Space-time Point Processes semi-parametric estimation with predictive measure information

Giada Adelfio*, Marcello Chiodi

Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università degli Studi di Palermo, Italia; giada.adelfio@unipa.it, marcello.chiodi@unipa.it
**Corresponding author*

Abstract. *In this paper, we provide a method to estimate the space-time intensity of a branching-type point process by mixing nonparametric and parametric approaches. The method accounts simultaneously for the estimation of the different model components, applying a forward predictive likelihood estimation approach to semi-parametric models.*

Keywords. *Nonparametric Estimation; Forward Predictive Likelihood; etasFLP; Point Process.*

1 Processes with evolutionary features

A better comprehension and interpretation of real natural phenomena described by evolutionary process often requires the identification, and then the estimation, of a function used to specify how the present depends on the past. A simplest process is a memoryless process, that has an evolution without after-effects: the number of arrivals occurring in any bounded interval of time after time t is independent of the number of arrivals occurring before time t. This assumption is sometimes unrealistic and unuseful in real contexts, where the dependence on the past is crucial also in order to make prediction (e.g. seismic data, diseases data, etc.). Moreover, if any one realisation of a type of random process consists of a set of isolated points either in time or geographical space, point processes provide a very useful theoretical tool to represent the evolution of some random value, or system, over time. In such kind of processes it is assumed that what happens now may depend on the past, but not on the future. This identifies a natural ordering for temporal point processes. However, if the space component can not be ignored, a space-time point process is defined, by specifying a stochastic model for the space component that is based on the space "coordinates" of the considered phenomenon and the time of the next event given the known times of previous events. The past in a point process is captured by the concept of the history of the process and to specify the mean number of events in a region conditional on the past the conditional intensity function is considered [8]. Therefore the importance of getting a reliable estimation of the conditional intensity function is obvious.

In exploratory contexts or to assess the adequacy of a specific parametric model, some kind of non-parametric estimation procedure could be useful [13] [7], though this can not be considered when we look for predictive properties of the estimated intensity function.

Among the processes that model and describe the evolutionary features of the reproduction observed process we can consider the branching models. In probability theory, a branching process is a Markov process in which each individual in the $n-th$ generation produces some random number of individuals in the $(n+1)-th$ generation, according to a probability distribution that does not vary from individual to individual. These models have been recently considered for the description of different applicative fields: biology ([4]), demography ([9]; [10]), epidemiology ([3]; [2]), wildfires distribution and size ([12]; [11]).

In general, the conditional intensity function of the branching model is defined as the sum of a term describing the large-time scale variation (spontaneous activity or background) and one relative to the small-time scale variation due to the interaction with the events in the past (induced activity or offsprings):

$$\lambda_{\boldsymbol{\theta}}(t,\mathbf{s}|\mathcal{H}_t) = \mu f(\mathbf{s}) + \tau_{\boldsymbol{\phi}}(t,\mathbf{s}) \tag{1}$$

with $\boldsymbol{\theta} = (\boldsymbol{\phi},\mu)'$, the vector of parameters of the induced intensity ($\boldsymbol{\phi}$) together with the parameter of the background general intensity (μ), $f(\mathbf{s})$ the space density, and $\tau_{\boldsymbol{\phi}}(t,\mathbf{s})$ the induced intensity, given by:

$$\tau_{\boldsymbol{\phi}}(t,\mathbf{s}) = \sum_{t_j<t} \nu_{\boldsymbol{\phi}}(t-t_j,\mathbf{s}-\mathbf{s}_j).$$

In such models, we have to simultaneously estimate the different components of the intensity function (large-time scale and small-time scale). If the large-time scale component $\mu f(\mathbf{s})$ in (1) is known, the parameters $\boldsymbol{\phi}$ can be usually estimated by Maximum Likelihood method. In applications, the large-time scale component $\mu f(\mathbf{s})$ is usually estimated trough non parametric techniques, like kernel estimators.

In this paper, we provide a method to estimate the space-time intensity of the generating point process of the different components like above, by mixing nonparametric and parametric approaches.

Therefore, we propose an estimation method of the space-time intensity of a branching-type point process (described in section 2 and named FLP, Forward Likelihood Predictive), that accounts simultaneously for the estimation of parametric and nonparametric ones, applying a forward predictive likelihood estimation approach to semi-parametric models [5].

Moreover, we have collected the new computing tools in an R package (etasFLP, [6]) wich fits a branching-type model to an earthquake catalog with a mixed estimation technique: FLP for nonparametric background seisimicity and maximum likelihood for the parametric components. The two estimation steps are alternated until convergence is obtained and for each event the probability of being a background event is estimated.

2 Forward predictive likelihood

Suppose that in a space-time point process the intensity function $\lambda(\cdot)$ depends on a set of parameters $\boldsymbol{\psi}$, such that $\lambda(\mathbf{z},\boldsymbol{\psi})$.

Let denote by $\hat{\psi}(H_{t_k}) \equiv \hat{\psi}(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_i, ..., \mathbf{z}_k)$ a generic estimator of ψ, based on observations until t_k.

Assume that a realization of the process is observed in the space region $\Omega_\mathbf{s}$ and the time interval $(T_0; T_{max})$. The log-Likelihood for the point process, given the k observed values $\mathbf{z}_i$ and computed using the estimator $\hat{\psi}(\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_i, ..., \mathbf{z}_k)$ is:

$$\log L(\hat{\psi}(H_{t_k}); H_{t_k}) = \sum_{i=1}^{k} \log \lambda(\mathbf{z}_i; \hat{\psi}(H_{t_k})) - \int_{T_0}^{T_{max}} \int_{\Omega_\mathbf{s}} \lambda(\mathbf{z}; \hat{\psi}(H_{t_k}))\, d\mathbf{s}\, dt \tag{2}$$

The ML estimation can not be directly used in a semi-parametric context: in fact, for example, considering the intensity (1), which contains a component that is usually estimated nonparametrically, the likelihood (2) would be maximized putting all the mass on the observed points.

In this paper, we use the method proposed in [5] and developed in [1] that measures the ability of the observations and estimation until t_k to give information on the next observation.

Let $\hat{\psi}(H_{t_k})$ be a vector of estimators, that could include smoothing constants in a semi-parametric context, based on the observed history up to t_k. Let $\log L(\hat{\psi}(H_{t_k}); H_{t_{k+1}})$ be the likelihood computed on the first $k+1$ observations, but using the estimates based on first k, defined as:

$$\log L(\hat{\psi}(H_{t_k}); H_{t_{k+1}}) = \sum_{i=1}^{k+1} \log \lambda(\mathbf{z}_i; \hat{\psi}(H_{t_k})) - \int_{T_0}^{t_{k+1}} \int_{\Omega_\mathbf{s}} \lambda(\mathbf{z}; \hat{\psi}(H_{t_k}))\, d\mathbf{s}\, dt \tag{3}$$

For example, in equation (3), $\lambda(\mathbf{z}_{k+1}; \hat{\psi}(H_{t_k}))$ could be the intensity of the $(k+1)th$ point estimated by a kernel method using the centers given by the previous k points.

Then, we use the difference between (2) and (3) to measure the *predictive information* of the first k observations on the $k+1$-th as:

$$\begin{aligned}
&\delta_{k,k+1}(\hat{\psi}(H_{t_k}); H_{t_{k+1}}) \equiv \\
&= \log L(\hat{\psi}(H_{t_k}); H_{t_{k+1}}) - \log L(\hat{\psi}(H_{t_k}); H_{t_k}) = \\
&= \sum_{i=1}^{k+1} \log \lambda(\mathbf{z}_i; \hat{\psi}(H_{t_k})) - \int_{T_0}^{t_{k+1}} \int_{\Omega_\mathbf{s}} \lambda(\mathbf{z}; \hat{\psi}(H_{t_k})) d\mathbf{s} dt - \\
&\quad - \sum_{i=1}^{k} \log \lambda(\mathbf{z}_i; \hat{\psi}(H_{t_k})) - \int_{T_0}^{t_k} \int_{\Omega_\mathbf{s}} \lambda(\mathbf{z}; \hat{\psi}(H_{t_k})) d\mathbf{s} dt = \\
&= \log \lambda(\mathbf{z}_{k+1}; \hat{\psi}(H_{t_k})) - \int_{t_k}^{t_{k+1}} \int_{\Omega_\mathbf{s}} \lambda(\mathbf{z}; \hat{\psi}(H_{t_k})) d\mathbf{s} dt.
\end{aligned} \tag{4}$$

This leads to a technique similar to cross-validation, but applied only to the future observations: in fact, each contribution $\delta_{k,k+1}$ is based only on the past observations $t_1, ..., t_k$.

Therefore, given n the number of observations, we choose $\widetilde{\psi}(H_{t_k})$ which maximizes:

$$FLP_{k_1,k_2}(\hat{\psi}) \equiv \sum_{k=k_1}^{n-1} \delta_{k,k+1}, \tag{5}$$

where k_1 is a fixed constant, for example $k_1 = \left[\frac{n}{2}\right]$.

The quantity in (5) can be used also to compare different kinds of intensity estimates obtained by considering the optimized values of the quantities $FLP_{k_1,k_2}(\psi)$.

In this paper, we use the measure defined in (5) to estimate the nonparametric component of models like (1), providing, on the basis of the measure in (5), better kernel estimates (in terms of MISE) of space-time intensity functions than classical methods.

References

[1] Adelfio, G. and Chiodi, M. (2014) Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stochastic Environmental Research and Risk Assessment*, doi=10.1007/s00477-014-0873-8

[2] Balderama, E., Schoenberg, F.P., Murray, E., and Rundel, P.W. (2012) Application of branching point process models to the study of invasive red banana plants in Costa Rica. *JASA* **107**(498), 467–476.

[3] Becker, N. (1977) Estimation for discrete time branching processes with application to epidemics. *Biometrics*. **33**(3), 515–522.

[4] Caron-Lormier, G., Masson, J.P., Menard, N., Pierre, J.S. (2006) A branching process, its application in biology: Influence of demographic parameters on the social structure in mammal groups. *Journal of Theoretical Biology* **238**, 564–574.

[5] Chiodi, M. and Adelfio, G. (2011) Forward Likelihood-based predictive approach for space-time processes. *Environmetrics*, **22**, 749–757.

[6] Chiodi, M. and Adelfio, G. (2014). etasFLP: Estimation of an ETAS model. Mixed FLP (Forward Likelihood Predictive) and ML estimation of non-parametric and parametric components of the ETAS model for earthquake description. *R package version 1.0.2.* http://CRAN.R-project.org/package=etasFLP

[7] Choi, E. and Hall, P. (1999) Nonparametric approach to analysis of space-time data on earthquake occurrences. *Journal of Computational and Graphical Statistics*, **8**(4), 733–748.

[8] Daley, D. J. and Vere-Jones, D. (2003) *An introduction to the theory of point processes.* New York: Springer-Verlag.

[9] Jagers, P., Klebaner, F. C. (2000) Population-size-dependent and age-dependent branching processes. *Stochastic Processes and their Applications* **87**, 235–254.

[10] Johnson, R. A., Taylor, J. R. (2008) Preservation of some life length classes for age distributions associated with age-dependent branching processes. *Statistics and Probability Letters* **78**, 2981–2987.

[11] Juan, P., Mateu, J. and Saez, M. (2012) Pinpointing spatio-temporal interactions in wildfire patterns. *Stochastic Environmental Research and Risk Assessment*, **26**, 8, 1131–1150.

[12] Schoenberg, F. P., Peng, R. and Woods, J. (2003) On the distribution of wildfire sizes. *Environmetrics*, **14**, 583–592.

[13] Silverman, B. (1986) *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London.