Crime in Memphis

July 20, 2024

1 Model

- ***Data: Crime Rate in Memphis*** The response variable $\mathbf{Y}$ is the crime rate for every Zip codes in Memphis and is recorded at different time points.
- ***Goal***:Developing a model for the response variable based on time varing covariates with the that allows acessing spatial and temporal associations in the response variable.
- ***Idea***: Somehow introduce Exchangibility. Idea is to group few neighbouring Zip codes and let's call that group/ collection of Zipcodes as "Area". For example, Areas can be counties of Memphis.

1.0.1 Notations

- i is the index for Areas.
- j is the index for Zipcodes.
- k is the index for Time.
- p_{ijk} is the crime rate for Zipcode j in i-th Area recorded at time k.
- $y_{ijk} = logit(p_{ijk})$
- X_{ijk} is the Covariate for Zipcode j in i-th Area recorded at time k.

1.0.2 The proposed Model

- $y_{ijk} = X_{ijk}\beta + a_{ik} + b_i + \epsilon_{ijk}$, where, $\epsilon_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma_\epsilon^2)$
- The spatial random effect $\mathbf{b} \sim CAR(A, \rho, \sigma_b)$ where $\mathbf{b} = (b_1, b_2, \cdots b_N)$, N is the number of Areas.
- The random effect a_{ik} accessing the temporal association for each Area i. Let $\mathbf{a}_i = (a_{i1}, a_{i2}, \cdots, a_{iT})$ * $\mathbf{a}_i \overset{\text{iid}}{\sim} N_T(\mathbf{0}_T, \Sigma)$. Σ is an Temporal AutoCorrelation matrix.

In order words, one can write the model as:

- $(y_{ijk} \mid a_{ik}, b_i; \beta) \overset{\text{iid}}{\sim} N(X_{ijk}\beta + a_{ik} + b_i, \sigma_\epsilon^2)$
- The spatial random effect $\mathbf{b} \sim CAR(A, \rho, \sigma_b)$ where $\mathbf{b} = (b_1, b_2, \cdots b_N)$, N is the number of Areas.
- The random effect a_{ik} accessing the temporal association for each Area i. Let $\mathbf{a}_i = (a_{i1}, a_{i2}, \cdots, a_{iT})$ * $\mathbf{a}_i \overset{\text{iid}}{\sim} N_T(\mathbf{0}_T, \Sigma)$. Σ is an Temporal AutoCorrelation matrix.

Note that, the normality assumption on the first line of the last paragraph can be replaced by other distributions (for example : Poisson etc) based on the different types of the response variable.