Bayesian spatiotemporal analysis of revascularization odds using splines

Giovani L. Silva¹ and C. B. Dean²

¹ Dep. Matemática - IST, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal
² Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

Abstract: Hierarchical Bayesian models are proposed for overdispersed longitudinal and spatially correlated binomial data. This class of models accounts for correlation among regions by using random effects and allows a flexible modelling of spatiotemporal odds by using smoothing splines. The aim is to identify temporal trends and produce smoothed maps including regional effects for revascularization odds of patients hospitalized for acute coronary syndrome in Quebec. Fitting these models requires Monte Carlo Markov chain methods.

Keywords: Disease mapping; binomial data; Hierarchical Bayesian model; Spatiotemporal smoothing; MCMC methods.

1 Introduction

Describing the spatiotemporal disparities in health utilization is critical for systems analysis and for assessing the distributive impact within systems of public policies in relation to spatial or health status inequalities. In that setting, a Poisson approximation to the binomial may not be appropriate as events, referring to specific choices of medical procedures for individuals diagnosed with a disease, are typically not rare or contagious.

This work presents a Bayesian approach of a spatial generalized additive mixed model developed by MacNab and Dean (2001) and adapted here for the spatiotemporal odds analysis and binomial data (Silva et al., 2006). Smoothing splines included in these models allow a flexible modelling of spatiotemporal odds. The aim is to identify temporal trends and produce smoothed maps including regional effects, as well as to provide a study of sensitivity of such types of analyses to prior assumptions.

An analysis of regional variation for revascularization odds of patients hospitalized for acute coronary syndrome (ACS) in Quebec illustrates the methods developed. The outcome is here the annual number of revascularizations at index hospitalization for individuals diagnosed with ACS between 1993 and 2000 in that Canadian province. Revascularizations at index hospitalization have been increasing because of changes in guidelines concerning these treatments in the recent past.
2 Spatiotemporal revascularization odds

2 Spatiotemporal odds model

Let $n_{it}$ denote the number of individual hospitalised for ACS in area $i$ and year $t$ and $Y_{it}$ be the associated number of revascularizations performed, $i = 1, \ldots, n$, $t = 1, \ldots, T$. Assume $Y_{it}$ has a binomial distribution with parameters $n_{it}$ and $\theta_{it}$ (probability of revascularization).

A general spatiotemporal odds model for area $i$ is given by

$$\logit \theta_{it} = \alpha_0 + S_0(t) + S_i(t) + b_i + h_i,$$

(1)

where $S_0(t)$ is the overall trend in rates, $S_i(t)$ is the regional specific trend, $b_i$ is the spatially correlated random effect for area $i$, and $h_i$ is an independent random effect.

For model (1), cubic B-splines can be assumed for the arbitrary smoothing functions $S_0(t)$ and $S_i(t)$. Spline smoothing may reveal nonlinear temporal effects for both the overall temporal and regional temporal components. Notice that a linear trend for $S_i(t)$ offers a simple interpretation of the spatiotemporal disparities in health system utilization at the small-area level.

3 Bayesian approach

Using an intrinsic conditional autoregressive (ICAR) model (Besag, York and Mollié, 1991) for spatially structured components $b_i$ in (1), the conditional distribution of $b_i$ is

$$b_i | \mathbf{b}_{-i}, \sigma_b^2 \sim \text{Normal}(\bar{b}_i, \sigma_b^2/n_i),$$

(2)

where $\bar{b}_i = \sum_{j \in N_i} b_j/n_i$, $N_i$ denotes the set of labels of the “neighbours” of area $i$, $n_i$ is the number of areas which are adjacent to area $i$, and $\sigma_b^2$ is a variance parameter.

For unstructured spatial heterogeneity $h_i$, we assume an independent normal distribution with mean zero and variance $\sigma_h^2$. Both variance components and B-spline coefficients are assigned highly dispersed, but proper inverse gamma and normal priors, respectively. For the intercept term, $\alpha_0$, a flat prior on the whole real line is designated to ensure the model is identifiable.

Fitting the current model requires Monte Carlo Markov chain (MCMC) methods to estimate quantities of interest, e.g., the ‘relative’ odds of revascularization in area $i$, defined by $\exp(b_i + h_i)$.

4 Analysis of revascularization odds

Several spatiotemporal models (1) were fitted for analysing the odds of revascularization in 139 Local Health Areas (LHA). For instance, in increasing level of complexity: model $M_1$ ($\logit \theta_{it} = \alpha_0 + \beta t$), model $M_2$...
(M_1 \text{ plus } \delta_i t + b_i + h_i), \text{ and model } M_3 (M_1 \text{ plus } S_0(t) + \delta_i t + b_i + h_i), \text{ where } S_0(t) \text{ is modelled by B-spline basis functions. Based on some measures of overall fit, e.g., Deviance Information Criterion (DIC), model } M_3 \text{ was considered the best model. Estimates of the model parameters were obtained via MCMC methods implemented in GeoBugs (Thomas et al., 2004). Based on model } M_3, \text{ the overall trend effect } S_0(t) \text{ seems to be well described by an approximately linear trend and indicates so much increasing odds of revascularization per year for both males and females. Figure 1 presents maps of the linear LHA temporal trend estimates in revascularization odds for males (top) and females (bottom) based on model } M_3, \text{ with cutpoints as quantiles. Montreal metropolitan region is highlighted on the upper left corner. We can observe that the highest LHA temporal trends are essentially located in south-western regions both for males and for females, specifically in greater Ottawa and Montreal cities.}

5 Conclusions

The class of models (1) provides smoothing for the nonlinear temporal and regional temporal effects in mapping rates over time what can yield informative interpretation of the data. For revascularization data analysis, they were crucial namely because it was possible to isolate small-area trends, for example, for Gatineau and Hull regions, which had a sudden increase in the revascularization number after 1996, the use of splines can smooth the steady changing trend along eight years.

Acknowledgments: This paper was partially supported by FCT, Merck Frosst, GEOIDE, FRQS, and NSERC.

References


FIGURE 1. Maps of the linear LHA temporal trend estimates in revascularization odds for males (top) and females (bottom) based on model $M_3$. 