Semiparametric Spatio-Temporal Varying Coefficient Model for Matched Case-Crossover Studies

Ana Maria Ortega-Villa, Inyoung Kim {Virginia Tech, Department of Statistics}

Introduction

- In matched case-crossover studies, effects from matching covariates are removed from conditional logistic regression by conditioning on the fixed number of sets of the case and control in the stratum.
- Matching covariates such as time and spatial location often play an important role in effect modification, which makes the estimations and predictions obtained by conditional logistic regression incorrect.
- We propose a semiparametric spatio-temporal varying coefficient model, developed under the Bayesian hierarchical model framework.
- This model is able to evaluate effect modification by time and spatial location in order to make correct statistical inference.
- We demonstrate the estimation using an epidemiological example of a 1-4 bi-directional case-crossover study of childhood aseptic meningitis with drinking water turbidity, from a study conducted by Kim *et al* (2003).

Objectives

The objective of our approach is to simultaneously evaluate the following three features:

1. Detect the parametric relationship between the predictor and binary outcomes
2. Evaluate semiparametric relationships between the predictor and time
3. Determine whether there is an effect modification due to spatial location, for a small number of locations among the subjects.

Methods

Estimating Parameters Under Bayesian Hierarchal Model Framework

To estimate a time varying coefficient $\beta_r(t)$ for each available region using regression splines of order $p > 1$:

$$\beta_r(t) = \alpha_{0,r} + \alpha_{1,r}t + \cdots + \alpha_{p,r}t^p + \sum_{l=1}^{L} \alpha_{p+l,r}(t - \xi_l)_+^p$$

Prior Distributions:

$$\begin{aligned}
\boldsymbol{\alpha_{1,r}} = (\alpha_{0,r}, \ldots, \alpha_{p,r})^T &\sim N(\mathbf{0}, \phi_1^{-1} I), \\
\boldsymbol{\alpha_{2,r}} = (\alpha_{p+1,r}, \ldots, \alpha_{p+L,r})^T &\sim N(\mathbf{0}, (\phi_{\alpha,r})^{-1} I) \\
\phi_{\alpha,r} &\sim \text{Gamma}(u_\alpha, v_\alpha)
\end{aligned}$$

where ϕ_1^{-1} is a fixed large number, u_α and v_α, are the hyperparameters of the Gamma distribution.

Testing for the Difference Between Regions

To test difference between the $\beta_r(t)$s for $r = 1, \ldots, R$:

Step 1: Find the paired differences between the $\beta_r(t)$.

Step 2: For each set of paired differences find a $(1 - \alpha)100\%$ Credible Interval.

Step 3: Find the proportion of these Credible Intervals that do not include zero.

Testing for the Existence of Time Varying Coefficients

To test an effect modification due to time:

Step 1: For each region estimate $\beta_r(t)$ and $\beta_r'(t)$, the first derivative of $\beta_r(t)$ with respect to time.

Step 2: find a $(1 - \alpha)100\%$ Credible Interval for $\beta_r'(t)$

Step 3: Determine the existence of a time varying function according to the following criteria:
The $\beta_r(t)$ function is increasing in regions where the credible bounds $\beta_r'(t)$ are above zero. The $\beta_r(t)$ function is decreasing in regions where the credible bounds $\beta_r'(t)$ are below zero. If the credible band covers a portion of the zero line, we conclude that there is no significant increase of decrease of the $\beta_r'(t)$ function.

This approach is a modification of (Chaudhuri and Marron, 1999) for the proposed Bayesian Paradigm.

Results

- We present the results of the estimation of a 1-4 bidirectional matched case-crossover study of aseptic meningitis in children.
- The data comes from Kim *et al* (2003), composed of a total of 31 matched case-crossover datasets. Each dataset contains 669 patients under the age of 15 in two urban communitites in South Korea.
- We are interested in observing an effect modification in the association between turbidity of drinking water and the relative risk of aseptic meningitis in children due to time and spatial effects.
- Figure 1 and Table 1 present results of estimations of differences among the two regions.
- Figure 1 (a) shows a visible difference in $\beta(t)$ for each region, which is confirmed by the estimations of $\beta_2(t) - \beta_1(t)$, shown in part (b).
- The proportions of the calculated credible intervals that do not include zero are presented in Table 1, also showing a significant difference between $\beta_1(t)$ and $\beta_2(t)$.
- Figure 2 presents $\beta(t)$ and $\beta'(t)$ for both regions.
- For region 1, we can conclude that even though $\beta_1(t)$ seems to be decreasing, $\beta_1'(t)$'s credible bounds show that this decrease is not significant.
- For region 2, we can conclude that $\beta_2(t)$ seems to be decreasing, $\beta_2'(t)$s credible bounds show that this decrease is significant.
- Therefore the effect of water turbidity on the presence of aseptic meningitis in children does not vary over time for subjects in region 1, and varies over time for subjects in region 2.

Conclusion

- The proposed method allows us to detect relationships between a predictor X and binary outcome, as well as determine effect modifications by time and/or spatial location.
- In the 1-4 bidirectional matched case-crossover study of aseptic meningitis in children in two provinces in South Korea, we conclude that the effect modification of water turbidity on the presence or absence of disease varies according to the location of the subject.
- Additionally, we conclude that there is no effect modification by time for subjects in region one, and there is an effect modification due to time, present for subjects in region two.

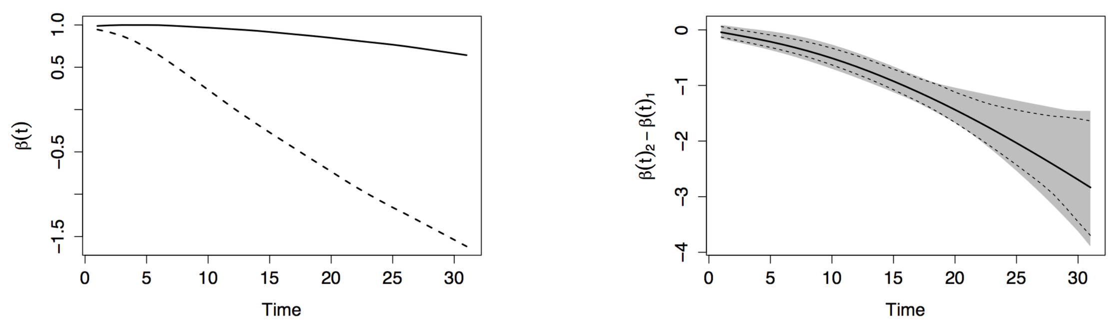
Tables and Figures

Table 1

	95%	99%
$\beta_2(t) - \beta_1(t)$	0.9355	0.8710
$\beta_2'(t) - \beta_1'(t)$	1	0.9032

Proportion of credible intervals not containing zero.

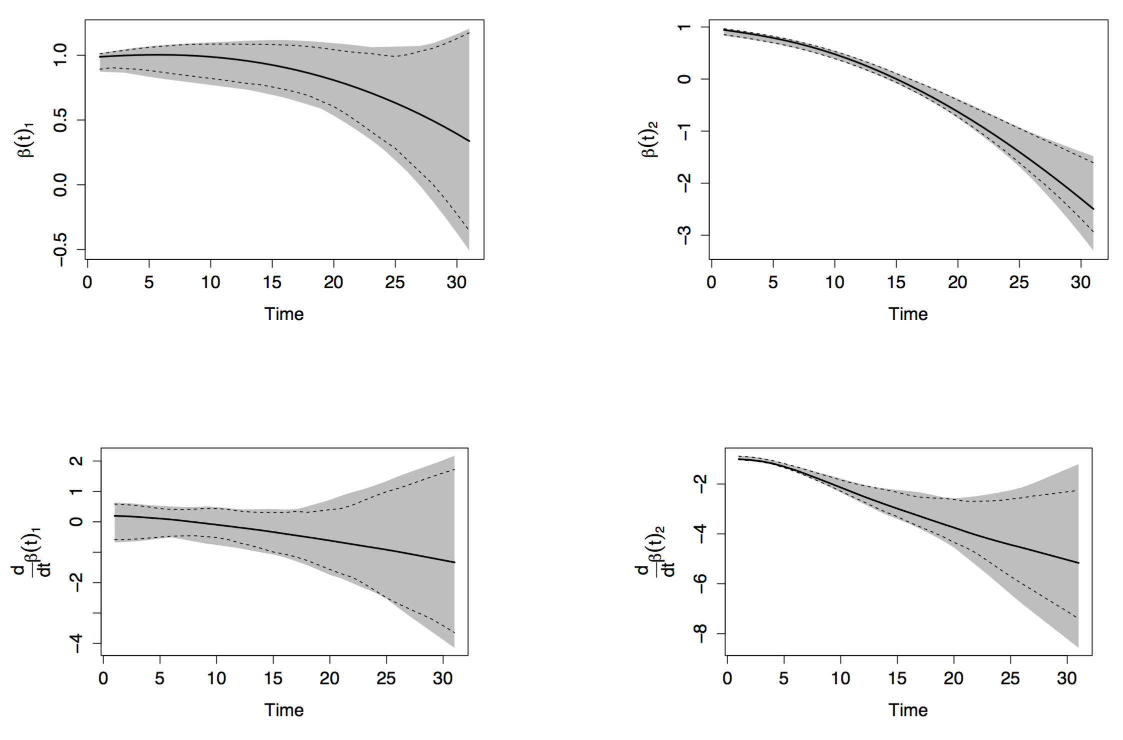
Figure 1



(a) (b)

(a) Estimated $\beta(t)$ functions. The solid line represents $\beta_1(t)$ and the dashed line represents $\beta_2(t)$. (b) Estimated $\beta_2(t) - \beta_1(t)$ and confidence bounds. (c) Estimated $\beta_2'(t) - \beta_1'(t)$and confidence bounds.

Figure 2



(a) (b)

The shaded area corresponds to 99% credible bounds, the dashed lines correspond to the 95% credible bounds and the solid line corresponded to the estimated function.

References

Chaudhuri, P. and Marron J. S. (1999) SiZer for exploration of structures in curves. *Journal of the American Statistics Association*, 94, 807-823.

Kim, H., Cheong, H. K., Park, S. K., and Bae, G. R. (2003). Drinking Water Turbidity and Aseptic Meningitis in Children in An Urban Community in Korea: Isee-567. *Epidemiology*. 14, 5, S110.