POINT PROCESSES

Frederic Paik Schoenberg

UCLA Department of Statistics, MS 8148

Los Angeles, CA 90095-1554

frederic@stat.ucla.edu

July 2000

A point process is a random collection of points falling in some space. In most applications, each point represents the time and/or location of an event. Examples of events include incidence of disease, sightings or births of a species, or the occurrences of fires, earthquakes, lightning strikes, tsunamis, or volcanic eruptions. When modeling purely temporal data, the space in which the points fall is simply a portion of the real line (Figure 1). Increasingly, spatial-temporal point processes are used to describe environmental processes; in such instances each point represents the time *and* location of an event in a spatial-temporal region (Figure 2).

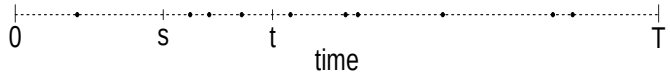

Figure 1: Temporal point process

DEFINITIONS

There are several ways of characterizing a point process. The mathematically-favored approach is to define a point process N as a random measure on a space S taking values in the non-negative integers $\mathbf{Z}^+$ (or infinity). In this framework the measure $N(A)$ represents the number of points falling in the

subset A of S. Attention is typically restricted to random measures that are finite on any compact subset of S, and to the case where S is a complete separable metric space (e.g. $\mathbf{R}^k$).

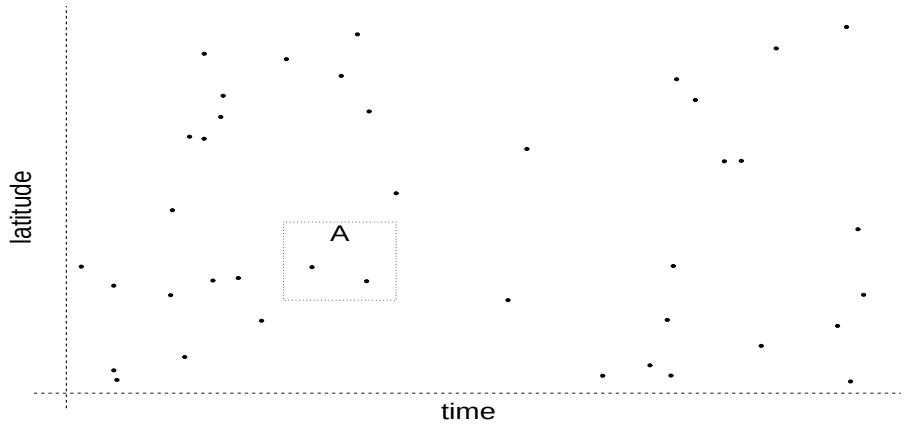

Figure 2: Spatial-temporal point process

For instance, suppose N is a temporal point process. For any times s and t, the measure N assigns a value to the interval (s, t). The value assigned by N is the number of points occurring between time s and time t; in Figure 1, this number is 3. For the set A in Figure 2, the value of $N(A)$ is 2.

The random measure definition above is perhaps not the most intuitive means of characterizing a point process, and several alternatives exist. Con-

sider the case of a temporal point process, e.g. the times of events occurring between time 0 and time T. One may characterize N as an ordered list $\{\tau_1, \tau_2, ..., \tau_n\}$ of event times. One may alternatively convey the equivalent information about N via the interevent times $\{u_1, u_2, ..., u_n\}$, where $\tau_0 = 0$ and $u_i = \tau_i - \tau_{i-1}$.

N may alternatively be described by a counting process $N(t)$, where for any t between 0 and T, $N(t)$ is the number of points occurring at or before time t. Note that this process $N(t)$ must be non-decreasing and right continuous, and take only non-negative integer values. One could thus define a temporal point process as any non-decreasing, right-continuous $\mathbf{Z}^+$-valued process. This sort of definition is used by Jacod (1975), Brémaud (1981), Andersen et al. (1993), and others.

The counting process definition of a point process lends itself naturally to discussions of martingales and distributional theory. On the other hand, the random measure formulation has the advantage of generalizing immediately to point processes in higher dimensions and in abstract spaces, and is thus

preferred.

Within the random measure framework, it is often convenient to note that a realization of a point process N can be written as the sum of Dirac delta measures δ_{τ_i}, where for any measurable set A, $\delta_{\tau_i}(A) = 1$ if A contains the point τ_i, and $\delta_{\tau_i}(A) = 0$ otherwise. In addition, one often works with integrals with respect to dN: $\int_A dN$ is simply the number of points in the set A, and for any function f on A, $\int_A f dN$ is defined simply as the sum $\sum_{i:\tau_i \in A} f(\tau_i)$.

Traditionally the points of a point process are thought to be indistinguishable, other than by their times and/or locations. Often, however, there is other important information to be stored along with each point. For example, one may wish to analyze a list of times of volcanic eruptions along with the sizes of the eruptions, or a catalog of arrival times of hurricanes along with the amounts of damage attributed to each. Such processes may be viewed as *marked* point processes, i.e. a random collection of points, where each point has associated with it a random variable, or *mark*.

The relationship between time series and point processes is worth noting. Many environmental datasets that are traditionally viewed as realizations of (marked) point processes could in principle also be regarded as time series, and vice versa. For instance, a sequence of earthquake origin times is typically viewed as a temporal point process; however, one could also store such a sequence as a time series consisting of zeros and ones, with the ones corresponding to earthquakes. Such a representation would typically be extremely cumbersome, requiring numerous zeros to store limited information. The time series representation is useful for processes taking on different values on a lattice of different time points; the point process representation is preferable for processes that take on values only at certain selected times, and where these time points may be anywhere on the real line, not necessarily on a lattice.

BASIC CLASSES OF POINT PROCESSES

Of the adjectives used to describe point processes, two of the most fundamental describe processes whose points are well-dispersed. A point process is

called *simple* if with probability one, all its points $\{\tau_i\}$ are distinct, i.e. $\tau_i \neq \tau_j$ for $i \neq j$. A point process N is *orderly* if for any t, $\frac{1}{\Delta t} P\{N[t, t+\Delta t] > 1\} \to 0$ as $\Delta t \to 0$.

A point process N may be called *self-exciting* if $\text{cov}\{N(A), N(B)\} > 0$ for any two adjacent disjoint sets A and B in S; N is *self-correcting* if instead this covariance is always negative. Thus the occurrence of points in a self-exciting point process causes other points to be more likely to occur, whereas in a self-correcting process, the points have an inhibitory effect.

The most important type of point process is the *Poisson process*, which is neither self-exciting nor self-correcting. The *Poisson* process is defined as a simple point process N such that the number of points in any set follows a Poisson distribution and the numbers of points in disjoint sets are independent. That is, N is a *Poisson* process if $N(A_1), \ldots, N(A_n)$ are independent Poisson random variables, for any disjoint, measurable subsets $A_1, \ldots, A_n$ of S.

The behavior of a temporal point process N is typically modelled by specifying its *conditional rate process* λ. λ may be thought of as the *frequency* with which events are expected to occur around a particular point in time, conditional on the prior history of the point process. In the statistical literature, λ is more commonly referred to as the conditional *intensity* rather than conditional *rate*. However, the term *intensity* is also used in various environmental sciences, e.g. in describing the size or destructiveness of an earthquake, so to avoid confusion, the term *rate* is preferred.

Formally, the conditional rate process λ associated with a temporal point process N may be defined by the limiting conditional expectation

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{E\{N[t, t+\Delta t] | H_t\}}{\Delta t}, \tag{1}$$

provided the limit exists. Here H_t is the entire history of the point process N up to time t. Some authors instead define λ via

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P\{N[t, t+\Delta t] > 0 | H_t\}}{\Delta t}; \tag{2}$$

for orderly point processes the two definitions are equivalent.

$\lambda(t)$ represents the infinitesimal expected rate of events at time t, given the entire history up to time t. As all finite-dimensional distibutions of N may be derived from the conditional rate (see Daley and Vere-Jones, 1988), in modeling N it suffices to set down a model for λ.

Although λ may be estimated nonparametrically (Diggle 1985; Guttorp and Thompson, 1990; Vere-Jones, 1992), it is more common to estimate λ via a parametric model. At present, two models most often used for many temporal environmental processes such as wildfire and earthquake origin times are stationary Poisson and renewal models (Johnson and Gutsell, 1994; Kagan, 1997). These models are described below.

In general, λ depends not only on t but also on the times τ_i of preceding events. When N is a Poisson process, however, λ is deterministic; i.e. $\lambda(t)$ depends only on t. A *stationary* Poisson process has constant conditional rate: $\lambda(t) = \alpha$, for all t. In the case of modeling environmental disturbances, this model incorporates the idea that the risk of an event is the same at

all times, regardless of where and how frequently such disturbances have occurred previously.

Another important elementary type of temporal point process is the *renewal* process. A renewal process on the line is a simple point process such that the interevent times $\{u_1, u_2, ..., u_n\}$ are independent (typically i.i.d.) random variables. In the i.i.d. case, the density function governing each interevent time is called the *renewal density*. For example, Nadeau et al. (1995) analyze clusters of Parkfield microearthquakes by modeling them as renewal processes with log-normal renewal density.

For a temporal renewal process with density f, the conditional rate is given by $\lambda(t) = s(t - \tilde{t})$, where $\tilde{t}$ is the time of the most recent event prior to time t, and $s(t)$ is the survivor function corresponding to f. That is, $s(t) = \frac{f(t)}{1-F(t)}$, where $F(t) = \int_0^t f(u)du$ is the cumulative distribution function corresponding to f. Note that f is ordinarily taken to be a density on the half-line, i.e. $f(t) = 0$ for $t < 0$.

Renewal models embody the notion that the hazard of an event occurring at a particular time depends only on the time since the most recent event. In fire hazard analysis, for example, such a model is consistent with the theory of fuel loading followed by complete fuel depletion in the event of a fire.

Self-exciting point process models are often used in epidemiology and seismology to model events that are clustered together in time. A commonly used example is the Hawkes model, where λ is given by

$$\lambda(t) = \mu(t) + \int_0^t \nu(t-u)dN(u), \tag{3}$$

the functions $\mu(t)$ and $\nu(t)$ representing the background rate and clustering density, respectively. An example where $\mu(t)$ is constant is offered by Hawkes and Adamopoulos (1973). Two forms of the clustering density ν that are used in modeling earthquakes (Ogata, 1988) are:

$$\nu(t) = \frac{\kappa}{(t+\phi)^\theta} \tag{4}$$

and

$$\nu(t) = \sum_{k=1}^{K} \phi_k t^{k-1} e^{-\theta t}, \tag{5}$$

which correspond with power-law and exponential decay in the clustering be-

havior over time.

Self-correcting models are used in ecology, forestry and other fields to model occurrences that are well-dispersed. Such models may be useful in describing births of species, for example, or in seismology for modeling earthquake catalogs after aftershocks have been removed (Vere-Jones, 1978). An important example is the Markovian point process model (so called because the conditional rate obeys the Markov property) where

$$\lambda(t) = f\{t, N[0,t)\}. \tag{6}$$

For example, the function f may be selected so that λ takes the form

$$\lambda(t) = exp\{\alpha + \beta(t - \rho N[0,t))\}, \tag{7}$$

where α, β and ρ are constants (e.g. Isham and Westcott, 1979; Ogata and Vere-Jones, 1984; Vere-Jones and Ogata, 1984). The parameters α and β govern the background rate and trend in the occurrences, while the product $\beta\rho$ represents the decrease in conditional rate of future events caused by each event, perhaps due to diminished fuel load in the case of wildfire modeling, or the release of strain energy in the seismological case.

The above models are used to describe purely temporal behavior of point processes. When N is a marked point process, e.g. information associated with the character or magnitude of each event is available, the marks may provide additional information and therefore the formulae for λ may be adjusted. For instance, Ogata (1988) suggests replacing the clustering density $\nu(t)$ by a function $\nu(t, m_i)$ which is a function not only of time t but also of the mark m_i associated with event i.

The definition of the conditional rate process can be extended to the more general case of a spatial-temporal marked point process. Unlike the one-dimensional case where it is most natural to condition on the *past*, various natural choices exist for conditioning in the spatial case; see Merzbach and Nualart (1986) or Schoenberg (1999) for details.

The spatial-temporal Poisson process may be useful to model processes without spatial or temporal interactions, i.e. processes for which the oc-

currence of an event at one location and time does not influence the occurrence of future events or events at other locations. Many environmental processes are thought to have important interactions however. In the case where points are clustered spatially and/or temporally, the Hawkes model, whose definition extends immediately to the spatial-temporal case, may be useful. An important example is the Neyman-Scott cluster process (Neyman and Scott, 1958). When points are well-dispersed, multi-dimensional self-correcting models may be useful; see Ogata and Tanemura (1986) for an application to Japanese black pine saplings and seedlings, where each tree appears to have an inhibitory effect on the growth of nearby trees.

The definition of the renewal process may also be extended to the plane; see Hunter (1974) or Daley and Vere-Jones (1988). Despite the popularity of renewal models for temporal point processes, two-dimensional renewal models appear to have been sparsely (if ever) used in the analysis of planar point process data.

TRANSFORMATIONS

Some important operations on point processes include superposition, thinning, and rescaling. Graphically, these operations correspond, respectively, to overlaying points of one process onto the plot of another point process, deleting certain points, and stretching out an axis of a point process.

Mathematically, superposition of point processes corresponds to addition, i.e. N_3 is the superposition of point processes N_1 and N_2 if $N_3(A) = N_1(A) + N_2(A)$ for any measurable set A in S. In the case of thinning, each point τ_i of N is deleted with some probability p_i; in the simple case all points are eliminated with equal probability ($p_i = p$). A temporal point process N is *rescaled* by a factor of α to form the process M if $N(0,t) = M(0,\alpha t)$ for all t.

The Poisson process frequently arises as a limiting process resulting from transformations of point process. For example, Palm (1943) showed that under quite general conditions, when k independent copies of a temporal point process are superposed and rescaled by a factor of k, the resulting process

converges to a Poisson process (for details see Kallenberg, 1983, or Daley and Vere-Jones, 1988). Similar results can be obtained for randomly thinned point processes (Westcott, 1976) or rescaled point processes (see Karr, 1991).

Other important operations involve transforming a point process into an image, e.g. via Dirichlet tessellation [see tessellation] or kernel smoothing.

ESTIMATION AND SIMULATION

The parameter vector θ for a point process model with conditional rate $\lambda(\theta)$ is usually estimated by maximizing the log-likelihood function

$$L(\theta) = \int_S log[\lambda(\theta)]dN - \int_S \lambda(\theta)d\mu, \tag{8}$$

with μ typically Lebesgue measure on S.

For example, for a temporal point process observed from time 0 to time T, one maximizes the log-likelihood function

$$L(\theta) = \int_0^T log[\lambda(t;\theta)]dN(t) - \int_0^T \lambda(t;\theta)dt. \tag{9}$$

The consistency, asymptotic normality, and efficiency of the maximum likelihood estimators, under various conditions, have been established (Brillinger, 1975; Ogata, 1978). Standard errors may be obtained via the Hessian of the log-likelihood, under certain limitations (Kutoyants, 1984; Dzhaparidze, 1985; Rathbun and Cressie, 1994).

Alternatively, simulations may be useful for obtaining approximate standard errors and for other types of inference. Lewis and Shedler (1979) and Ogata (1981) devised an effective simulation method for point processes based on random thinning theory. The procedure, which works for point processes whose conditional rate λ is bounded (or locally bounded), involves simulating a (locally) stationary Poisson process and thinning it, keeping each point τ_i with probability $\lambda(\tau_i)$.

MODEL EVALUATION

A useful technique for evaluating point process models is via random rescaling. The method essentially involves rescaling the observed point process N at time t by a factor of $\hat{\lambda}(t)$, where $\hat{\lambda}$ is an estimate of the conditional rate λ. The resulting process M is a stationary Poisson process with unit rate, under quite general conditions (Meyer, 1971; Brémaud, 1972; Papangelou, 1972; Aalen, 1975). The technique has been extended to the multi-dimensional case (Merzbach and Nualart, 1986; Nair, 1990; Schoenberg, 1999).

It can be shown that the rescaled process M is Poisson with unit rate *if and only if* the model is correct; i.e. $\hat{\lambda} = \lambda$ almost everywhere with probability one (Schoenberg, 1999). Hence a useful method for assessing the fit of a point process model is to examine whether the *rescaled* point process looks like a Poisson process with unit rate. Several tests exist for this purpose, with different uses depending on the alternative hypotheses (e.g. Saw, 1975; Diggle, 1979; Dijkstra et al., 1984; Lisek and Lisek, 1985; Lawson, 1988; Arsham, 1987; Andersen et al., 1993). Some of the most useful are tests based on second and higher order properties (Bartlett, 1964; Davies, 1977; Ripley,

1979; Heinrich, 1991).

References

[1] Aalen, O. (1975). *Statistical Inference for a Family of Counting Processes.* Ph.D. thesis, Statistics Department, University of California, Berkeley.

[2] Andersen, P., Borgan, O., Gill, R., and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

[3] Arsham, H. (1987). A modified one-sided K-S confidence region narrower in the tail. *Comm. Stat. A* **16**, 17-28.

[4] Bartlett, M. (1964). The spectral analysis of two-dimensional point processes. *Biometrika* **51**, 299-311.

[5] Brémaud, P. (1972). *A Martingale Approach to Point Processes.* Ph.D. thesis, Electrical Engineering Department, University of California, Berkeley.

[6] Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

[7] Brillinger, D. (1975). Statistical inference for stationary point processes. *Stochastic Processes and Related Topics* **1**, 55-100.

[8] Daley, D., and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes.* Springer, Berlin.

[9] Davies, R. (1977). Testing the hypothesis that a point process is Poisson. *Adv. Appl. Probab.* **9**, 724-746.

[10] Diggle, P. (1979). On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrics* **35**, 87-101.

[11] Diggle, P. (1985). A kernel method for smoothing point process data *Appl. Stat.* **34**, 138-147.

[12] Dijkstra, J., Rietjens, T. and Steutel, F. (1984). A simple test for uniformity. *Statistica Neerlandica* **38**, 33-44.

[13] Dzhaparidze, K. (1985). On asymptotic inference about intensity parameters of a counting process. *Proc. 45th Session Int. Stat. Inst.* **4** (23.2), 1-15.

[14] Guttorp, P., and Thompson, M. (1990). Nonparametric estimation of intensities for sampled counting processes. *J. Royal Stat. Soc. B* **52**, 157-173.

[15] Hawkes, A., and Adamopoulos, L. (1973). Cluster models for earthquakes – regional comparisons. *Bull. Int. Statist. Inst.* **45** (3), 454-461.

[16] Heinrich, L. (1991). Goodness-of-fit tests for the second moment function of a stationary multidimensional Poisson process. *Statistics* **22**, 245-278.

[17] Hunter, J. (1974). Renewal theory in two dimensions: basic results. *Adv. Appl. Probab.* **6**, 376-391.

[18] Isham, V., and Westcott, M. (1979). A self-correcting point process. *Stoch. Proc. Appl.* **8**, 335-347.

[19] Jacod, J. (1975). Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31**, 235-253.

[20] Johnson, E. and Gutsell, S. (1994). Fire frequency models, methods and interpretations. *Advances in Ecological Research* **25**, 239-287.

[21] Kagan, Y. (1997). Statistical aspects of Parkfield earthquake sequence and Parkfield prediction experiment. *Tectonophysics* **270**, 207-219.

[22] Kallenberg, O. (1983). *Random Measures, third ed.* Akademie-Verlag, Berlin.

[23] Karr, A. (1991). *Point Processes and Their Statistical Inference, second ed.* Dekker, NY.

[24] Kutoyants, Y. (1984). Parameter Estimation for Stochastic Processes. Helderman, Berlin.

[25] Lawson, A. (1988). On tests for spatial trend in a non-homogeneous Poisson process. *J. Appl. Stat.* **15**, 225-234.

[26] Lewis, P., and Shedler, G. (1979). Simulation of nonhomogeneous Poisson

processes by thinning. *Naval Research Logistics Quarterly* **26**, 403-413.

[27] Lisek, B., and Lisek, M. (1985). A new method for testing whether a point process is Poisson. *Statistics* **16**, 445-450.

[28] Merzbach, E., and Nualart, D. (1986). A characterization of the spatial Poisson process and changing time. *Ann. Probab.* **14**, 1380-1390.

[29] Meyer, P. (1971). Démonstration simplifée d'un théorème de Knight. *Lecture Notes in Mathematics* **191**, 191-195.

[30] Nadeau, R., Foxall, W. and McEvilly, T. (1995). Clustering and periodic recurrence of microearthquakes on the San Andreas Fault at Parkfield, California. *Science* **267**, 503-507.

[31] Nair, M. (1990). Random space change for multiparameter point processes. *Ann. Probab.* **18**, 1222-1231.

[32] Neyman, J., and Scott, E. (1958). Statistical approach to problems in cosmology (with discussion). *J. Royal Statist. Soc. B* **20**, 1-43.

[33] Ogata, Y. (1978). The asymptotic behavior of maximum likelihood estimators for stationary point processes. *Ann. Inst. Stat. Math.* **30**, 243-262.

[34] Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Trans. Inf. Theory* **IT-27**, 23-31.

[35] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.* **83**, 9-27.

[36] Ogata, Y., and Tanemura, M. (1986). Likelihood estimation of interaction potentials and external fields of inhomogeneous spatial point patterns. In *Pacific Statistical Congress*, I. Francis, B. Manly, and F. Lam, eds. Elsevier Science, North-Holland, pp. 150-154.

[37] Ogata, Y., and Vere-Jones, D. (1984). Inference for earthquake models: a self-correcting model. *Stoch. Proc. Appl.* **17**, 337-347.

[38] Palm, C. (1943). Intensitätsschwankungen im Fernsprechverkehr. *Ericsson Techniks* **44**, 1-189.

[39] Papangelou, F. (1972). Integrability of expected increments of point processes and a related random change of scale. *Trans. Amer. Math. Soc.* **165**, 483-506.

[40] Rathbun, S., and Cressie, N. (1994). Asymptotic properties of estimators for the parameters of spatial inhomogeneous Poisson point processes. *Adv. Appl. Prob.* **26**, 122-154.

[41] Ripley, B. (1979). Tests of 'randomness' for spatial point patterns. *J. Royal Stat. Soc. B* **41**, 368-374.

[42] Schoenberg, F. (1999). Transforming spatial point processes into Poisson processes. *Stoch. Proc. Appl.* **81**, 155-164.

[43] Vere-Jones, D. (1978). Earthquake predicton – a statistician's view. *J. Phys. Earth* **26**, 129-146.

[44] Vere-Jones, D. (1992). Statistical methods for the description and display of earthquake catalogs. In *Statistics in the Environmental and Earth Sciences*, A. Walden and P. Guttorp, eds. Edward Arnold, London, pp. 220-246.

[45] Vere-Jones, D., and Ogata, Y. (1984). On the moments of a self-correcting process. *J. Appl. Prob.* **21**, 335-342.

[46] Westcott, M. (1976). Simple proof of a result on thinned point processes. *Ann. Probab.* **4**, 89-90.