

Nonparametric Self-Exciting Point Processes: Benefits and Uses

James Molyneux
Assistant Professor
Department of Statistics

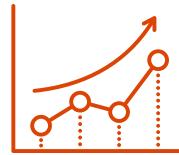
Slides: <https://github.com/jimmylovestea/wnar-2021-slides>





Outline

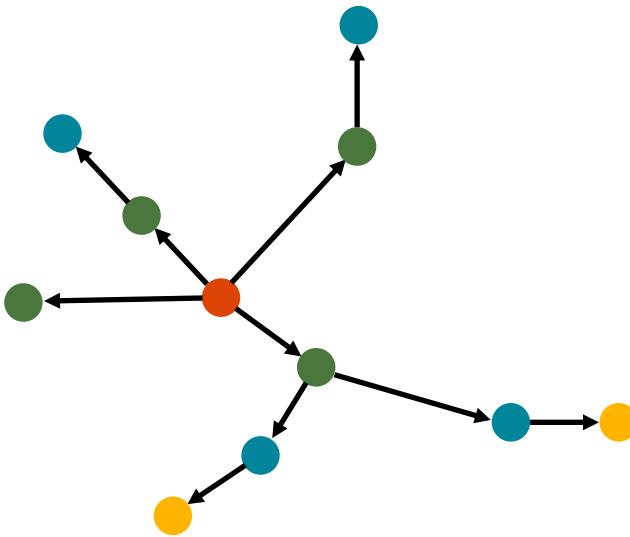
- Self-exciting phenomenon
- Self-exciting point processes
- Recent nonparametric innovations
- Application: Hector Mine earthquake sequence
- Closing thoughts



**Self-exciting
phenomenon**



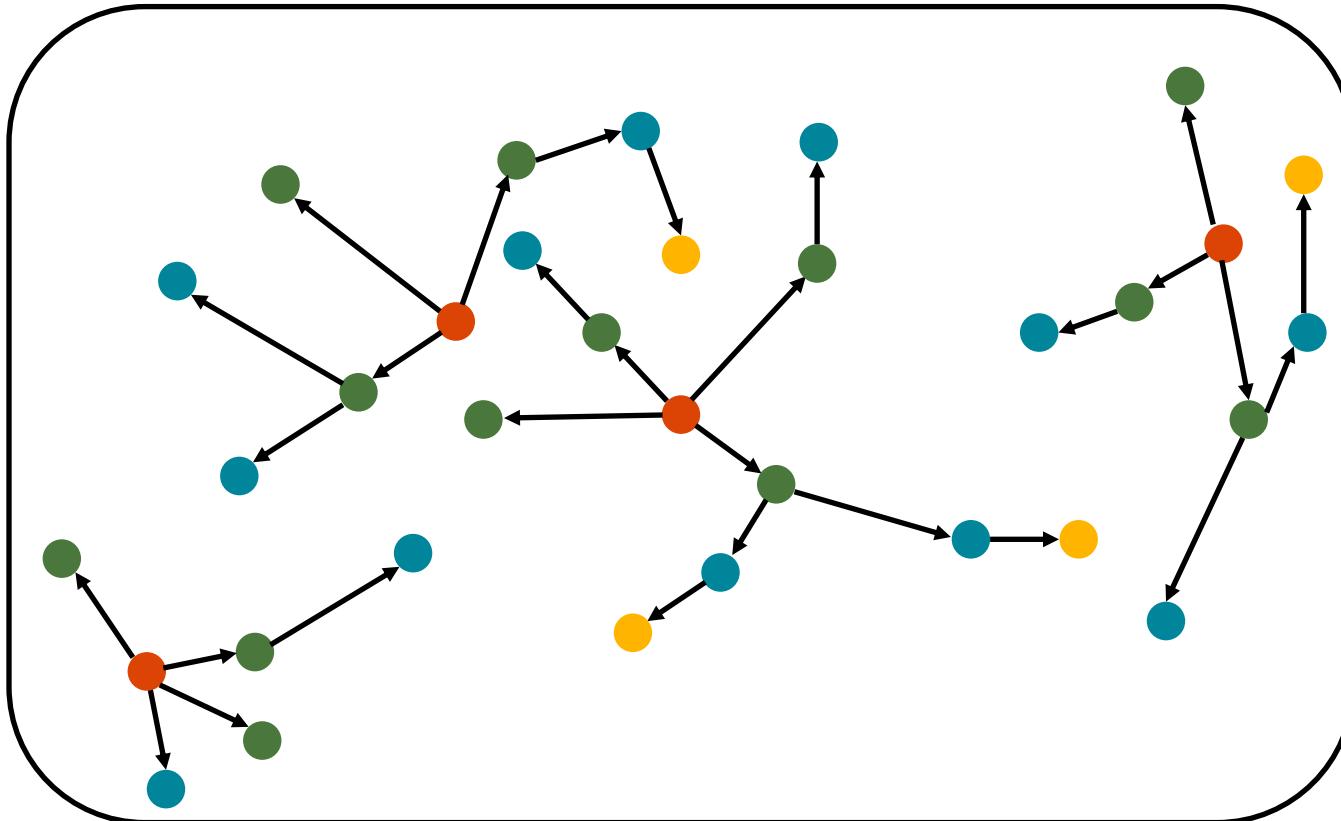
Self-exciting phenomenon



Time: $t_0 \rightarrow t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow \dots \rightarrow T$



Self-exciting phenomenon

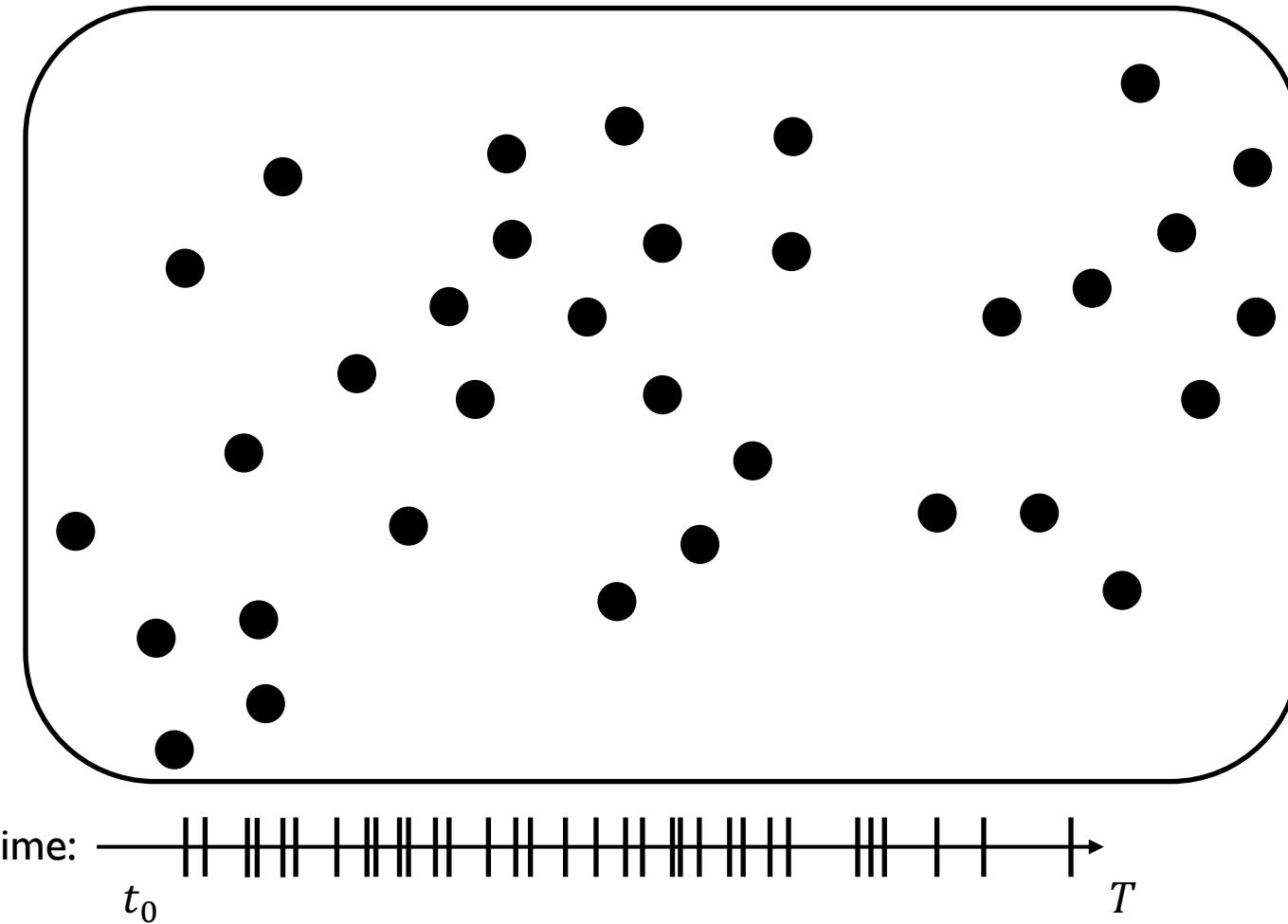


Time:





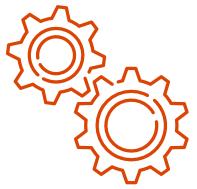
Self-exciting phenomenon





Self-exciting phenomenon

- Invasive Species ([Balderama, 2012](#))
 - Invasive species introduced by animals, floods, winds, people, etc.
 - Reproduce and spread
- Crime ([Mohler, 2011](#))
 - Successful burglary occurs in a neighborhood
 - Successive burglaries attempted nearby
- Disease Epidemics ([Park, 2020](#); [Bertozzi, 2020](#))
 - A person is infected with a disease
 - The disease spreads from person to person
- Earthquakes ([Ogata, 1999](#))
 - Mainshock earthquake occurs
 - Aftershocks occur afterwards



Self-exciting point processes



Self-exciting point processes

- Conditional Intensity : $\lambda(x, y, t | \mathcal{H}_t)$
 - Describes the infinitesimal rate of occurrences at time t and location (x, y)
- Background rate : $\mu(x, y)$
 - Function which describes the rate of *background* events at location (x, y)
- Triggering function : $v(x - x_i, y - y_i, t - t_i)$
 - Function which describes the amount of *excitation* from previous events

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} v(x - x_i, y - y_i, t - t_i)$$



Self-exciting point processes

- In applications, we often let the triggering function be *separable*
- The triggering function, $\nu(x - x_i, y - y_i, t - t_i)$, is a product of:
 -
 -
 -

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} \nu(x - x_i, y - y_i, t - t_i)$$



Self-exciting point processes

- In applications, we often let the triggering function be *separable*
- The triggering function, $\nu(x - x_i, y - y_i, t - t_i)$, is a product of:
 - Time : $g(t - t_i)$
 -
 -

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} g(t - t_i)$$



Self-exciting point processes

- In applications, we often let the triggering function be *separable*
- The triggering function, $v(x - x_i, y - y_i, t - t_i)$, is a product of:
 - Time : $g(t - t_i)$
 - Space : $h(x - x_i, y - y_i)$
 -

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} g(t - t_i) h(x - x_i, y - y_i)$$



Self-exciting point processes

- In applications, we often let the triggering function be *separable*
- The triggering function, $\nu(x - x_i, y - y_i, t - t_i)$, is a product of:
 - Time : $g(t - t_i)$
 - Space : $h(x - x_i, y - y_i)$
 - Marks : $k(m_i)$

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} g(t - t_i) h(x - x_i, y - y_i) k(m_i)$$



Self-exciting point processes

- In applications, we often let the triggering function be *separable*
- The triggering function, $\nu(x - x_i, y - y_i, t - t_i)$, is a product of:
 - Time : $g(t - t_i)$
 - Space : $h(x - x_i, y - y_i)$
 - Marks : $k(m_i)$
- Marks can be some additional information about an event
 - Earthquake magnitudes
 - Crime types
 - Plant size

$$\lambda(x, y, t | \mathcal{H}_t) = \mu(x, y) + \sum_{i: t_i < t} g(t - t_i)h(x - x_i, y - y_i)k(m_i)$$



Self-exciting point processes

- For earthquakes, the conditional intensity at time t and location (x, y) is:

- $\mu(x, y) = \tilde{\mu}(x, y)$

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

- $h(x - x_i, y - y_i) = ((x - x_i)^2 + (y - y_i)^2 + d)^{-q}$

- $k(m_i) = K e^{\alpha(m_i - m_0)}$

- Estimate parameters $\{\tilde{\mu}, c, p, d, q, K, \alpha\}$ using maximum likelihood estimation

$$\lambda(x, y, t | \mathcal{H}_t) = \tilde{\mu}(x, y) + \sum_{i: t_i < t} (t - t_i + c)^{-p} ((x - x_i)^2 + (y - y_i)^2 + d)^{-q} K e^{\alpha(m_i - m_0)}$$



**WHAT IF WE'RE UNSURE WE'VE
SPECIFIED OUR MODEL CORRECTLY?**



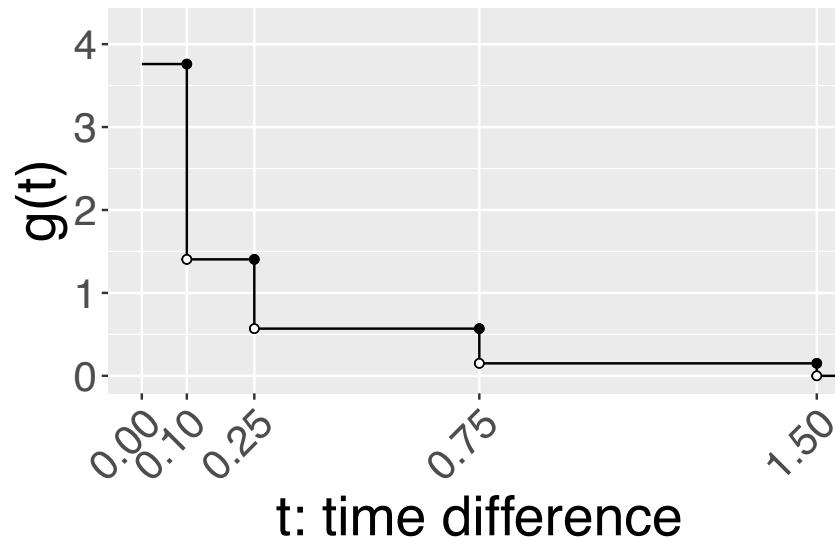
Self-exciting point processes

- Formulate the model nonparametrically ([Marsan and Lengliné, 2008](#))
 - Introduce a latent variable, Z , which describes whether an event is a *background* event or a *triggered* event
 - Compute probability event i is a background event
 - Compute probability event i is triggered by previous event j
 - Use probabilities to update $\mu(x, y)$
 - Use probabilities to update *histogram estimators* of $g(t - t_i)$, $h(x - x_i, y - y_i)$, $k(m_i)$



Self-exciting point processes

- Histogram estimators:
 - Take domain of *pairwise* differences of points in time and space
 - Split the domains into disjoint sets of bins/intervals
 - Compute step functions (i.e. constant values) for each bin/interval based on the probabilities of points which fall into each bin/interval
 - Do something similar for the marks
- Example:





Self-exciting point processes

- With estimated histogram estimators and estimated background rate, $\mu(x, y)$
 - Update probability each event is a *background* event
 - Update probabilities each event was *triggered* by a previous event
- Iterate until convergence



**Recent
nonparametric
innovations**



Recent nonparametric innovations

- Peter Boyd, Ph.D. Candidate of Statistics, Oregon State University



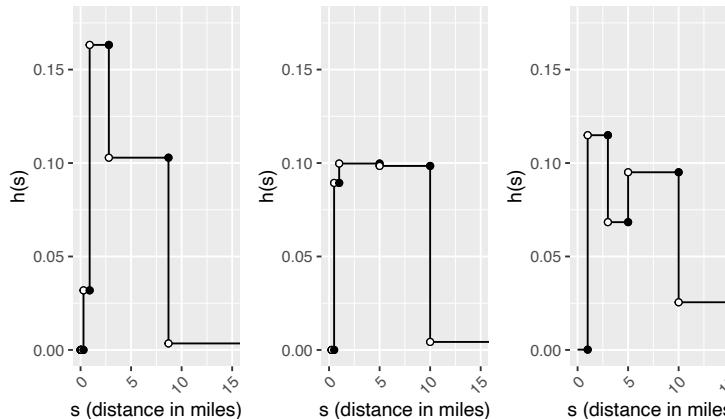
:[boydpe](#)



:[Peter Boyd](#)

- Does how we decide to break up the domains of our histogram estimators matter?

- Spoiler alert: Yes



- Is there some way to resolve this issue which would also let us estimate the triggering functions in a less “clunky” manner?
 - Spoiler alert: Also yes



Recent nonparametric innovations

- Rather than specify a set of fixed bins, what if we instead:
 - Generate lots of random bins of varying size
 - Use these random bins to estimate the expected rates of triggering which occurs in each bin
 - Smooth these estimated expected rates to give us a sense of what happens, on average
- Benefits:
 - No longer need for users to specify, and then justify, bins themselves
 - Random bins of varying sizes give us a better *overall* sense of the variability of the triggering
 - Still flexible enough to focus on where we think more/less triggering occurs

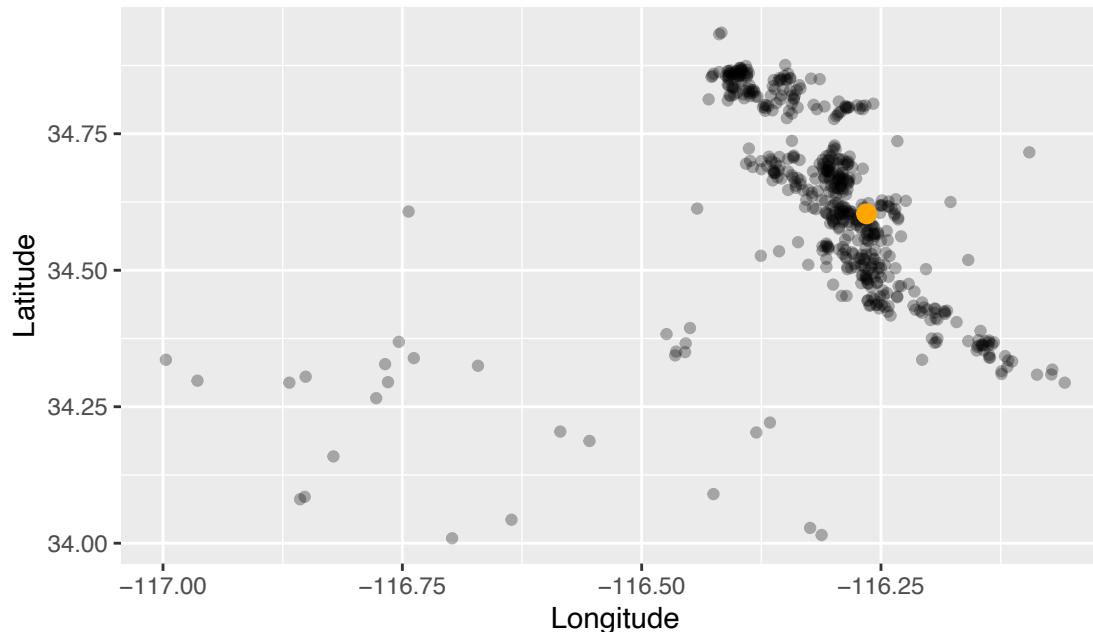


**Application: Hector
Mine earthquake
sequence**



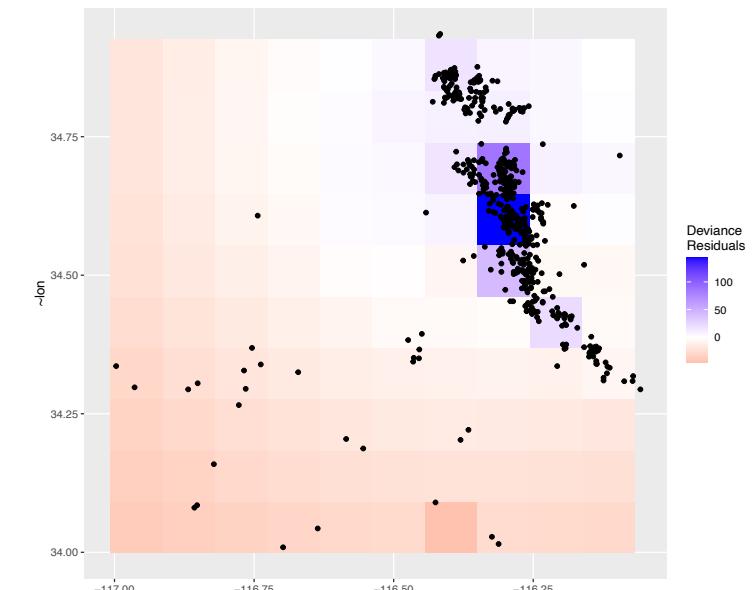
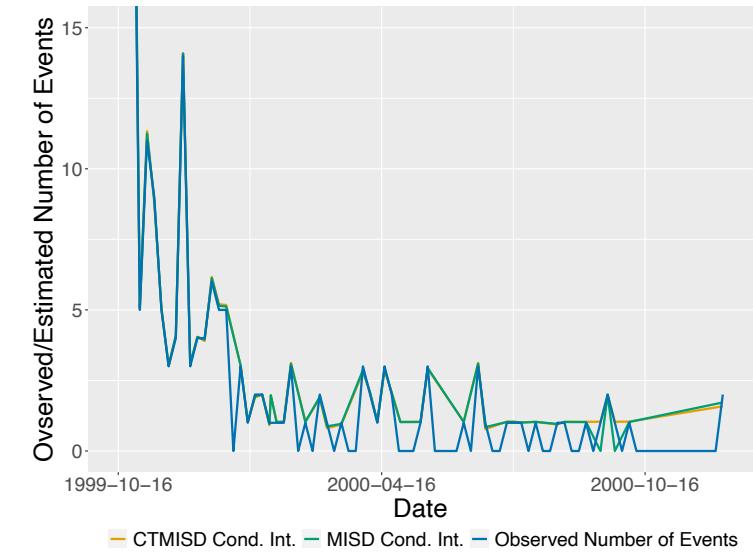
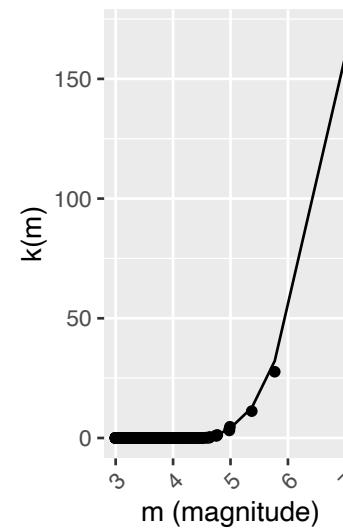
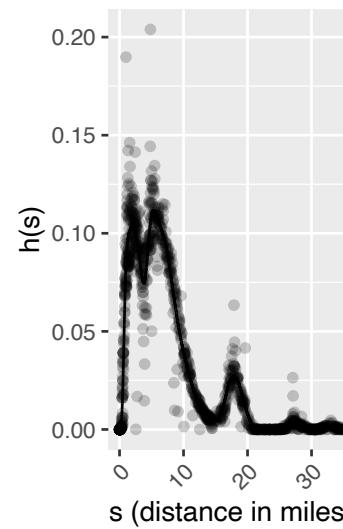
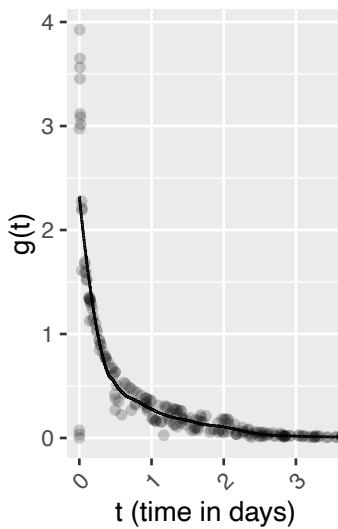
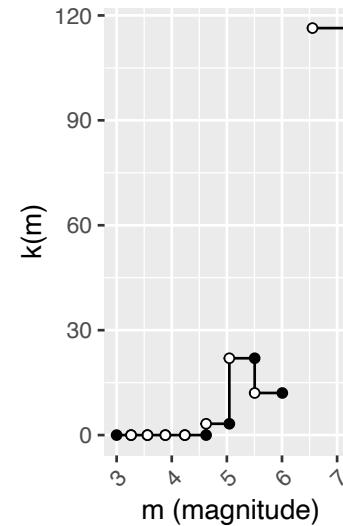
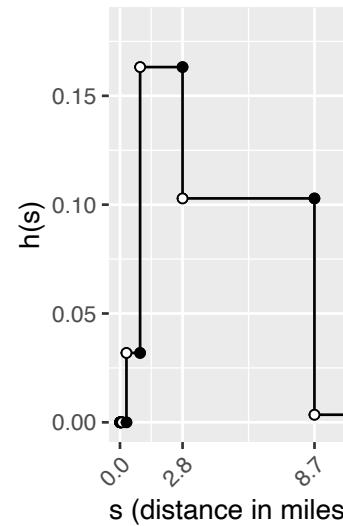
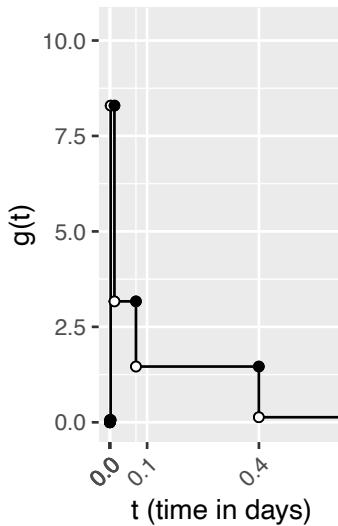
Application: Hector Mine earthquake sequence

- Large, magnitude 7.1 earthquake which occurred on October 16, 1999
- ~50 miles southeast of Barstow, CA
- Felt throughout Southern California and Las Vegas
- A “fan favorite” for folks in statistical seismology





Application: Hector Mine earthquake sequence





Closing thoughts



Closing thoughts

- Self-exciting point processes are great models for data where points cluster in time or space.
- Nonparametric methods alleviate need to specify functional forms of temporal/spatial decay.
- Recent advances make nonparametric methods less “clunky” to implement/use.
- Check out [Peter's Github](#) for:
 - An R package which implements the histogram estimator version of the model.
 - Soon: An R package which implements the new method.
- Be on the lookout for our article about the new method

THANK YOU

✉ : molyneuj@oregonstate.edu

🐱 : [jimmylovestea](#)

🐦 : [jimmylovestea](#)



Oregon State
University