,

$$\lambda \int_A E_x^! g(X) d\mathbf{x} = \int_A \mathbf{E}[g(X) \lambda(\mathbf{x}; X)] d\mathbf{x}$$

for all Borel sets A . In fact the Georgii–Nguyen–Zessin formula (Georgii, 1976; Nguyen and Zessin, 1979) states that

$$\lambda E_x^! g(X) = \mathbf{E}[g(X) \lambda(\mathbf{x}; X)] \tag{16.7}$$

for almost all $\mathbf{x} \in \mathbb{R}^d$. Thus, expectations with respect to \mathcal{P} may be translated into reduced Palm expectations and vice versa.

16.1.6 Finite Point Processes

Most point patterns encountered in practice are observed in a bounded region. Sometimes this region is dictated by the application; more often, the spatial process of interest extends over a space that is too large to be mapped exhaustively and data are recorded in a smaller “window” chosen for convenience. In any case, the resulting map contains a *finite* number of points.

The distribution of point processes whose realizations are almost surely finite can be described as follows (Daley and Vere-Jones, 1988; Reiss, 1993). For convenience, suppose that (D, d) is equipped with a Borel measure ν and assume $0 < \nu(D) < \infty$. Then, it suffices to specify

- A discrete probability distribution $(p_n)_{n \in \mathbb{N}_0}$ for the number of points
- A family of symmetric probability densities $\pi_n(\mathbf{x}_1, \dots, \mathbf{x}_n)_{n \in \mathbb{N}}$ with respect to the n -fold product of ν for the points themselves (i.e., for point locations, and marks if applicable)

The symmetry requirement for the π_n is needed to make sure that the patterns generated by the π_n are permutation invariant, in other words, do not depend on the order in which their points are listed.

It should be noted that for many point processes, p_n cannot be expressed in closed form. If it can, the model specification is algorithmic, so that realizations are easily obtained. Note that if ν is diffuse, the point process is simple.

Janossy densities (Janossy, 1950) are defined in terms of $(p_n, \pi_n)_{n \in \mathbb{N}_0}$, as

$$j_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = n! p_n \pi_n(\mathbf{x}_1, \dots, \mathbf{x}_n), \quad n \in \mathbb{N}_0. \quad (16.8)$$

In an infinitesimal sense, $j_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\nu(\mathbf{x}_1) \dots d\nu(\mathbf{x}_n)$ is the probability of finding exactly n points, one at each of infinitesimal regions centered at $\mathbf{x}_1, \dots, \mathbf{x}_n$. For $n = 0$, Equation (16.8) is conventionally read as $j_0(\emptyset) = p_0$. As $\int_D \dots \int_D j_n(\mathbf{x}_1, \dots, \mathbf{x}_n) d\nu(\mathbf{x}_1) \dots d\nu(\mathbf{x}_n) = n! p_n$, one may retrieve p_n and, hence, π_n from j_n .

Example 16.18

For the binomial point process introduced in Example 16.1, $p_m = 1\{m = n\}$ for some fixed $n \in \mathbb{N}$. As the points are uniformly and independently scattered over some compact set D , $\pi_n \equiv |D|^{-n}$ with respect to Lebesgue measure. For $m \neq n$, π_n may be defined arbitrarily.

The homogeneous Poisson process of Example 16.2 is described by $p_n = \exp[-\lambda |D|] (\lambda |D|)^n / n!$ and, for each $n \in \mathbb{N}_0$, $\pi_n \equiv |D|^{-n}$ with respect to the n -fold product of Lebesgue measures. Consequently, $j_n = \lambda^n \exp[-\lambda |D|]$.

Although in most applications it is not realistic to assume that points are scattered randomly, Poisson processes are supremely useful as benchmarks. Indeed, one may construct a wide range of point process models by specifying their probability density (Radon–Nikodym derivative) with respect to a Poisson process with finite, diffuse intensity measure ν . However, not all finite point processes are absolutely continuous with respect to a given Poisson model; counterexamples on the plane equipped with Lebesgue measure are those that place points deterministically or in equidistant pairs.

Let X be a point process defined by its probability density f . The probability of the event $\{N(D) = n\}$ is

$$p_n = \frac{e^{-\nu(D)}}{n!} \int_D \cdots \int_D f(\mathbf{x}_1, \dots, \mathbf{x}_n) d\nu(\mathbf{x}_1) \dots d\nu(\mathbf{x}_n)$$

for each $n \in \mathbb{N}$, with $p_0 = e^{-\nu(D)} f(\emptyset)$ equal to the probability of the empty configuration \emptyset . If $p_n > 0$, conditionally on X containing exactly n points, their joint probability distribution is proportional to $f(\mathbf{x}_1, \dots, \mathbf{x}_n)$. It follows that $j_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = e^{-\nu(D)} f(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is also proportional to f .

Example 16.19

Return to the setup of Example 16.2, and let ν be Lebesgue measure restricted to a compact set $D \subseteq \mathbb{R}^2$. A comparison of the Janossy densities of the homogeneous Poisson process X with intensity λ to those of a unit intensity Poisson process yields that X must have density

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \exp[(1 - \lambda)|D|] \lambda^n, \quad n \in \mathbb{N}_0,$$

with respect to the distribution of the unit intensity Poisson process (see also Example 16.18). It is not hard to verify that f indeed defines the distribution of X .

The binomial point process has density

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = n! e^{|D|} |D|^{-n}$$

for configurations consisting of n points, and $f \equiv 0$ on $\{N(D) \neq n\}$ with respect to the distribution of a unit intensity Poisson process on D .

For finite point processes that are defined in terms of a density with respect to some Poisson process, conditional intensities exist and are easy to compute (Ripley and Kelly, 1977).

Theorem 16.5

Let X be a finite point process on a complete, separable metric space (D, d) with probability density f with respect to the distribution of a Poisson process on D with finite, diffuse intensity measure ν and assume that $f(\mathcal{X}) > 0$ implies $f(\mathcal{X}') > 0$ for all $\mathcal{X}' \subseteq \mathcal{X}$. Then, X has Papangelou conditional intensity

$$\lambda(\mathbf{x}; \mathcal{X}) = \frac{f(\mathcal{X} \cup \{\mathbf{x}\})}{f(\mathcal{X})} \tag{16.9}$$

for $\mathbf{x} \in D \setminus \mathcal{X}$ and configurations \mathcal{X} such that $f(\mathcal{X}) > 0$.

For processes that are absolutely continuous, the conditional intensity provides a fourth way of defining the distribution. Indeed, by Equation (16.9),

$$f(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \propto \prod_{i=1}^n \lambda(\mathbf{x}_i; \{\mathbf{x}_1, \dots, \mathbf{x}_{i-1}\})$$

regardless of the order in which the points are labeled.

In summary, finite point processes are usually easier to deal with than most infinite ones. Their distribution is known once we have specified either $(p_n, \pi_n)_{n \in \mathbb{N}_0}$ or $j_n, n \in \mathbb{N}_0$. If the point process is absolutely continuous with respect to a Poisson process, alternative modeling strategies are to specify f or $\lambda(\cdot; \cdot)$.

16.1.7 Gibbs Measures by Local Specification

Infinite point processes are not easily specified by a density with respect to a (homogeneous) Poisson process. Indeed, even homogeneous Poisson processes on \mathbb{R}^d with different intensities are not absolutely continuous with respect to each other (see, e.g., van Lieshout (2000)). Nevertheless, one may try to define a point process X by specification of a family of conditional densities for the finite restrictions of X to bounded Borel sets.

Again, assume that (D, d) is a complete, separable metric space equipped with a diffuse, locally finite Borel measure ν . Define a family of probability densities $f_B(\mathcal{X} \cap B \mid \mathcal{X} \cap B^c)$, B a bounded Borel set, with respect to the law of a Poisson process with intensity measure ν restricted to B . Here \mathcal{X} ranges through \mathcal{N}^{lf} . Then, for $F \in \mathcal{N}^{\text{lf}}$,

$$\sum_{n=0}^{\infty} \frac{e^{-\nu(B)}}{n!} \int_B \cdots \int_B 1_F(\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \cup (\mathcal{X} \cap B^c)) f_B(\mathbf{x}_1, \dots, \mathbf{x}_n \mid \mathcal{X} \cap B^c) d\nu(\mathbf{x}_1) \cdots d\nu(\mathbf{x}_n)$$

can be interpreted as the probability that X falls in F conditional on the event $\{X \cap B^c = \mathcal{X} \cap B^c\}$.

It is not at all obvious that the family $f_B(\cdot \mid \cdot)$ defines a point process distribution \mathcal{P} . A necessary condition is that the *Dobrushin–Landford–Ruelle equations*

$$\mathcal{P}(F) = \int \sum_{n=0}^{\infty} \frac{e^{-\nu(B)}}{n!} \int_B \cdots \int_B 1_F(\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \cup (\mathcal{X} \cap B^c)) f_B(\mathbf{x}_1, \dots, \mathbf{x}_n \mid \mathcal{X} \cap B^c) d\nu(\mathbf{x}_1) \cdots d\nu(\mathbf{x}_n) d\mathcal{P}(\mathcal{X}) \tag{16.10}$$

are satisfied for all $F \in \mathcal{N}^{\text{lf}}$ and all bounded Borel sets $B \subseteq \mathbb{R}^d$, in which case \mathcal{P} is said to be a *Gibbs point process with local specification* $f_B(\cdot \mid \cdot)$.

An extensive overview of conditions for the existence of a solution to Equation (16.10) was given in Preston (1976), typically phrased in terms of bounds on the log conditional densities or intensities, and restrictions on the points in B^c that may affect f_B . For later developments, see, for example, Georgii and Häggström (1996) or Glötzl (1980). Note that even if there is a solution, it may not be unique, a phenomenon known as “phase transition” (see Preston (1976)).

Example 16.20

Let ν be Lebesgue measure on \mathbb{R}^2 and consider the *area interaction process* defined by

$$f_B(\mathcal{X} \cap B \mid \emptyset) = \alpha_B(\emptyset) \beta^{n(\mathcal{X})} \exp[-\beta |U_r(\mathcal{X})|]$$

for bounded Borel sets B in the plane. Here, the parameter β is assumed to be strictly positive, $U_r(\mathcal{X}) = \cup_{\mathbf{x}_i \in \mathcal{X}} \bar{b}(\mathbf{x}_i, r)$ is the union of closed balls around each of the points in \mathcal{X} , and $\alpha_B(\emptyset)$ is the normalizing constant that makes $f_B(\cdot \mid \emptyset)$ a probability density.

Note that a ball of radius r intersects B only if its center is at most r away from the boundary of B or, in other words, falls in $B_{\oplus r} = \{\mathbf{b} + \mathbf{c} : \mathbf{b} \in B; \mathbf{c} \in \bar{b}(0, r)\}$. Hence, we may define, for any configuration \mathcal{X} and bounded Borel set $B \subseteq \mathbb{R}^2$,

$$f_B(\mathcal{X} \cap B \mid \mathcal{X} \cap B^c) = \alpha_B(\mathcal{X} \cap B^c) \beta^{n(\mathcal{X})} \exp[-\beta |U_r(\mathcal{X}) \cap B_{\oplus r}|],$$

where again $\alpha_B(\mathcal{X} \cap B^c)$ is the normalizing constant. It can be shown that there exists a stationary solution to the Dobrushin–Landford–Ruelle equations (Baddeley and van Lieshout, 1995). Moreover, it is known that multiple solutions exist when β is large, whereas the solution is unique for β small (Ruelle, 1971; see also Chayes, Chayes, and Kotecky 1995; Dereudre, 2008; and Georgii and Häggström, 1996). It is plausible that the occurrence of a phase transition is monotone in β , but to the best of our knowledge no rigorous proof exists to date.

References

- A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, 46:601–619, 1995.
- C. Berg, J.P.R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups*. Springer, Berlin, 1984.
- N.R. Campbell. The study of discontinuous phenomena. *Proceedings of the Cambridge Philosophical Society*, 15:117–136, 1909.
- J.T. Chayes, L. Chayes, and R. Kotecky. The analysis of the Widom–Rowlinson model by stochastic geometric methods. *Communications of Mathematical Physics*, 172:551–569, 1995.
- G. Choquet. Theory of capacities. *Annals of the Fourier Institute*, V:131–295, 1953/1954.
- D.R. Cox and V. Isham. *Point Processes*. Chapman & Hall, London, 1980.
- D. Daley and D.J. Vere-Jones. *An Introduction to the Theory of Point Processes*, 1st ed., Springer, New York, 1988. *Second edition, Volume I: Elementary Theory and Methods*, 2003. *Second edition, Volume II: General Theory and Structure*, 2008.
- D. Dereudre. Existence of quermass processes for nonlocally stable interaction and nonbounded convex grains. Preprint, Université de Valenciennes et du Hainaut-Cambrésis, France, 2008.
- A.K. Erlang. The theory of probabilities and telephone conversations (in Danish). *Nyt Tidsskrift for Matematik*, B20:33–41, 1909.
- H.-O. Georgii. Canonical and grand canonical Gibbs states for continuum systems. *Communications of Mathematical Physics*, 48:31–51, 1976.
- H.-O. Georgii and O. Häggström. Phase transition in continuum Potts models. *Communications of Mathematical Physics*, 181:507–528, 1996.
- E. Glötzl. Lokale Energien und Potentiale für Punktprozesse (in German). *Mathematische Nachrichten*, 96:198–206, 1980.
- P. Grabarnik and S.N. Chiu. Goodness-of-fit test for complete spatial randomness against mixtures of regular and clustered spatial point processes. *Biometrika*, 89:411–421, 2002.
- K.-H. Hanisch. On inversion formulae for n -fold Palm distributions of point processes in LCS-spaces. *Mathematische Nachrichten*, 106:171–179, 1982.
- J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester, U.K., 2008.
- L. Janossy. *Cosmic Rays*. Oxford University Press, Oxford, U.K., 1948.
- L. Janossy. On the absorption of a nucleon cascade. *Proceedings of the Royal Irish Academy of Sciences*, A53:181–188, 1950.
- O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, 1975.
- J. Kerstan and K. Matthes. Verallgemeinerung eines Satzes von Sliwnjak. (in German). *Académie de la République Populaire Roumaine. Revue Roumaine de Mathématiques Pures et Appliquées*, 9:811–830, 1964.
- A.Y. Khinchin. *Mathematical Methods in the Theory of Queueing*. Griffin, London, 1960.
- G. Kummer and K. Matthes. Verallgemeinerung eines Satzes von Sliwnjak. II (in German). *Académie de la République Populaire Roumaine. Revue Roumaine de Mathématiques Pures et Appliquées*, 15:845–870, 1970.
- G. Matheron. *Random Sets and Integral Geometry*. John Wiley & Sons, New York, 1975.
- K. Matthes, J. Kerstan, and J. Mecke. *Infinitely Divisible Point Processes*. Wiley, Chichester, U.K., 1978.
- B. McMillan. Absolutely monotone functions. *Annals of Mathematics*, 60:467–501, 1953.
- G. Mönch. Verallgemeinerung eines Satzes von A. Rényi (in German). *Studia Scientiarum Mathematicarum Hungarica*, 6:81–90, 1971.
- J. Neyman. On a new class of “contagious” distributions applicable in entomology and bacteriology. *Annals of Mathematical Statistics*, 10:35–57, 1939.
- X.X. Nguyen and H. Zessin. Integral and differential characterization of the Gibbs process. *Mathematische Nachrichten*, 88:105–115, 1979.
- T. Norberg. Existence theorems for measures on continuous posets, with applications to random set theory. *Mathematica Scandinavica*, 64:15–51, 1989.
- C. Palm. Intensitätsschwankungen im Ferngesprächverkehr (in German). *Ericsson Techniks*, 44:1–189, 1943.

- F. Papangelou. The conditional intensity of general point processes and an application to line processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 28:207–226, 1974.
- S.D. Poisson. *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précédées des Règles Générales du Calcul des Probabilités* (in French). Bachelier, Paris, 1837.
- C. Preston. *Random Fields*. Springer, Berlin, 1976.
- R.-D. Reiss. *A Course on Point Processes*. Springer, New York, 1993.
- A. Rényi. Remarks on the Poisson process. *Studia Scientiarum Mathematicarum Hungarica*, 2:119–123, 1967.
- B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188–192, 1977.
- D. Ruelle. Existence of a phase transition in a continuous classical system. *Physical Review Letters*, 27:1040–1041, 1971.
- C. Ryll-Nardzewski. Remarks on processes of calls. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics*, 2:455–465, 1961.
- Y.M. Slivnyak. Some properties of stationary flows of homogeneous random events (in Russian). *Teorija Veroyatnostei i ee Primenenija*, 7:347–352, 1962.
- O.V. Smirnova (Ed.) *East-European Broad Leaved Forests* (in Russian). Nauka, Moscow, 1994.
- S.K. Srinivasan. *Stochastic Point Processes and Their Applications*. Griffin, London, 1974.
- M.N.M. van Lieshout. *Markov Point Processes and Their Applications*. Imperial College Press, London, 2000.
- M.N.M. van Lieshout and A.J. Baddeley. A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50:344–361, 1996.

17

Spatial Point Process Models

Valerie Isham

CONTENTS

17.1	Model Construction.....	283
17.1.1	The Probability Generating Functional.....	286
17.1.2	Thinning.....	287
17.1.3	Translation.....	288
17.1.4	Superposition.....	289
17.2	Poisson Processes.....	289
17.2.1	Nonhomogeneous Poisson Process.....	290
17.2.2	The Cox Process.....	291
17.3	Poisson Cluster Processes.....	292
17.3.1	The Compound Poisson Process.....	293
17.3.2	The Neyman–Scott Poisson Cluster Process.....	293
17.4	Markov Point Processes.....	294
	References.....	297

17.1 Model Construction

This chapter describes various classes of parametric point process models that are useful for scientific work. In any particular application, there will generally be an underlying physical mechanism that generates the point events that are observed, and which may be fully or partially understood. In building a model for these events, the modeler often seeks to represent or reflect that physical process, albeit in a highly simplified way. Thus, for example, if the point events are the locations of seedling trees, it will be natural to build a model that takes into account the positions of parent trees, the clustering of the seedlings around these, and perhaps also the prevailing wind direction, even if the exact process of seed generation and dispersal is not represented. Such a model is often termed “mechanistic,” and has interpretable parameters that relate to physical phenomena. In contrast, “descriptive” models aim to represent the statistical properties of the data and their dependence on explanatory variables without necessarily worrying about the physical mechanisms involved. For example, a model that involves inhibition between nearby events can be used to model the positions of ants’ nests (Harkness and Isham, 1983; see also Chapter 19, Section 19.4.2 of this volume for further discussion). The inhibition reflects competition for resources, but is not modeled directly.

It will often be simplest to construct mechanistic models of complex phenomena in terms of simpler point process models whose properties are well understood, and which are used as “building blocks.” The definitions of these simple models may reflect either particular generic mechanisms for generating events (as in the hypothetical seedling example above),