Point process methodology for on-line spatio-temporal disease surveillance

Peter Diggle, Barry Rowlingson and Ting-li Su

This paper concerns the development of on-line spatio-temporal disease surveillance system for non-specific gastroenteric disease in the county of Hampshire, UK. They applied point process methodology for the surveillance purpose, in particular, using non-stationary log-Gaussian Cox process to model spatio-temporal intensity, including deterministic component for spatial and temporal variation in the normal disease pattern, and an unobserved stochastic component for localized departure from normal pattern. They give the methods for estimation and also the prediction of potential anomalies, thus alert public health departement to take follow-up action.

The model is a spatial temporal log-Gaussian Cox process model. Cox process is a Poisson process with a varying intensity which is itself a stochastic process, while log-Gaussian Cox process indicates that intensity process is a log Gaussian process. Conditional on an unobserved stochastic process R(x,t), cases form an inhomogeneous Poisson point process with intensity

$$\lambda(x,t) = \lambda_0(x)\mu_0(t)R(x,t).$$

$\lambda_0(x)$ represents spatial variation in the intensity of reported cases. Similarly, $\mu_0(t)$ is temporal variation in the spatially averaged incidence rate. Since $\lambda_0(x)$ is constrained to integrate to 1 over the region, $\mu_0(t)$ represents the time variation in the mean number of incident cases per day, and $\lambda_0(x)$ is a probability density function for the spatial distribution of incidence averaged over time. $R(x,t)$ is the latent stationary, log-Gaussian stochastic process, which is modelled as a stationary, unit-mean log-Gaussian stochastic process,

$$R(x,t) = \exp\{S(x,t)\},$$

where $S(x,t)$ is a stationary Gaussian process with mean -0.5σ^2, variance σ^2, which guarantees $E[\exp\{S(x,t)\}] = 1$ for all x and t, and correlation function

$$\rho(u,v) = \text{Corr}\{S(x,t), S(x-u,t-v)\} = r(u)\exp(-v/\theta).$$

To fit the model, they need do estimation for $\lambda_0(x), \mu_0(t)$ and parameters that determine Gaussian process $S(x,t)$.

For spatial density $\lambda_0(x)$, an adaptive bandwidth kernel estimator

$$\hat{\lambda}_0(x) = n^{-1} \sum_{i=1}^{n} h_i^{-2} \phi\{(x - x_i)/h_i\}$$

is used, where $x_i : i = 1, \cdots, n$ are the case-locations and $\phi(x)$ is a Gaussian kernel. Because there are very severe variations in population density across the county, the estimator allows a different value of the bandwidth h_i to be associated with each observed case location x_i. $h_i = h_0\{\tilde{\lambda}_0(x_i)/\tilde{g}\}^{-0.5}$. The adaptive bandwidth is estimated by

$$\tilde{\lambda}_0(x_i) = n^{-1} \sum_{i=1}^{n} h_0^{-2} \phi\{(x - x_i)/h_0\},$$

and $\tilde{g}$ is the geometric mean of $\tilde{\lambda}_0(x_i)$.

$\mu_0(t)$ represents the unconditional expectation of the numbers of cases on day t. Poisson log-linear regression model is fitted.

$$\log \mu_0(t) = \delta_{d(t)} + \alpha_1 \cos(\omega t) + \beta_1 \sin(\omega t) + \alpha_2 \cos(2\omega t) + \beta_2 \sin(2\omega t) + \gamma t,$$

where $d(t)$ is indicator for the day of the week, and $\omega = 2\pi/365$ would correspond to annual periodicity in incidence rates.

To estimate parameters of stochastic component $S(x,t)$, they used moment-based methods, assuming a separable correlation structure, $\rho(u,v) = \rho_x(u)\rho_t(v)$, and matching empirical and theoretical descriptors of the spatial temporal covariance structure.

Consider the parameters of the spatial covariance structure of $S(x,t)$, if an exponential correlation function $\rho_x(u) = \exp(-|u|/\phi)$ is used, then the theoretical pair correlation $g(u) = \exp\{\sigma^2 \exp(-|u|/\phi)\}$, and the time averaged kernel estimator

$$\hat{g}(u) = \frac{1}{2\pi u T|W|} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{i \neq j} \frac{K_h(u - ||x_i - x_j||) w(x_i, x_j)}{\hat{\lambda}_0(x_i)\hat{\mu}_0(t)\hat{\lambda}_0(x_j)\hat{\mu}_0(t)}.$$

T is length of the study-period, and W is the study area. Because events outside W are not recorded in the data, to avoid substantial negative bias in $\hat{\lambda}_0(x)$ near the boundary of the study region, w_{ij}, Ripley's edge-correcton is applied, which equals the proportion of the circumference of the circle with center x_i and radius r_{ij} which lies within W. K here is the Epanecnikov kernel function. Minimization criterion for estimating σ^2 and ϕ is

$$\int_0^{u_0} [\{\log \hat{g}(u)\} - \{\log g(u)\}]^2 du.$$

For the time component, temporal correlation is assumed an exponential form $\rho_t(v) = \exp(-|v|/\theta)$. To estimate the parameter θ, let N_t denote the numbers of incident cases on day t, the theoretical and empirical descriptors are

$$C(t, v; \theta) = \text{Cov}(N_t, N_{t-v}) = \mu_0(t)\mathbf{1}(v = 0) + \{\mu_0(t)\mu_0(t - v)\}$$
$$\times \left\{ \int_W \int_W \lambda_0(x_1)\lambda_0(x_2) \exp[\sigma^2 \exp(-v/\theta) \exp(-u/\phi)] dx_1 dx_2 - 1 \right\},$$
$$\hat{C}(t, v) = N_t N_{t-v} - \hat{\mu}_0(t)\hat{\mu}_0(t - v).$$

Minimization criterion is

$$\sum_{v=1}^{v_0} \sum_{t=v+1}^{n} \{\hat{C}(t, v) - C(t, v; \theta)\}^2.$$

The fitted model is then used for prediction. First they generate a sample from the predictive distribution of the surface $S(x, t)$, hence (R(x,t)), conditional on the observed spatio-temporal pattern of incident cases up to and including time t on a fine grid of locations $x_k, k = 1, \cdots, m$. Follows that they calculate the predictive probability $p(x_k, t; c) = \mathbf{P}\{R(x_k, t) > c|data\}$, c is a critical threshold value. A high predictive probability would suggest some follow-up means may be required.

In conclusion, they illustrate how to estimate the normal pattern of spatial temporal variations in the distribution of incident cases, and quickly identify any anomalous variations from the normal pattern. Widerly speaking, statistical evidence can be combined with other form of evidence to trigger earlier response to an emerging problem than typically achieved by current surveillance systems.