

Directional derivatives for map surveillance (A. Clark)

Goal: Monitor a set of spatio-temporal coordinates for any change in relative risk of disease over time at any location.

- Relative risk is difficult, use the conditional probability of a case.
- A common measure of change is the **directional derivative**:

$$p_i(x, t) = \lim_{\Delta t \rightarrow 0} \frac{p(x, t + \Delta t) - p(x, t)}{\Delta t}$$

- If $p(x, t)$ is increasing then $p(x, \theta) > 0$ and if $p(x, t)$ is decreasing then $p(x, \theta) < 0$.
- Checks for any difference, not just a difference in the means.
- Can be estimated by the directional derivative of the Nadaraya-Watson estimator

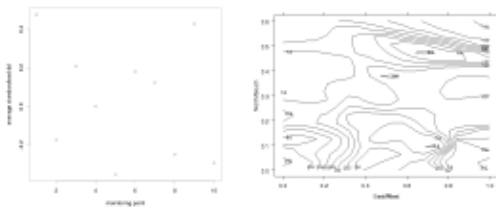
$$\frac{\partial}{\partial t} \left[\frac{\sum K(x - x_i; h_x) W(t - t_i; h_t) y_i}{\sum K(x - x_i; h_x) W(t - t_i; h_t)} \right]$$

where we only use points prior to time t and K and W are kernel functions. K is symmetric about 0 and W is peaked at zero and decays quickly to the left. The smoothing parameters h_x and h_t control the amount of dependence. If they are small then the estimated change is only based on points that are close in both space and time. If they are large then the estimates are pulled towards a common mean.

Example: Monthly dates and place of birth and birth abnormalities in a region of the United Kingdom.

We want to assess if there is any evidence that the risk is changing over time. We choose to monitor at times 0.2, 0.2999, 0.3888, 0.4777, ..., 1.000.

As a summary, we report the average standardised directional derivative at each of the ten monitoring points and we also display a contour plot of the standardised directional derivative at the first monitoring point.



- The contour plot shows that the changes conditional probability surface vary over space, with the largest change occurring in the north-east of the study region.
- The trend plot shows the relative risk has not changed considerably over the time period since no trend is present.

However, we may have isolated areas where the conditional probability has changed over time. In order to assess this, it is necessary to produce a map of the aggregated change over the time period. This is the subject of future work

Bayesian model diagnostics: surveillance residuals (C.L. Vidal Rodeiro)

Space-time models for counts of disease in a surveillance context can be used with the purpose of assessing their ability to detect changes in risk. In surveillance, it is important that the models are capable of describing the overall behaviour of the disease in space and time and also that they will be sensitive to changes in the spatio-temporal structure.

In order to monitor the process and detect the changes in risk patterns, several measures can be used, in particular **residuals** and summary functions of residuals, which will allow us to determine whether the observed data for a new year are representative of the data we might expect under a model when it is fitted for previous years.

We examine *surveillance residuals* which are computed as the difference between the observed data for a new year and the data we might expect under a model when it is fitted for previous years.

Counts of disease in fixed spatial sites $i=1, \dots, m$ and temporal periods $j=1, \dots, T$ are denoted as y_{ij} , e_{ij} and $\theta_{ij}^{(s)}$ denote, respectively, the expected number of cases and the relative risk in the i th region in time period j . The **surveillance residual** can be obtained as:

$$r_{ij}^s = y_{ij} - \frac{1}{G} \sum_{g=1}^G E(y_{ij} | \theta_{ij}^{(s)}) \quad i=1, \dots, m \quad j=2, \dots, T$$

where $E(y_{ij} | \theta_{ij}^{(s)})$ is the expected value from the posterior predictive distribution and in the context of MCMC sampling $\{\theta_{ij}^{(s)}\}$ is a set of parameter values sampled from the posterior distribution. In the case of count data they can be approximated by

$$r_{ij}^s = y_{ij} - \frac{1}{G} \sum_{g=1}^G e_{ij}^{(g)} \quad i=1, \dots, m \quad j=2, \dots, T$$

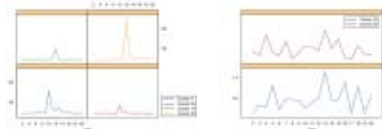
Example: We use as study region the U.S. state of South Carolina and a time period of 20 years. Three local jumps of different intensities were generated in different years (two jumps in year 10 and one jump in year 12).

The model used to fit the data assumes the following expression for the relative risk

$$\log \theta_{ij} = u_i + v_i + t_j$$

where u_i is a spatially correlated random effect, v_i is an uncorrelated spatial random effect and t_j is a temporal random effect.

In order to study the pointwise *surveillance residuals*, we randomly selected four counties which are located in the regions where the jumps in risk are located for years 10 and 12 (picture on the left). Also, two counties located in the region where the risk does not change were selected (picture on the right). It is clear that something unusual happened in counties 4 and 25 in year 10 and in counties 14 and 43 in year 12, because the residuals are much bigger than for other years and therefore the data for those particular years are not representative of what is expected under the model. There is not an unusual behaviour across the years in counties where changes in risk were not generated. The surveillance residuals for these counties highly vary over the years.



These residuals inform us that unusual events are taking place in certain counties in specific years but they do not say anything about the nature of the events. To further assess the distribution of residuals it could be useful to apply the equivalent of a parametric bootstrap in the Bayesian setting.

Conclusions

- There is considerable scope for development of new methods within the general area of surveillance of disease maps.
- There is a need to develop spatial methods which are sensitive to the sequential nature of the surveillance task. This could be via updating algorithms or through the sequential methods discussed.
- Ultimately it would be useful to develop methods which could be employed easily or routinely within a public health surveillance context.
- This development may require methods development, dissemination and the incorporation of methods into a suitable surveillance system which can be used by public health analysts.

Bayesian alarm functions and sequential posterior inference (A.B. Lawson)

- y_t is the current data (counts usually) for a monitored site (could be a small area or address).
- $y_{1:T}$ is the cumulative data on the disease up to and including time t .
- A parameter vector θ is defined.
- Syndromic variables, x_t , are also available.

Define the complete data and ancillary (syndromic) vector as

$$D_t = \begin{pmatrix} y_t \\ x_t \\ \vdots \end{pmatrix}$$

Posterior definition

Conditioning on x_t

A sequential posterior can be identified as:

$$P(\theta_{T+1}, y_T) = P(\theta_{T+1}, x_T) \propto f(y_T | \theta, x_T) P(\theta_{T+1}, x_{T+1})$$

where $P(\theta_{T+1}, x_{T+1})$ is the posterior up to and including time $T-1$. The equivalent (posterior) predictive distribution is given by:

$$P(y_T | y_{1:T-1}) = \int f(y_T | \theta, x_T) P(\theta_{T+1}, x_{T+1}) d\theta$$

Within an MCMC sampler this can be approximated via:

$$\approx \frac{1}{G} \sum_{g=1}^G f(y_T | \theta_{T+1}^{(g)}, x_T)$$

where $\theta_{T+1}^{(g)}$ is the sampled parameter vector for the g th iteration from the posterior at $T-1$. This is called **recursive Bayesian learning**.

Bayesian version of the optimal surveillance alarm function

Define a **frequentist** alarm function for the current time (s) as:

$$P(x_s) = \sum_{k=1}^s \pi_k \prod_{l=1}^k \frac{f(x(l) | \mu^k)}{f(x(l) | \mu^1)} \prod_{l=1}^s \pi_l$$

Here, the function is designed to detect any change (of μ^k to μ^l) on the range $k=1, \dots, s$. π_k is the probability of a jump at k given there hasn't been one before. Often for discrete times the geometric distribution is used for π_k .

A **Bayesian** version of this would have

$$P(x_s) = \sum_{k=1}^s h(k) \frac{\prod_{l=1}^k \int f(x(l) | \mu^k) g(\mu^k | \mu)}{\prod_{l=1}^k \int f(x(l) | \mu^1) g(\mu^1 | \mu)} \sum_{l=1}^s h(l)$$

$h(k)$ is the probability of a jump at k , and $g(\mu^k | \mu)$ is the conditional prior distribution of the new μ value given the time u .

Note that for an alm which is simply concerned with the jump at the present time (s) (and only then) the alarm function simplifies down to the **Bayes Factor**:

$$BF = \frac{f(x(s) | \mu^s) g(\mu^s | \mu)}{f(x(s) | \mu^1) g(\mu^1 | \mu)}$$

Otherwise the alarm function is a weighted product of posteriors for the $s-k+1$ time points with weights

$$w_k = h(k) / \sum_{l=1}^s h(l)$$

Syndromic Vector Monitoring

Adopting the notation above, the vector density for D_t yields $f(x(u) | \mu) \equiv f(D_u | \theta)$, and we generalize the jump to a vector form. In this case,

$$P(D_t) = \sum_{k=1}^t w_k \frac{\prod_{l=1}^k f(D_l | \theta^k) g(\theta^k | \mu)}{\prod_{l=1}^k f(D_l | \theta^1) g(\theta^1 | \mu)}$$

This alarm could be extended to include dependence on previous observed data. Thus, redefine $g(\theta | \mu)$ as $g(\theta | D_{t-1})$ and so the alarm becomes:

$$P(D_t) = \sum_{k=1}^t w_k \frac{\prod_{l=1}^k f(D_l | \theta^k) g(\theta^k | D_{l-1})}{\prod_{l=1}^k f(D_l | \theta^1) g(\theta^1 | D_{l-1})}$$

where $U-1$ denotes times up to but not including u .

This leads to a compact definition of the alarm as a weighted sum of products of posteriors for the data up to and including u :

$$P(D_t) = \sum_{k=1}^t w_k \frac{\prod_{l=1}^k P(\theta^k | D_l)}{\prod_{l=1}^k P(\theta^1 | D_l)}$$