Journal of Statistical Planning and Inference 142 (2012) 2009–2016

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Estimation of small area event rates and of the associated standard errors

Georgios Papageorgiou *, Malay Ghosh *

University of Florida, Department of Statistics, 103 Griffin/Floyd Hall - P.O. Box 118545, Gainesville, FL 32611-8545, United States

ARTICLE INFO

Article history:
Received 9 April 2011
Received in revised form
13 February 2012
Accepted 16 February 2012
Available online 2 March 2012

Keywords:
Area level measurements
Count data
Disease mapping
Mean squared error
Optimal estimating equations

ABSTRACT

We develop empirical best estimators for small area event rates based on the hierarchical Poisson model with log-normal mixing distribution, when the basic data consists of area level measurements. We derive an approximate expression to the mean squared error of the estimators and we provide a method for estimating this expression.

1. Introduction

It has long been recognized that direct small area estimates, usually calculated based on a small number of observations, result in unreliable estimation of the corresponding parameters of interest. By now the literature on model based small area estimation methods, that thrive by 'borrowing strength' from neighboring areas, is quite vast. Ghosh and Rao (1994) provide a review of such methods based on hierarchical models, showing the usefulness of empirical Bayes and empirical best linear unbiased prediction (EBLUP) for small area estimation. Applications, for continuous responses, include estimation of per capita income (Fay and Herriot, 1979) and estimation of mean acreage under a crop (Battese et al., 1988).

Small area estimation techniques for the analysis of discrete data and in particular for the estimation of event rates have also been proposed by several authors. These, among many, include the quasi-empirical Bayes estimators of Raghunathan (1993) who applied his methodology for estimating hospital admission rates, and the empirical Bayes estimators of Clayton and Kaldor (1987) who applied their methodology for estimating relative rates of lip cancer.

Here we develop empirical best estimators for small area event rates based on the hierarchical Poisson model with log-normal mixing distribution. Specifically, we assume that the observed small area counts, y_i, are independent realizations from Poisson distributions with conditional means $E(Y_i|\lambda_i)=\lambda_i$. The proposed estimators, that are developed in the spirit of Ghosh and Maiti (2004), are shrinkage estimators in the sense that the direct estimates of λ_i, given by y_i, are shrunk towards a regression surface, $x_i^T\beta$, the mean of the mixing distribution, $E\log(\lambda_i)=x_i^T\beta$. Whereas Ghosh and Maiti (2004) consider the one parameter natural exponential family with quadratic variance functions, they considered only conjugate

* Corresponding authors.
E-mail addresses: gpapageo@stat.ufl.edu (G. Papageorgiou), ghoshm@stat.ufl.edu (M. Ghosh).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2012.02.048

priors. For instance, for the Poisson case, they considered only the gamma priors. In contrast, we consider the non-conjugate log-normal priors for Poisson parameters. In fact, our approach is more well-suited for the log-linear regression models which have become the cornerstone of research for the analysis of count data.

In Section 2 we present the model and the development of the estimators of the small area rates. These are obtained as best linear unbiased predictors (BLUPs) assuming that the model parameters are known, and then replacing them by appropriate estimators, obtained using the extended quasi-likelihood approach of Godambe and Thompson (1989). In Section 3 we derive a second order approximation to the mean squared error (MSE) of these estimators and we provide a method for estimating this expression. The paper concludes with a brief discussion. Some of the long algebraic derivations are provided in the Appendix.

Simulation study results that show that the proposed approximation to the MSE provides a reliable measure of the uncertainty associated with the estimators and results from an application to a real data set are available from the authors.

2. The model and development of estimators

Let y_i denote the observed number of events in the ith small area, $i=1,\ldots,N$. Our model assumes that conditionally on the unobserved small area rates, denoted by λ_i, the responses Y_i are independently distributed according to the Poisson distribution with mean λ_i. The responses y_i can serve as unbiased estimators of the respective rates, λ_i, but this estimation procedure is often unreliable indicating the need of developing estimators that borrow strength from related areas. We achieve this by making a distributional assumption on the λ_i. Specifically, we assume that the natural logarithm of the λ_i independently follow a normal distribution with mean $x_i^T\beta$ and variance σ^2, where x_i is the p-dimensional design vector of the ith area and β is the vector of the regression coefficients. Thus, the assumed model, in symbols, is written as

$$Y_i|\lambda_i \stackrel{\text{ind}}{\sim} \text{Poisson}(\lambda_i) \quad \text{and} \quad \log(\lambda_i) \stackrel{\text{ind}}{\sim} N(x_i^T\beta,\sigma^2),\ i=1,\ldots,N. \tag{1}$$

Henceforth, the parameters will be collectively denoted by $\eta=(\beta^T,\sigma^2)^T$.

We estimate the small area means by the respective best linear unbiased predictors (BLUPs) of λ_i given on Y_i, which we will denote by $\tilde{\lambda}_i(\eta)\equiv\tilde{\lambda}_i$, $i=1,\ldots,N$. These are given by

$$\tilde{\lambda}_i = E(\lambda_i)+\frac{\text{cov}(Y_i,\lambda_i)}{\text{var}(Y_i)}\{Y_i-E(Y_i)\}, \tag{2}$$

where, based on model (1) it can be shown that, $\mu_i\equiv E(Y_i)=\exp[x_i^T\beta+\sigma^2/2]$, $\mu_{2i}\equiv\text{var}(Y_i)=\mu_i(1+\mu_i\phi)$, and $\text{cov}(Y_i,\lambda_i)=\mu_i^2\phi$, where $\phi(\sigma^2)\equiv\phi=\exp[\sigma^2]-1$. Substituting these in (2) we obtain that

$$\tilde{\lambda}_i=\mu_i+\frac{\mu_i\phi}{\mu_i\phi+1}(Y_i-\mu_i), \tag{3}$$

which equivalently can be written as a weighted average of μ_i and Y_i: $\tilde{\lambda}_i=w_iY_i+(1-w_i)\mu_i$, where $w_i=\mu_i\phi/(\mu_i\phi+1)$.

However, the above cannot be used as such because both β and σ^2 are unknown and we thus need to estimate them from the marginal distributions of Y_i, $i=1,\ldots,N$. We achieve that by using the extended quasi-likelihood estimation theory of Godambe and Thompson (1989). This theory, as applied here, requires knowledge of the first four marginal moments of Y_i, which we can obtain based on model (1).

We begin by defining the elementary unbiased estimating functions, $g_i=(g_{1i},g_{2i})^T$, where $g_{1i}=Y_i-\mu_i=Y_i-\exp[x_i^T\beta+\sigma^2/2]$ and $g_{2i}=(Y_i-\mu_i)^2-\text{var}(Y_i)=(Y_i-\mu_i)^2-\mu_i(1+\mu_i\phi)$. Next, we calculate the matrix of the expectations of the partial derivatives of g_i with respect to the parameters

$$D_i^T=\begin{bmatrix}-E(\frac{\partial g_{1i}}{\partial\beta}) & -E(\frac{\partial g_{2i}}{\partial\beta})\\ -E(\frac{\partial g_{1i}}{\partial\sigma^2}) & -E(\frac{\partial g_{2i}}{\partial\sigma^2})\end{bmatrix}=\mu_i\begin{bmatrix}x_i & x_i(1+2\mu_i\phi)\\ \frac{1}{2} & \frac{1}{2}+\mu_i(2\phi+1)\end{bmatrix}.$$

Further, let

$$\Sigma_i=\text{var}(g_i)=\begin{bmatrix}\mu_{2i} & \mu_{3i}\\ \mu_{3i} & \mu_{4i}-\mu_{2i}^2\end{bmatrix},$$

where $\mu_{ri}=E(Y_i-\mu_i)^r$, for $r=1,\ldots,4$.

The estimates of the unknown parameters are obtained as solutions to the optimal estimating equations $S_N(\eta)\equiv\sum_{i=1}^N D_i^T\Sigma_i^{-1}g_i=0$. Note that

$$\Sigma_i^{-1}=\delta_i^{-1}\begin{bmatrix}\mu_{4i}-\mu_{2i}^2 & -\mu_{3i}\\ -\mu_{3i} & \mu_{2i}\end{bmatrix},$$

where $\delta_i=\mu_{4i}\mu_{2i}-\mu_{2i}^3-\mu_{3i}^2$. Hence, the optimal estimating equations can be written as

$$\sum_i\mu_i\delta_i^{-1}[\{\mu_{4i}-\mu_{2i}^2-(1+2\mu_i\phi)\mu_{3i}\}g_{1i}-\{\mu_{3i}-(1+2\mu_i\phi)\mu_{2i}\}g_{2i}]x_i=0, \tag{4}$$