Journal of Environmental Statistics
February 2010, Volume 1, Issue 2. http://www.jenvstat.org

Diagnostics for Nonparametric Estimation in Space-Time Seismic Processes

Giada Adelfio
Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli"
University of Palermo

Marcello Chiodi
Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli"
University of Palermo

Abstract

In this paper we propose a nonparametric method, based on locally variable bandwidths kernel estimators, to describe the space-time variation of seismic activity of a region of Southern California. The flexible estimation approach is introduced together with a diagnostic method for space-time point process, based on the interpretation of some second-order statistics, to analyze the dependence structure of observed data and suggest directions for fit improvement. In this paper we review a diagnostic method for space-time point processes based on the interpretation of the transformed version of some second-order statistics. The method is useful to analyze dependence structures of observed data and suggests directions for fit improvement also when a more flexible estimator of the conditional intensity function is used, such as the kernel one. In particular locally variable bandwidths kernel estimators are used to describe space-time variations of seismic activity of a region of Southern California.

Keywords: Point process, second-order statistics, residual analysis, kernel estimator, seismic process.

1. Introduction

The definition of effective stochastic models to describe adequately the seismic activity of a fixed area is of great interest in seismology, since a reliable description of earthquake occurrence might suggest useful ideas on the mechanism of such a complex phenomenon.

Earthquake catalogs are conventionally analyzed as point processes; in this context second-order statistics may have a relevant role in the study and the comprehension of a seismic process and its realization, since features like self-similarity, long-range dependence and fractal dimension need often to be analyzed for a more realistic description of seismicity.

The diagnostic method introduced in Adelfio and Schoenberg (2009) is here used to assess the fitting of a space-time non-homogeneous Poisson process estimated by kernel smoothing.

The residual analysis approach proposed in this paper is based on the interpretation of a transformed version of second-order statistics obtained by weighting data by a quantity proportional to the inverse of the conditional intensity function. As shown in this paper, this diagnostic tool enables one to detect properties of clustering and inhibition of observed data and therefore may suggest directions for fitting, even if the conditional intensity function of the model is estimated by kernel estimators. Indeed, because of their particularly adaptive properties to anomalous behavior in data, kernel density estimators with variable bandwidths (i.e. locally variable smoothing parameters) are here used to interpret dependence features of seismic activity of a given area of Southern California.

For a better exposition of the application results of the diagnostic method in this context, some formal definitions are reported. Point processes and the conditional intensity function are reviewed in section 2. The diagnostic method is introduced in section 3; the base of the model used to describe the seismic activity of the analyzed area is reviewed in section 4. In section 5 the description of nonparametric estimation procedures for point processes and the use of the proposed diagnostic approach are shown. Conclusions are provided in section 6.

2. Point processes and conditional intensity function

Point processes are well studied objects in probability theory and a powerful tool in statistics for modeling and analyzing the distribution of real phenomena, such as seismicity, the study of epidemic diffusions, the spatial distribution of some pathologies as a function of pollution sources location, etc. Point processes can be specified mathematically in several ways, for instance, by considering the joint distributions of the counts of points in arbitrary sets or defining a complete intensity function. For temporal point processes with no simultaneous occurrences, an alternative specification considers the joint distributions of intervals between successive points starting from a fixed origin. These point processes are frequently used as models for random events in time, such as the arrival of customers in a queue (queueing theory).

More formally, a point process is a random point patterns on a set $X \subseteqq \mathbb{R}^d, d \geq 1$. For a mathematical definition of a point pattern specified as a locally finite counting measure with positive values in $\mathbb{Z}$ see Daley and Vere-Jones (2003).

Let N be a point process on a spatial-temporal domain $X \subseteq \mathbb{R}^d \times \mathbb{R}_+$ and $\ell(x)$ be the Lebesgue measure of x; its conditional intensity function is defined by:

$$\lambda(t, \mathbf{s}|\mathcal{H}_t) = \lim_{\ell(\delta t, \delta \mathbf{s}) \to 0} \frac{E[N([t, t+\delta t) \times [\mathbf{s}, \mathbf{s}+\delta \mathbf{s})|\mathcal{H}_t)]}{\ell(\delta t \delta \mathbf{s})} \tag{1}$$

where $\mathcal{H}_t$ is the space-time occurrence history of the process up to time t, i.e. the σ-algebra of events occurring at times up to but not including t; $\delta t, \delta \mathbf{s}$ are time and space increments respectively, and $E[N([t, t+\delta t) \times [\mathbf{s}, \mathbf{s}+\delta \boldsymbol{s})|\mathcal{H}_t)]$ is the history-dependent expected value of occurrence in the volume $\{[t, t+\delta t) \times [\mathbf{s}, \mathbf{s}+\delta \mathbf{s})\}$. The conditional intensity function is a function of the point history and it is itself a stochastic process depending on the past up to time t.

Assuming that such a limit exists for each point $(t, \mathbf{s})$ in the space-time domain and that the point process N is simple, the conditional intensity process uniquely characterizes the finite-dimensional distributions of the point process Daley and Vere-Jones (2003).

If the conditional intensity function is independent of the past history, but dependent only on the current time and the spatial locations, (1) identifies an inhomogeneous Poisson process. A process with constant conditional intensity provides a stationary Poisson process.

3. Second-order residual analysis

A major difficulty in residual analysis for point processes, is to find a good definition of residuals.

Some approaches require the transformation of data into a residual point pattern (result of a thinning or rescaling procedure (Meyer 1971)) and the use of tests to assess the consistency of the residuals with the homogeneous Poisson process (Schoenberg 2003). The residuals ought to be a homogeneous Poisson process, if the model used for the transformation is indeed the model that generated the data. Although literature provides several approaches to assess the consistence of observed points with homogeneous Poisson processes (Diggle (1983), Bartlett (1964), Ripley (1976), Baddeley (1984), Heinrich (1991)), these kinds of approaches often constitute only a starting point of a more complex analysis, necessary to build and fit a more realistic process. An alternative is represented by the weighted measures here introduced. References with earlier work are included in Adelfio and Schoenberg (2009).

The proposed method deals with the definition and the interpretation of a weighted version of second-order statistics (such as autocorrelation, K-function, spectrum, fractal dimension and R/S statistic), just weighting the contribution of each observed point by the inverse of the conditional intensity function that identifies the process.

Weighted second-order statistics, and hence second-order residuals, are so obtained.

Second-order residual analysis directly applies to data, since its definition does not require a homogeneity assumption or previous transformation of data into residuals. Therefore, if we assume as known the conditional intensity function, or at least a good (ideally consistent) estimate be available, second-order weighted measures overcome some limitations that characterize the rescaling or the thinning procedures. Indeed, the probabilities used in the thinning method to retain points are replaced here by weights in order to offset the inhomogeneity of the process including all the observed points as in Zhuang, Ogata, and Vere-Jones(2002).

The proposed diagnostic measures can be applied to processes of any dimension, assuming that the statistics here discussed can be computed and interpreted to analyze features such as clustering and inhibition.

Some distributional theoretical results related to the weighted version of some second-order statistics are provided in Adelfio and Schoenberg (2009), by an easily generalizable approach based on the theory of martingales. In that work, it is shown that some weighted second-order statistics behave as the corresponding ones (not weighted) of a homogeneous Poisson process.

Second-order residual analysis is here applied to give directions for a better space-time seismic model definition, by using nonparametric estimation methods. Some definitions and theoretical results relative to the second-order residual analysis are reviewed in Appendix A. Examples of applications to seismic data are also provided in Adelfio and Chiodi (2009).

4. A point process describing seismic occurrence

Earthquake sequences may be considered as a stochastic point process in a multidimensional space, since each earthquake may be identified by a point in space (epicentral coordinates), in time (origin time) and magnitude domain (where magnitude is a logarithmic measure of earthquake strength). The seismic activity is traditionally considered as the superposition of two different components: the background seismicity and the space-time clustered one, induced by main earthquakes. Earthquake sequences are schematically composed of mainshocks and events that could occur before and after the mainshock, named foreshocks and aftershocks, respectively; isolated events constitute the background.

Although most earthquake catalogs exhibit a variability character as their dominant statistical feature, it is not possible to study statistical properties of a random process without a general model of the process, that includes, for instance, considerations on stationarity or space-time interactions.

Indeed a number of statistical models have been proposed for representing the intensity function of earthquakes. Simpler models assume that earthquakes occur in space and time according to a stationary point process, such that conditional rate becomes a constant. In seismology, however, the stationarity hypothesis might be acceptable only with respect to time and for main earthquakes, after aftershocks are removed, because epicenters usually display a substantial degree of spatial heterogeneity and clustering. On the other hand, univariate analysis is not informative about multivariate seismic processes. For these reasons, the description of seismic events often requires the definition of more complex models than stationary Poisson processes and the relaxation of the assumption of statistical independence of earthquakes (Adelfio Chiodi, De Luca, and Luzio 2006). Therefore, second-order properties of point processes may have a relevant role in the study and the comprehension of the seismic process and its realization; second-order moments may give us valuable information about the interdependence among earthquakes, and constitute a first step beyond a Poissonian model in which the events are assumed to be statistically independent.

Hence, when aggregation is present, generalizations of the simple Poisson process, such as self-exciting point processes, are useful to model events that are clustered together, and self-correcting processes when regularities are observed, e.g. the strain-release model (Daley and Vere-Jones 2003). A widely used model is the epidemic type aftershocks-sequences (ETAS) model (Ogata (1988),Ogata (1998)) , that is a self-exciting point process, describing earthquake activity in a given region during a period of time, through a branching structure.

4.1. Space-time ETAS model

The ETAS model is completely characterized by its conditional intensity function, which provides a quantitative evaluation of the future seismic activity and is proportional to the probability that an event with magnitude m will take place at time t, with epicenter of coordinates (x, y), that is a position on the surface of the earth expressed in latitude and longitude. It is defined as the sum of a term describing spontaneous activity (background) and one relative to the induced seismicity. The main hypothesis of the model states that all events, both mainshocks or aftershocks, have the possibility of possessing offspring. The

ETAS conditional intensity function is:

$$\lambda_{\boldsymbol{\theta}}(x, y, t, m|\mathcal{H}_t) = J(m)\left[\mu(x, y) + \sum_{t_j<t} \nu_{\boldsymbol{\theta}}(t - t_j, x - x_j, y - y_j|m)\right] \tag{2}$$

with $\nu_{\boldsymbol{\theta}}(t-t_j, x-x_j, y-y_j|m) = g(t-t_j; \boldsymbol{\theta})f(x-x_j, y-y_j|m, \boldsymbol{\theta})$ and $\boldsymbol{\theta}$ a vector of parameters. The ETAS model is a special case of a marked Hawkes process, with marks given by the magnitude values, such that $m \in \mathcal{M} = \mathbb{R}$; the marks distribution $J(m)$ corresponds to the Gutenberg-Richter frequency magnitude law (Gutenberg and Richter 1944).

In the ETAS model, the background seismicity $\mu(x, y)$ is assumed to be stationary in time, while the temporal distribution of the seismicity triggered by previous events is modeled as a non-stationary Poisson process according to the modified Omori's formula (Utsu 1961). In this model, the occurrence rate of aftershocks at time t following the earthquake of time τ, is described by:

$$g(t - \tau) = \frac{K}{(t - \tau + c)^p}, \quad \text{with} \quad t > \tau \tag{3}$$

with K a normalizing constant, c and p characteristic parameters of the seismic activity of the given region; p is useful for characterizing the pattern of seismicity, indicating the decay rate of aftershocks in time.

To describe the spatial clustering of aftershocks, conditioned on the magnitude of the generating event, Ogata (1998) suggests the use of the following spatial density function:

$$f(x - x_j, y - y_j|m_j) = \left\{\frac{(x - x_j)^2 + (y - y_j)^2}{e^{\alpha(m_j - m_0)}} + d\right\}^{-q}$$

where α measures the influence on the relative weight of each sequence, m_0 is the completeness threshold of magnitude, i.e. the lower bound for which earthquakes with higher values of magnitude are surely recorded in the catalog, d and q are two parameters related to the spatial influence of the mainshock.

5. Nonparametric estimation and diagnostics

In seismic models, parametric estimation suffers from many drawbacks related to the definition of a reliable mathematical model from the geophysical theory and to the sensitivity of statistical estimates to the composition of the sample, that is the space-time region under study. Many of the disadvantages of the parametric modeling can be avoided by making use of nonparametric techniques, such as kernel density methods (Silverman 1986). Therefore a flexible procedure based on kernel estimators is proposed, useful when a clear discrimination between principal and secondary events is not available.

The kernel estimator of an unknown density f in $\mathbb{R}^d$ is defined as:

$$\hat{f}(z_1, \ldots, z_d) = \frac{1}{nh_{z_1} \cdots h_{z_d}} \sum_{i=1}^{n} K\left(\frac{z_1 - z_{i1}}{h_{z_1}}, \cdots, \frac{z_d - z_{id}}{h_{z_d}}\right)$$

where $K(z_1, \ldots, z_d)$ denotes a multivariate kernel function operating on d arguments $(z_1, \ldots, z_d)$ centred at $(z_{i1}, \ldots, z_{id})$ and $(h_{z_1}, \ldots, h_{z_d})$ are the smoothing parameters (bandwidths) of kernel functions.

The kernel estimator can be obtained as the sum of n surfaces, each corresponding to a single point: the kernel function, that is a density function (usually Gaussian), defines the shape of these surfaces.

The application of this technique requires the parameters $\boldsymbol{h} = (h_{z_1}, \ldots, h_{z_d})$ to be optimized in order to achieve the best compromise according to some criteria between the surface smoothness and the level of detail in the phenomenon representation.

In Adelfio, Chiodi, De Luca, Luzio, and Vitale (2006b) the seismicity of the Southern Tyrrhenian Sea is described by the use of Gaussian kernels and the optimum value of $\boldsymbol{h}$ is chosen such as to minimize the mean integrated square error (MISE) of the estimator $\hat{f}(\cdot)$. In particular the authors used the value $\boldsymbol{h}_{opt}$ that Silverman (1986) obtained minimizing the MISE of $\hat{f}(\cdot)$ assuming multivariate normality:

$$\boldsymbol{h}_{opt} = A(K)n^{\frac{-1}{d+4}} \tag{4}$$

with $A(K)$ a function of the used kernel density. This is referred to as Silverman's rule.

In order to get a more realistic and sensitive estimation, in this paper locally variable smoothing parameters are estimated. The variable bandwidths estimation procedure is described by the main following steps:

1. let $z_1, \ldots, z_d$ be the d arguments of the multivariate kernel function centered at $z_{i1}, \ldots, z_{id}$, $i = 1, \ldots, n$;

2. choose a suitable integer between 10 and 100 for the parameter n_p p (Zhuang et al. 2002).

3. For each event calculate a bandwidth value

$$\boldsymbol{h}_i = (h^i_{z_1}, \ldots, h^i_{z_d}) \tag{5}$$

as the radius of the smallest circle centered at the location of the ith event $(z_{i1}, \ldots, z_{id})$ that includes at least n_p other events.

5.1. Details of the application

In this work second-order residuals are used to get a better description of the seismic activity of Southern California occurring between January 1984 and June 2004 in a rectangular area around Los Angeles, between longitude -122 and -114 and latitude 32 and 37; the dataset consists of 2030 earthquakes with magnitude not smaller than 3.5; the estimated fractal dimension for the whole set, based on the correlation integral estimator (Grassberger and Procaccia 1983), is 1.4729. (fig. 1). The space-time kernel intensity estimator is defined by the superposition of the separable kernel densities, as in Adelfio and Ogata (2010):

$$\lambda(t, x, y) = \sum_{i=1}^{n} \phi(x - x_i, y - y_i)\gamma(t - t_i)$$

with ϕ and γ spatial and temporal probability density functions, respectively. Therefore, the complex estimation issue of a self-exciting process now reduces to the simple estimation of the

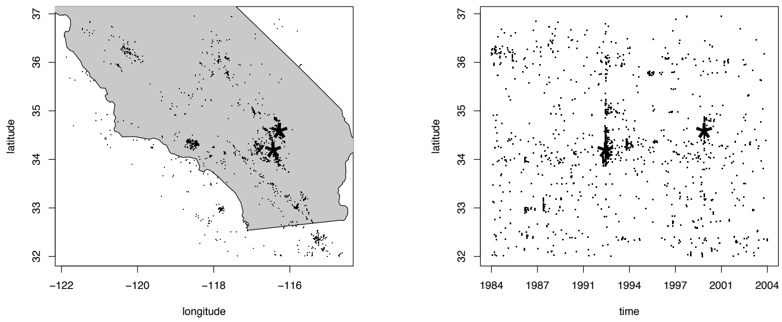

Figure 1: Earthquakes (on the left epicenters and on the right latitude versus time plot) occurred in the east coast of the of Southern California, in the region defined by $32° \sim 37°$ N and $-122° \sim -114°$ E for all depth and for the time span 1984-2004, with magnitude equal or larger than 3.5. (Big-stars: $m \geq 7$).

intensity function of an inhomogeneous Poisson process identified by a space-time Gaussian kernel density, assuming independence between spatial and temporal densities.

As introduced in the previous section, the choice of an appropriate bandwidth could be critical in intensity estimation issues. Indeed, too high bandwidths may provide very smooth estimates, blurring local clustered features of the observed seismicity. On the other hand, too low bandwidths tend to preserve noisy seismicity features.

In this section we compare two estimates of the seismicity rate by using both a global (invariant in the analyzed region) and a local bandwidth approach. First a space-time kernel density estimate of the seismic activity with constant bandwidth values in space and time (see eq. (4)) and Gaussian kernel functions (denoted by approach a) is estimated.

Afterwards, a nonparametric kernel density with variable bandwidth values (5) and Gaussian kernel is estimated (denoted by approach b). As a direction for future work we are considering the possibility of using asymmetric kernel estimator, since the Gaussian one, especially for the temporal distribution of triggered seismicity, might overestimate the seismic activity before strong earthquakes.

In approach b smoothing parameters are obtained as described by (5) fixing $n_p = 10$ (Zhuangetal. 2002).

To study the goodness of fit of the two alternative approaches the weighted spectral density (fig. 2) is analyzed (see Appendix A.1 for more details). Because of the strong clustering nature of data, evident from fig. 1, approach a tends to smooth more than expected. On the other hand the nonparametric model with variable bandwidth values (approach b) seems to account for the residual correlation (due to the presence of aftershocks) not taken into account by approach a.

Looking at the plot on the right in fig. 2 we could say that by fitting locally variable bandwidths kernel estimators, the estimated weighted spectral density behaves as the one of a temporal homogeneous Poisson process.

Moreover if we apply approach b, the estimated weighted correlation dimension (defined in

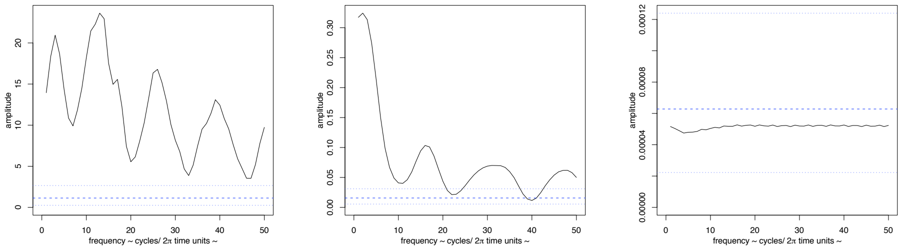

Figure 2: Original periodogram (on the left), weighted periodogram for constant bandwidth approach (approach a; in the middle) and for varying bandwidths model (approach b; on the right), with 95%-bounds of a homogeneous Poisson process (dotted lines).

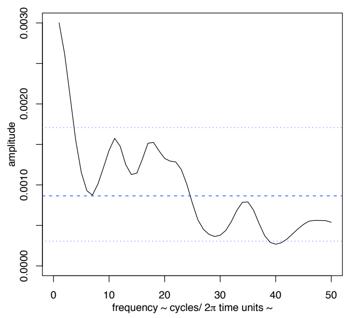

Figure 3: Weighted periodogram for ETAS model, with 95%-bounds of a homogeneous Poisson process (dotted lines).

Appendix A.2) is $D_W = 1.91$, not significatively different from 2, that is the expected value of the correlation dimension for a homogeneous Poisson process in space. For the model with constant bandwidth values the estimated weighted correlation dimension is $D_W = 1.45$. These results suggest that the variable bandwidth model seems to provide a realistic and reliable estimate of the observed seismic activity also for the spatial domain.

Indeed in case of low clustering, a constant smoothing value for all the observed region could provide a valid description of the activity of that area; on the other hand, if events are strongly clustered, the introduction of variable smoothing, accounting for the high production of offspring, seems to be preferable.

It is also interesting to compare the two approaches, and the corresponding estimated intensity functions, with a parametric estimate of ETAS model, obtained by using the method of Zhuang et al. (2002) applied to Southern California data.

The weighted periodogram obtained with the ETAS model is plotted in fig. 3 and shows an inadequate fitting in the temporal domain. Moreover the weighted fractal dimension in space for the ETAS model is $D_W = 1.53$, suggesting the inadequacy of the model for the description of such earthquakes, characterized by strong clustered features, mostly in space.

From the previously provided diagnostic results the variable kernel approach provides a more realistic description of the seismic activity of the studied area both than the constant bandwidth approach and the ETAS model. Indeed approach b seems to follow more adequately the seismic activity of the observed area, characterized by highly variable changes both in space and in time; because of its flexibility, it also provides a better fitting to local space-time changes as suggested by data. Approach a, conversely, is characterized by global bandwidths and generates an over-smooth rate, that is an estimated decreasing activity in the centers of some clusters and an increasing activity in some regions around; as a consequence some clusters of high activity may not be clearly identified.

The ETAS model seems inadequate to describe the observed seismicity. To improve its fitting to data, more complex and parameterized models should be defined. Indeed, the ETAS model imposes some restrictions on the observed seismic process, such as constant p-values of (3) in space, that induce a too quick decay of the intensity function in correspondence of events that are strongly clustered in time. Also the use of constant K-values of (3) may not be reasonable when data are strongly clustered; indeed in this case higher value of K accounting for the high production of offsprings may be worthwhile.

On the other hand, the nonparametric variable-bandwidth kernel density estimator does not constrain the process to have predetermined properties and provides a reasonable characterization of the studied area that presents highly clustered seismicity identified by a complex intensity function.

6. Conclusion

In this paper we compute two estimates of the space-time rate of the seismic activity of Southern California by using both a global and a local smoothing estimation approach. The fit is assessed using the proposed diagnostic approach for point processes and compared with a parametric model. The diagnostic method is based on the definition of a transformed version of some second-order statistics and it requires the estimate of the conditional intensity function of the fitted model to offset the inhomogeneity of the process.

As suggested by the diagnostic approach, an inhomogeneous Poisson process, characterized by a space-time intensity function estimated by a flexible variable-bandwidth kernel estimator, reasonably describes the highly variable and clustered seismic activity of the observed area.

This estimation approach seems helpful in this context, since earthquakes occurrence is characterized by inhomogeneous activity regions, that is by some areas with high clustered activity and others characterized by uniform seismicity; therefore, a different degree of smoothness seems to describe these features better. Moreover it does not constrain the generating process to fit some predetermined properties and shows some computational advantages related to easily interpretable results. In other words, kernel estimators approach provides a simple and effective tool of finding structure in datasets without the imposition of a parametric model, with the advantage of being very intuitive and relatively simple to analyze mathematically. On the other hand, although nonparametric models fit usually better than parametric ones, because their synthesis is founded on the observed data, a purely nonparametric approach does not enable one to estimate and interpret parameters with physical meaning, on the contrary of parametric models like the ETAS one.

Acknowledgments

Our sincerest gratitude to the editors and the two referees for their very useful comments.

References

Adelfio G, Chiodi M (2009). “Second-order diagnostics for space-time point processes with applica- tion to seismic events.” Environmetrics, **20**, 895-911.

Adelfio G, Chiodi M, De Luca L, Luzio D (2006a). “Nonparametric clustering of seismic events.” Data Analysis, Classification and the Forward Search-Proceedings of the Meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society, University of Parma, June 6-8, 2005, pp. 397-404.

Adelfio G, Chiodi M, De Luca L, Luzio D, Vitale M (2006b). “Southern-Tyrrhenian seismicity in space-time-magnitude domain.” Annals of Geophysics, **49**(6), 1245-1257.

Adelfio G, Ogata Y (2010). “Hybrid kernel estimates of space-time earthquake occurrence rates using the Etas model.” Annals of the Institute of Statistical Mathematics, **62**(1), 127-143.

Adelfio G, Schoenberg FP (2009). “Point process diagnostics based on weighted second-order statis- tics and their asymptotic properties.” Annals of the Institute of Statistical Mathematics, **61**(4), 929-948.

Baddeley A, Møller J, Waagepetersen R (2000). “Non and semi-parametric estimation of interaction in inhomogeneous point patterns.” Statistica Neerlandica, **54**(3), 329-350.

Baddeley AJ, Silverman BW (1984). “A cautionary example on the use of second-order methods for analyzing point patterns.” Biometrics, 40, 1089-1093.

Bartlett MS (1964). “The spectral analysis of two dimensional point processes.” Biometrika, **51**(3/4), 299-311.

Daley DJ, Vere-Jones D (2003). An introduction to the theory of point processes. second edition. New York: Springer-Verlag.

Denker M, Keller G (1986). “Rigorous statistical procedures for data from dynamical systems.” Journal of Statistical Physics, **44**, 67-94.

Diggle PJ (1983). Statistical analysis of spatial point patterns. Academic Press Inc., London.
Grassberger P, Procaccia I (1983). “Measuring the strangeness of a strange attractor.” Physica D, **9**, 189-208.

Gutenberg B, Richter CF (1944). “Frequency of earthquakes in California.” Bulletin of the Seismo- logical Society of America, **34**, 185-188.

Heinrich L (1991). “Goodness-of-fit tests for the second moment function of a stationary multidimen- sional Poisson process.” Statistics, **22**, 245-278.

Meyer P (1971). “Démonstration simplifée d’un théorème de Knight.” Lecture Notes in Mathematics, **191**, 191-195.

Ogata Y (1988). “Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes.” Journal of the American Statistical Association, **83**(401), 9-27.

Ogata Y (1998). “Space-time point-process models for earthquake occurrences.” Annals of the Insti- tute of Statistical Mathematics, **50**(2), 379-402.

Ripley BD (1976). "The second-order analysis of stationary point processes." Journal of Applied Probability, **13**(2), 255-266.

Schoenberg FP (2003). "Multi-dimensional residual analysis of point process models for earthquake occurrences." Journal American Statistical Association, **98**(464), 789-795.

Silverman BW (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.

Utsu T (1961). "A statistical Study on the Occurrence of aftershocks." Geophys. Mag., **30**, 521-605.

Zhuang J, Ogata Y, Vere-Jones D (2002). ÒStochastic Declustering of Space-Time Earthquake Occur- rences.Ó Journal of the American Statistical Association, **97**(458), 369-379.

A. Appendix: The weighted process and its second-order characteristics

According to the scheme introduced, to define the residual measures, the definition of weighted process is required. Therefore, let N be a point process defined on $S \in \mathbb{R}^d$, $d \geq 1$. For any point s in S, let $\lambda(s|\mathcal{F})$ be the conditional intensity function of the process with respect to some filtration $\mathcal{F}$ on S, for simplicity denoted by $\lambda(s)$, assumed positive and bounded away from zero. For any set S, N_w is defined as a real-valued random measure such that:

$$N_w(S) = \int_S \frac{1}{\lambda_*(s)} dN$$

with $\frac{1}{\lambda_*(s)} = \frac{\lambda_{\min}}{\lambda(s)}$, and assuming that the positive constant $\lambda_{\min} \leq \inf\{\lambda(s); s \in S\}$ does exist. Therefore second-order statistics of N_w are defined generalizing the definition of second-order statistics for N and some theoretical properties of these statistics are proved on the basis of the martingale property of N_w (Adelfio and Schoenberg 2009).

Some of the second-order statistics useful to describe observed point patterns are here listed. The spectrum mainly used to describe point processes that exhibit long-range dependence in time is reported in Section A.1, together with theoretical results developed. In Section A.2 the definition of the reduced second-order measure of a point process and fractal dimension are given, being useful quantities in the description of attractive and repulsive features for spatial point processes.

A.1. Spectrum and weighted spectrum

Consider a temporal point process $\{N(t), 0 < t < T\}$, such that $N(t_1), \ldots,\ N(t_p)$ denote the number of events that occur in the intervals $(0, t_1], \ldots,\ (0, t_p]$.

Let $N(dt_1), \ldots,\ N(dt_p)$ denote the number of events that occur in the small intervals $(t_1, t_1 + dt_1], \ldots,\ (t_p, t_p + dt_p]$. Suppose $N(t)$ be a stationary point process, such that the joint distributions of the variables $N(dt_1), \ldots, N(dt_k)$ are unaffected by simple translations of $t_1, \ldots, t_k$.

Suppose that the process is orderly, that is the probability that a small region contains more than one point is very small.

Indicating with $\mu_N(t) = E[N(dt)]$ the mean measure of the process, we know that:

$$\mu_N(t) = \lambda(t)dt = E[N(dt)] \doteq Pr[\text{there is an event } \text{ in}(t, t + dt)]$$

It follows that $\lambda(t)$ is the intensity of the process $N(t)$. Moreover, given $s = t + h$, with $h \in \mathbb{R}_+$ define:

$$\begin{aligned} dC(h)\ dt &= \mathrm{cov}[N(ds), N(dt)] \\ &\doteq Pr[\text{there is an event of in } (s, s+ds) \\ &\quad \text{and an event of in } (t, t+dt)] - \lambda(t)\lambda(s)dtds \end{aligned}$$

When a function $c(h)$ exists, with $c(h) = dC(h)/dh$, for $h \neq 0$, then $c(h)$ defines the autocovariance density.

The spectrum of a process is obtained applying the Fourier transform to the covariance function. If $\lim_{|dt|\to 0} \frac{E[N(dt)]}{|dt|} = \lambda$, such that the complete covariance density is $c^{(c)}(h) = \lambda\delta(h)+c(h)$, with $\delta(\tau)$ the Dirac delta function and $c(h)$ continuous at the origin, the complete spectral density function for N is given by:

$$f_N(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ih\omega}c^{(c)}(h)dh = \frac{1}{2\pi}\left\{\lambda + \int_{-\infty}^{\infty} e^{-ih\omega}c(h)dh\right\}$$

Therefore, for a Poisson process with constant intensity function λ, $c(h) = 0$ and the power spectrum reduces to:

$$f_N(\omega) = \frac{1}{2\pi}\lambda$$

Intuitively, the area bounded by the spectrum in $-\pi$ and π is the variance of the process. Considering a generic frequency ω in $[-\pi, \pi]$, the quantity $\int_{-\omega}^{\omega} f(\omega)d\omega$ is the portion of the process variance explained by cycles with frequency less than ω.

A process with a high spectral density at small frequencies has the greater part of the variance concentrated at low frequency cycles. In other words, the dependence of the process at short-term is low while it is stronger at long-term.

The spectral density of the weighted process is here defined by:

$$f_{N_w}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} c_w^{(c)}(h)\exp(-i\omega h)dh \tag{6}$$

where $c_w^{(c)}(h) = \lambda_{\min}\delta(h) + c_w(h)$ is the complete covariance density of the process N_w, with $\delta(\cdot)$ the Dirac delta function and $c_w(h)$ continuous at the origin. Since $c_w(h) = 0$ (Adelfio and Schoenberg 2009) the spectral density of the weighted process (that is the spectral density calculated on the weighted points by using the inverse of the conditional intensity function multiplied by its minimum value) reduces to

$$f_{N_w}(\omega) = \frac{\lambda_{\min}}{2\pi}$$

that is the power spectrum of Poisson process with constant rate $\lambda_{\min}$.

A.2. The weighted correlation integral

Let N be a time point process defined on $[0, T] \in \mathbb{R}$, with Lebesgue measure T and let $I_{ij}(\delta)$ be the indicator variable $I(|t_i - t_j| \leq \delta)$, with t_k $\forall k = 1, \dots, n$ points of the state space.

The weighted correlation integral, for a time point process N with realizations $t_1, t_2, \ldots, t_n$ on $[0, T]$, can be written as:

$$\hat{C}_W(\delta) = \frac{1}{(\lambda_{\min} T)^2} \sum_{i}^{n} \omega_i \sum_{j \neq i}^{n} \omega_j I(|t_i - t_j| \leq \delta)$$

with $\omega_k = \frac{\lambda_{\min}}{\lambda(t_k)}, \forall k$ and $\lambda(t)$ the conditional intensity function of the process with respect to some filtration $\mathcal{H}_t$ on $[0, T]$.

As shown in Adelfio and Schoenberg (2009), the weighted correlation integral is the equivalent of the inhomogeneous or weighted K-function (Baddeley, Møller, and Waagepetersen 2000) for temporal point process models.

The asymptotic normality of the usual correlation integral for the i.i.d. case and under mixing conditions was proved in Denker and Keller (1986). In Adelfio and Schoenberg (2009) a martingale approach has been used to prove the asymptotic normality of the weighted version of the correlation integral and therefore its consistence with the one of a homogeneous Poisson process; this result has been extended to spatial point processes too.

The weighted correlation dimension, denoted with D_w, is defined as the slope of the plot of $\log \hat{C}_W(\delta)$ versus $\log \delta$ for sufficiently small δ.

Affiliation:

Giada Adelfio
Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli"
viale delle Scienze, ed. 13
90128, Palermo, Italy
E-mail: adelfio@unipa.it
URL: http://dssm.unipa.it/adelfio/giada-home-ita.htm/

Journal of Environmental Statistics http://www.jenvstat.org
Volume 1, Issue 2 Submitted: 2009-11-02
February 2010 Accepted: 2009-12-22