Fay-Herriot Model: Connection to James-Stein and Application to Census Data

Background / Chosen Paper

To produce more reliable small-area estimates, many model-based approaches have been proposed. As part of their work with the census bureau, Fay and Herriot (1979) proposed a very useful model to develop estimates of small area means. This model, originally used to allocate federal funds to local governments in the United States, has become a foundational small-area model. For my project, I plan to analyze paper and also reference a paper from Chung and Datta (n.d.) to help implement the Fay-Herriot (FH) model in a Bayesian Hierarchical Modeling framework.

Relation to Linear Models/Aspects Not Covered in Class

For $i = 1, \ldots, m$, let Y_i be the direct estimate of small area characteristic θ_i obtained from a survey. In other words, the survey contains information from m subregions and we are interested in an estimate for each region i. We also have $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ be p-dimensional vectors of covariates and the corresponding coefficients. The FH Model can be written as:

$$Y_i = \theta_i + e_i, \qquad \theta_i = \boldsymbol{x}_i^T \beta + v_i, \qquad i = 1, \ldots, m$$

where $e_i \sim_{ind} N(0, D_i)$ (sampling error) and $v_i \sim_{iid} N(0, \sigma_v^2)$ (small area effects). For this project, the sampling variances D_i are assumed to be known but the regression parameters $\boldsymbol{\beta}$ and model error variance σ_v^2 are unknown quantities.

In a special case when $D_i = D_j$ for $i \neq j$ and an improper uniform prior is assumed, the Bayes estimator for θ_i is closely related to the **James-Stein** (J-S) estimator. Fay and Herriot (1979) derived their estimator using this relationship and using an Empirical Bayes technique that estimates σ_v^2 based on the data (the resulting prior is a point mass). Their derivation along with the theoretical link between these estimators will be discussed. This theoretical discussion be informative to my classmates since the J-S estimator was briefly mentioned in 206B class but the link between the J-S and Bayes Estimator was not actually discussed.

Chosen Computational Component

Fay and Herriot (1979) predated the MCMC revolution. For this project, we use a Hierarchical Bayes (HB) approach to FH model as described by Chung and Datta.

We have

$$Y_i|\theta_1, \ldots, \theta_m, \boldsymbol{\beta}, \sigma_v^2 \sim_{iid} N(\theta_i, D_i) \qquad \theta_i|\boldsymbol{\beta}, \sigma_v^2 \sim_{iid} N(x_i^T \boldsymbol{\beta}, \sigma_v^2) \qquad \pi(\boldsymbol{\beta}, \sigma_v^2) \propto g(\boldsymbol{\beta}, \sigma_v^2)$$

I plan to use American Community Survey (ACS) data to show an application of this model to estimate the poverty rates in counties or states with a small population in a select region of the US. This will include a comparison to the direct estimates and a generalized least squares model. I will also conduct posterior inference on θ_i and σ_v^2, which is the mean poverty rate in each county. A comparison of the posterior of σ_v^2 will be compared to the estimated used by original FH method.

References

Chung, H. C., & Datta, G. S. (n.d.). Bayesian hierarchical spatial models for small area estimation. , 30.

Fay, R. E., & Herriot, R. A. (1979, June). Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association*, *74*(366a), 269–277. Retrieved 2022-05-14, from http://www.tandfonline.com/doi/abs/10.1080/01621459.1979.10482505 doi: 10.1080/01621459.1979.10482505