A Marked Spatio-Temporal Point Process Model for California Earthquakes

by

Stephen L. Rathbun
Department of Statistics
326 Thomas Building
Penn State University
University Park, PA 16802
rathbun@stat.psu.edu

Acknowledgement: I thank David Brillinger who graciously provided the data for this project

Spatio-Temporal Point Processes

Alternative Approaches:

- Treat time as an additional spatial axis.
- Treat the spatio-temporal point process as a marked temporal point process whose marks are comprised of the locations of the events.

The theory of marked point processes is reviewed by:

Brémaud (1980). *Point Processes and Queues.* Springer- Verlag, New York.

Sigman, K. (1995). *Stationary Marked Point Processes.* Chapman and Hall, New York.

Data:

$$\{(\mathbf{s}_i, t_i, z_i) \, : \, i = 1, \cdots, N\}$$

where

- Location of ith earthquake: $\mathbf{s}_i \in D$
- Time of ith earthquake: $t_i \in [0, T]$
- Magnitude of ith earthquake: $z_i > z_0$
- Study region: $D \subset \Re^d$
- Time interval of study: $[0, T]$
- Cut-off magnitude below which shocks are excluded from data: z_0

Let

$$N(A \times B \times C)$$

denote the number of earthquakes occuring:

- In a region $A \subset \Re^d$,
- At times beloning to the set $B \subset \Re$,
- With magnitudes belonging to the set $C \subset \Re_+$

Marked Spatio-Temporal Point Processes

Assumption: The marginal temporal point process is *orderly*; that is,

$$\Pr\{N[t, t + \delta) > 1\} = o(\delta)$$

for all $\delta \in \Re$.

Conditional Intensity Function:

$$\Lambda(\mathbf{s},t,z) = \lim_{\substack{|d\mathbf{s}|\to 0 \\ \delta\to 0 \\ \varepsilon\to 0}} \frac{\Pr\{N\{d\mathbf{s} \times [t,t+\delta) \times [z,z+\varepsilon)\} > 0|\mathcal{H}_t\}}{\delta\varepsilon|d\mathbf{s}|}$$

where $\mathcal{H}_t$ is the σ-algebra generated by

$$\left\{N\{A \times (u,t] \times C\} \, : \, A \in \mathcal{A};\; 0 < u \le t;\; C \in \mathcal{C}\right\}.$$

where
- $\mathcal{A}$ denotes the Borel sets on $\Re^d$
- $\mathcal{C}$ denotes the Borel sets on $\Re_+$

Modeling Paradigms:

1. *Self-Exciting Point Process.* (Hawkes 1971)

   Captures the tendency for earthquakes to spawn aftershocks, clustered about their parent events.

2. *Stress-Release Model.* (Zheng and Vere-Jones 1994)

   Captures the release of stress following each seismic event, reducing the intensity of the point process until stress builds up again. This is an extension of the self-correcting point process model of Isham and Westcott (1979).

Self-Exciting Point Process Models

- Self-Exciting [Temporal] Point Process (Hawkes 1971):

$$\Lambda(t) = \rho(t) + \sum_{\{i:t_i<t\}} g(t-t_i),$$

  where $\rho(\cdot)$ and $g(\cdot)$ are nonnegative functions such that

$$\mu = \int_0^\infty g(t)dt < 1.$$

- Marked Spatio-Temporal Self-Exciting Point Process (Musmeci and Vere-Jones 1992; Ogata 1993, 1998; Rathbun 1993, 1996):

$$\Lambda(\mathbf{s},t,z) = f(z)\left\{\rho(\mathbf{s}) + \sum_{\{i:t_i<t\}} \omega(z_i)\varphi(\mathbf{s}-\mathbf{s}_i)g(t-t_i)\right\}$$

  Can be viewed as a branching process with immigration (Hawkes and Oakes 1974), where:
  - The distribution of earthquake magnitudes is given by the probability density function $f(\cdot)$;
  - "Immigrants" occur according to a Poisson process with intensity $\rho(\cdot)$ satisfying

$$\int_D \rho(\mathbf{s})d\mathbf{s} < \infty; \text{ for } |D| < \infty;$$

  - Parent shocks of magnitude z produce a Poisson number of offspring with mean

$$\omega(z)\int_0^\infty g(t)dt$$

    where $g(z)$ is typically an increasing function of shock magitude;
  - Offspring are spatially distributed about their parents according to the bivariate density function $\varphi(\cdot)$;
  - Offspring are temporally distributed about their parents according to a nonnegative function $g(\cdot)$.

Notes:

- A stationary version of the spatio-temporal self-exciting process exists provided that:

$$E_Z\left\{\omega(z)\int_0^\infty g(t)dt\right\} < 1$$

(Musmeci and Vere-Jones 1992).

- The ergodicity of the spatio-temporal self-exciting process follows from Theorem 10.3.IX of Daley and Vere-Jones (1988).

References:

Muscmeci, F., and Vere-Jones, D. (1992). *Annals of the Institute of Statistical Mathematics* **44,** 1-11.

Ogata, Y. (1993). *Bulletin of the International Statistical Institute* **55,** 249-250

Rathbun, S.L. (1993). *Bulletin of the International Statistical Institute* **55,** 379-396.

Rathbun, S.L. (1996). *Journal of Statistical Planning and Inference* **51,** 55-74.

Self-Correcting Point Process Models

- Self-Correcting [Temporal] Point Process (Isham and Westcott 1979):

$$\Lambda(t) = \exp\{\alpha + \beta(t - \gamma N[0,t))\};\ \alpha,\beta,\gamma > 0$$

- Spatio-Temporal Self-Correcting Point Process (Rathbun 1993, 1996):

$$\Lambda(\mathbf{s},t) = \exp\left\{\alpha(\mathbf{s}) + \beta\left(t - \gamma\sum_{\{i:t_i<t\}}\kappa(\mathbf{s}-\mathbf{s}_i)\right)\right\}$$

where the kernel $\kappa(\bullet)$ controls range over which stress is reduced by an earthquake.

References:

Isham, V., and Westcott, M. (1979). *Stochastic Processes and their Applications* **8,** 335-347.

Rathbun, S.L. (1993). *Bulletin of the International Statistical Institute* **55,** 379-396.

Rathbun, S.L. (1996). *Journal of Statistical Planning and Inference* **51,** 55-74.

- Stress-Release Model (Zheng and Vere-Jones 1994):

  Let

$$X(t) = X(0) + \beta t - \sum_{\{i:t_i<t\}} S(z_i)$$

denote the tectonic stress at time t, where $S(z)$ is the amount of stress released by an earthquake of magnitude z. Then the conditional intensity is:

$$\Lambda(t,z) = f(z)\psi\{X(t)\}$$

- Spatio-Temporal Stress-Release Model:

  Let

$$X(\mathbf{s},t) = X(\mathbf{s},0) + \beta(\mathbf{s})t - \sum_{\{i:t_i<t\}} S(z_i)\kappa(\mathbf{s}-\mathbf{s}_i)$$

denote the tectonic stress at location **s** at time t.

  - The function $\beta(\mathbf{s})$ allows the rate at which stress builds up to depend on location;
  - The kernel $\kappa(\bullet)$ controls range over which stress is reduced by an earthquake.

Then the condition intnsity is:

$$\Lambda(\mathbf{s},t,z) = f(z)\psi\{X(\mathbf{s},t)\}$$

Reference:

Zheng, X., and Vere-Jones, D. (1994). *Tectonophysics* **229,** 101-121.

Marked Self-Exciting-Stress-Release Process

Consider the conditional intensity function:

$$\Lambda(\mathbf{s},t,z) \\ = f(z)\left\{\psi\{X(\mathbf{s},t)\} + \sum\nolimits_{\{i:t_i<t\}} \omega(z_i)\varphi(\mathbf{s}-\mathbf{s}_i)g(t-t_i)\right\}$$

This process can be viewed as a branching process with imigration, where

- ”Immigrants” occur according to a stress-release process with tectonic stress:
$$X(\mathbf{s},t) = X(\mathbf{s},0) + \beta(\mathbf{s})t - \sum_{\{i:t_i<t\}} S(z_i)\kappa(\mathbf{s}-\mathbf{s}_i)$$
- The distribution of earthquake magnitudes is given by the probability density function $f(\cdot)$;
- Parent shocks of magnitude z produce a Poisson number of offspring with mean
$$\omega(z)\int_0^\infty g(t)dt$$
where $g(z)$ is typically an increasing function of shock magitude;
- Offspring are spatially distributed about their parents according to the bivariate density function $\varphi(\cdot)$;
- Offspring are temporally distributed about their parents according to a nonnegative function $g(\cdot)$.

Parameter Estimation:

Data:

$$\{(\mathbf{s}_i,t_i,z_i) : i = 1,\cdots,N(T)\}$$

Parametric Family of Intensity Functions:

$$\{\Lambda(\mathbf{s},t,z;\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$$

Log-Likelihood:

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{N(T)} \log\Lambda(\mathbf{s}_i,t_i,z_i;\boldsymbol{\theta}) - \int_0^T\int_D\int_{z_0}^\infty \Lambda(\mathbf{s},t,z;\boldsymbol{\theta})dzd\mathbf{s}dt$$

Statistical Inference

Temporal Point Processes: (Ogata 1978)

Spatio-Temporal Point Processes: (Rathbun 1996)

Suppose that N is a spatio-temporal point process with intensity function $\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})$ and whose marginal temporal point process is orderly, stationary, and ergodic. Then under suitable recularity conditions, the maximum likelihood estimator

$$\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}$$

and

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}\right) \xrightarrow{\mathcal{L}} N\left(\mathbf{0}, I(\boldsymbol{\theta})^{-1}\right)$$

where $I(\boldsymbol{\theta})$ is the $p \times p$ matrix whose (i,j)th element is

$$I_{ij}(\boldsymbol{\theta}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \int_D \int_{z_0}^{\infty} \Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) dz d\mathbf{s} dt$$

where

$$\Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) = \frac{\left(\frac{\partial}{\partial\theta_i}\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})\right)\left(\frac{\partial}{\partial\theta_j}\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})\right)}{\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})}$$

Assumptions:

(A$_1$) The conditional intensity $\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})$ is $\mathcal{H}_t$-predictable, continuous in $\boldsymbol{\theta}$, $\Lambda(\mathbf{s},t,z;\boldsymbol{\theta}) > 0$ almost surely for all $\boldsymbol{\theta} \in \Theta$, and almost all $\mathbf{s} \in D$, $t > 0$, and $z > z_0$. In addition,

$$\int_0^T \int_D \int_{z_0}^{\infty} \Lambda(\mathbf{s},t,z;\boldsymbol{\theta}) dz d\mathbf{s} dt < \infty;\ 0 \le T < \infty$$

with probability one.

(A$_2$) The derivatives

$$\dot{\Lambda}_i(\mathbf{s},t,z;\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_i}\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})$$

and

$$\ddot{\Lambda}_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) = \frac{\partial^2}{\partial\theta_i\partial\theta_j}\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})$$

exist and are continuous in $\boldsymbol{\theta}$ for all $\boldsymbol{\theta} \in \Theta$, for all $i,j = 1,\cdots,p$, and almost all $\mathbf{s} \in D$, $t > 0$, and $z > z_0$.

(A$_3$) For compact subsets $K \subset \Theta$, and all $i,j = 1,\cdots,p$,

$$\sup_{\boldsymbol{\theta}\in K}\sup_{t} E_{\boldsymbol{\theta}}\left\{\int_D \int_{z_0}^{\infty} \frac{(\ddot{\Lambda}_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}))^2}{\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})} dz d\mathbf{s}\right\} < \infty$$

and

$$\sup_{\boldsymbol{\theta}\in K}\sup_{t} E_{\boldsymbol{\theta}}\left\{\int_D \int_{z_0}^{\infty} \frac{(\dot{\Lambda}_i(\mathbf{s},t,z;\boldsymbol{\theta})\dot{\Lambda}_j(\mathbf{s},t,z;\boldsymbol{\theta}))^2}{\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})} dz d\mathbf{s}\right\} < \infty$$

(A$_4$) For all $i,j = 1,\cdots,p$,

$$\frac{1}{T}\int_0^T \int_D \int_{z_0}^{\infty} \Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) dz d\mathbf{s} dt \xrightarrow{p}_u I_{ij}(\boldsymbol{\theta})$$

as $T \to \infty$, where

$$\Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) = \frac{\dot{\Lambda}_i(\mathbf{s},t,z;\boldsymbol{\theta})\dot{\Lambda}_j(\mathbf{s},t,z;\boldsymbol{\theta})}{\Lambda(\mathbf{s},t,z;\boldsymbol{\theta})}$$

(A$_5$) For all $c > 0$, and for all $i,j = 1,\cdots,p$,

$$\sup_{\boldsymbol{\theta},\boldsymbol{\theta}'\in K} \left\{\frac{1}{T}\int_0^T \int_D \int_{z_0}^{\infty} |\Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}) - \Delta_{ij}(\mathbf{s},t,z;\boldsymbol{\theta}')| dz d\mathbf{s} dt; \right.$$

$$\left. \sqrt{T}|\boldsymbol{\theta} - \boldsymbol{\theta}' \le c|\right\} \xrightarrow{p}_u 0$$

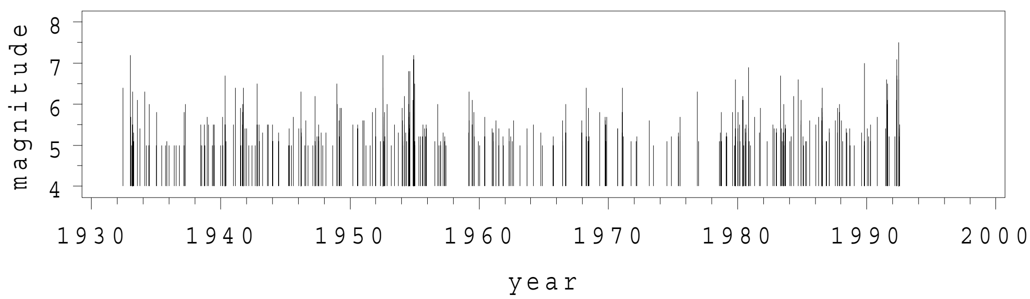
as $T \to \infty$.

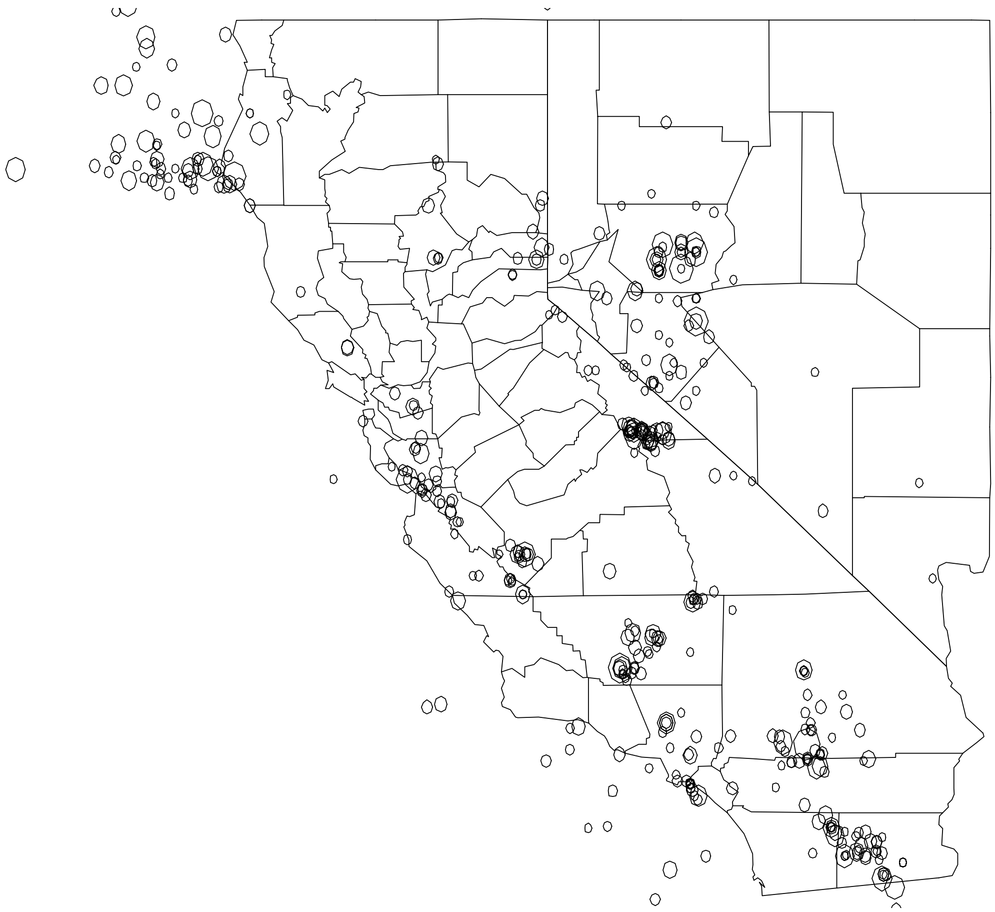
Note: Proof involves proving that the conditions of Sweeting's (1980) Theorem 1 are satisfied.

References:

Rathbun, S.L. (1996) *Journal of Statistical Planning and Inference* **51,** 55-74.
Sweeting, T.J. (1980). *Annals of Statistics* **8,** 1375-1381.

Example: Sequence of 432 earthquakes magnifude at least 5.0 occuring in and about California between 1932 and 1992. These data were kindly provided by David Brillinger.





Models:

1. Self-Exciting Point Process:

$$\Lambda(\mathbf{s},t,z) = f(z)\left\{\gamma_0\rho_0(\mathbf{s}) + \sum_{\{i:t_i<t\}} \omega(z_i)\varphi(\mathbf{s}-\mathbf{s}_i)g(t-t_i)\right\}$$

2. Self-Exciting-Stress-Release Point Process

$$\Lambda(\mathbf{s},t,z)$$
$$= f(z)\left\{\psi(\mathbf{s},t,z) + \sum\nolimits_{\{i:t_i<t\}} \omega(z_i)\varphi(\mathbf{s}-\mathbf{s}_i)g(t-t_i)\right\}$$

Selection of Model Components:

- Earthquake magnitudes are exponentially distributed (Gutenberg and Richter 1944):

$$f(z) = \alpha\exp\{-\alpha(z-z_0)\}$$

- Number of aftershocks is governed by (Ogata 1988):

$$\omega(z) = \exp\{\beta_0 + \beta_1 z_i\}$$

- Temporal distribution of aftershocks determined by the modified Omori function (Ogata 1988):

$$g(t) = (c+t)^{-(p+1)}$$

- Spatial distribution of aftershcoks determined by the Gaussian density function (Musmeci and Vere-Jones 1992):

$$\varphi(\mathbf{h}) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{h}\|^2\right\}$$

- Stress-Release component of the model is:

$$\psi\{\mathbf{s},t,z\} = \rho_0(\mathbf{s})\gamma_0\exp\left\{\gamma_1 t - \gamma_1\sum_{\{i:t_i<t\}} S(z_i)\kappa(\mathbf{s}_i - \mathbf{s})\right\}$$

where

$$S(z) = 10^{-\eta z},$$

and

$$\kappa(\mathbf{h}) = \exp\left\{-\xi\|\mathbf{h}\|^2\right\}$$

Following Zheng and Vere-Jones (1994), take $\eta = 0.75$, which corresponds to taking $S \propto E^{1/2}$ with the Gutenberg-Richter energy relation $\log_{10}E = 1.5z + \text{constant}$.

Parameter Estimation

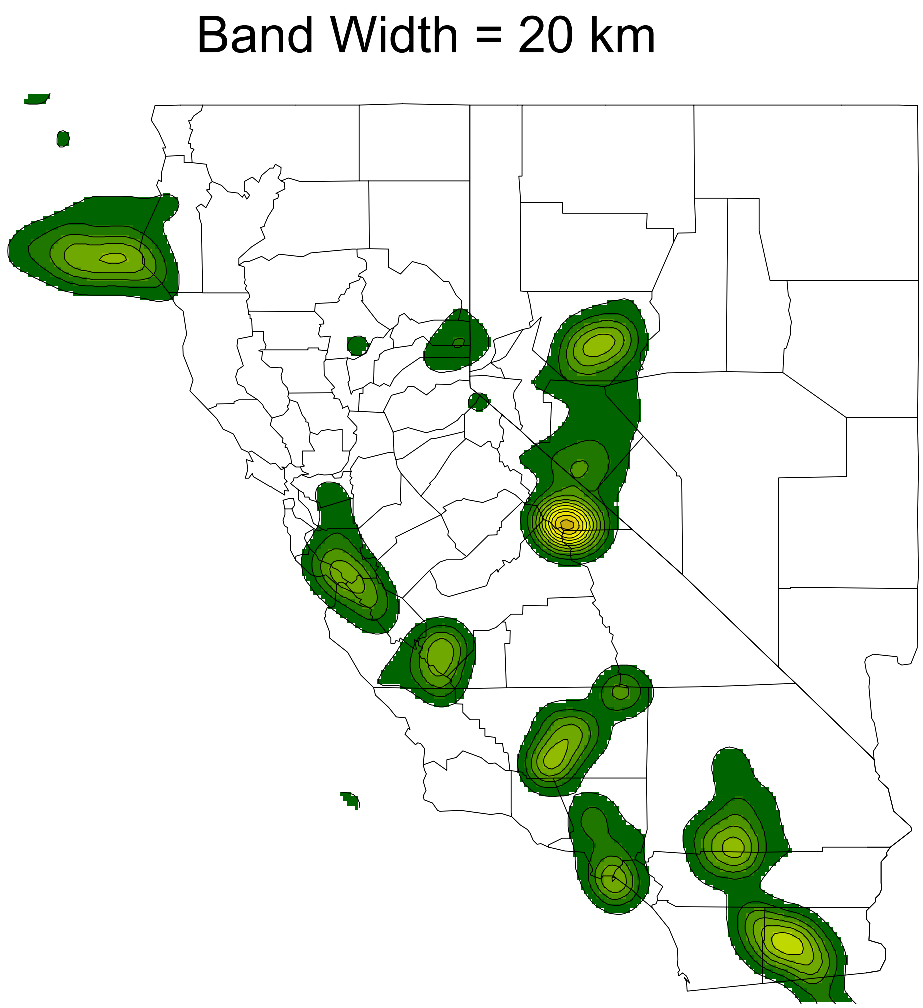
- Following the suggestion of Musmeci and Vere-Jones (1992), $\rho_0(\mathbf{s})$ is estimated using the edge-corrected kernel estimator:

$$\widehat{\rho}_0(\mathbf{s}) = \frac{1}{Tp_h(\mathbf{s})}\sum_{i=1}^{N(T)} h^{-2}\zeta\left(\frac{\mathbf{s}_i - \mathbf{s}}{h}\right)$$

where the kernel $\zeta(\bullet)$ is any two-dimensional density function, h is the bandwidth, and

$$p_h(\mathbf{s}) = \int_D \zeta\left(\frac{\mathbf{s}-\mathbf{u}}{h}\right)d\mathbf{u}$$

corrects of edge effects (Diggle 1985).



- Maximum likelihood estimates of the remaining parameters are obtained following the substitution of $\widehat{\rho}_0(\mathbf{s})$ for $\rho_0(\mathbf{s})$ in the expressions for the conditional intensity function.

Parameter Estimates: Self-Exciting Point Process

Parameter	Estimate	95% Confidence Interval
γ_0	0.574	0.503, 0.652
β_0	-10.15	-10.31,- 10.00
β_1	1.17	1.14, 1.20
c	3.95×10^{-6}	$1.85, 7.83 \times 10^{-6}$
p	1.94×10^{-7}	$0, 11.94 \times 10^{-7}$
σ	5.69	5.13, 6.33

Log Likelihood = -3628.90