ARTICLE IN PRESS

Journal of Statistical Planning and Inference xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

A class of general adjusted maximum likelihood methods for desirable mean squared error estimation of EBLUP under the Fay–Herriot small area model

Masayo Yoshimori Hirose

Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo, Japan

ARTICLE INFO

Article history:
Received 30 December 2016
Received in revised form 18 July 2018
Accepted 18 July 2018
Available online xxxx

Keywords:
Adjusted maximum-likelihood method
Empirical best linear unbiased prediction
Fay–Herriot model
Linear mixed model
Mean squared error

ABSTRACT

The empirical best linear unbiased prediction (EBLUP) estimator is utilized for efficient inference in various research areas, especially for small-area estimation. In order to measure its uncertainty, we generally need to estimate its mean squared prediction error (MSE). Ideally, an EBLUP-based method should not only provide a second-order unbiased estimator of MSE of EBLUP but also maintain strict positivity in estimators of both model variance parameter and MSE of EBLUP. Fortunately, the MSE estimators proposed in Yoshimori and Lahiri (2014) and Hirose and Lahiri (2017) achieve the three desired properties simultaneously. As far as we know, no other MSE estimator does so.

In this paper, we therefore seek an adequate class of general adjusted maximum-likelihood methods that simultaneously achieve the three desired properties of MSE estimation. To establish that the investigated class does so, we reveal the relationship between the general adjusted maximum-likelihood method for the model variance parameter and the general functional form of the second-order unbiased MSE estimator, maintaining strict positivity. We also compare the performance of several MSE estimators in our investigated class and others through a Monte Carlo simulation study. The results show that the MSE estimators in our investigated class perform better than those in others.

1. Introduction

In recent decades, there has been high demand for reliable statistics on smaller geographic areas and sub-populations where large samples are not available. Considering the limited number of observations, a direct design-based estimator (direct estimator) is not reliable for such "small areas"—as they are called. Even in such a situation, an explicit model-based approach can achieve more accurate estimates by borrowing strength from related areas. Accordingly, a methodology based on this approach has been developed for small-area estimation. For a comprehensive overview of small-area estimation, refer to Rao and Molina (2015).

The Fay–Herriot model (Fay and Herriot, 1979), in particular, is widely used as an aggregated level model for small-area inference as follows:

For $i = 1, \ldots, m$,

$$\text{Level 1}: y_i|\theta_i \overset{ind}{\sim} N(\theta_i, D_i);$$
$$\text{Level 2}: \theta_i \overset{ind}{\sim} N(x_i'\beta, A). \quad (1)$$

E-mail address: masayo@ism.ac.jp.

https://doi.org/10.1016/j.jspi.2018.07.006
0378-3758/

The level-1 model takes into account the sampling distribution of the direct estimator y_i for the ith small area. The true small-area mean for the ith area, denoted by θ_i, is linked to the area-specific auxiliary variables $x_i = (x_{i1}, \ldots, x_{ip})'$ in the level-2 model. In practice, the coefficient vector β in $\mathbb{R}^p$ and the model variance parameter A in this linking model are unknown. The assumption of the known D_i often follows from the asymptotic variances of the transformed direct estimates (Efron and Morris, 1975) or from empirical variance modeling (Fay and Herriot, 1979). This model can be rewritten as a specific linear mixed model:

$$y_i = \theta_i + e_i = x_i'\beta + u_i + e_i,\ \ i = 1, \ldots, m,$$

where u_i and e_i are mutually independent with the normality assumption $u_i \overset{iid}{\sim} N(0, A)$ and $e_i \overset{ind}{\sim} N(0, D_i)$. It is well known that, among all linear unbiased predictors $\hat{\theta}_i$ of θ_i, the best linear unbiased predictor (BLUP) yields the minimum mean squared prediction error (MSE), which is defined as $E[(\hat{\theta}_i - \theta_i)^2]$, where the expectation is defined with respect to the joint distribution of $y = (y_1 \ldots, y_m)'$ and $\theta = (\theta_1 \ldots, \theta_m)'$ under the Fay–Herriot model (1). We give the form of BLUP as follows:

$$\hat{\theta}_i^B = (1 - B_i)y_i + B_i x_i'\tilde{\beta},$$

where $B_i = \frac{D_i}{A+D_i}$ is called the shrinkage factor, which can shrink toward $x_i'\tilde{\beta}$ from the direct estimator y_i with $\tilde{\beta} = \tilde{\beta}(A) = (X'V^{-1}X)^{-1}X'V^{-1}y, X = (x_1, \ldots, x_m)'$ and $V = diag\{A + D_1, \ldots, A + D_m\}$.

Since A is unknown in practice, the following empirical best linear unbiased predictor (EBLUP) of θ_i is generally used for small-area inference (A is replaced with its consistent estimator, $\hat{A}$, in $\hat{\theta}_i^B$):

$$\hat{\theta}_i^{EB} = (1 - \hat{B}_i)y_i + \hat{B}_i x_i'\hat{\beta},$$

where $\hat{B}_i = \frac{D_i}{\hat{A}+D_i}$ and $\hat{\beta} = \tilde{\beta}(\hat{A})$. Hereafter, the consistent estimator $\hat{A}$ also denotes an even-translation-invariant estimator for all β and y that achieve unbiasedness in the EBLUP, as in Kackar and Harville (1981). Such conditions are satisfied by several estimators of the model variance parameter A, which are obtained by methods of moments estimators (Fay and Herriot, 1979; Prasad and Rao, 1990) and standard maximum likelihood methods such as the profile maximum likelihood method (PML) and the residual maximum likelihood method (REML). In particular, the REML estimator of A is preferred in terms of its higher-order asymptotic accuracy for large m. Let $\hat{A}_{RE}$ denote the REML estimator of A, obtained as

$$\hat{A}_{RE} = \underset{0 \le A < \infty}{\arg\max}\, L_{RE}(A|y),$$

where the residual likelihood function is

$$L_{RE}(A|y) = |X'V^{-1}X|^{-1/2}|V|^{-1/2} \exp\{-y'Py/2\}$$

and $P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$.

As mentioned above, an EBLUP is widely used as an efficient estimator based on a specific linear mixed model. It would also be quite important to measure the MSE of EBLUP as its uncertainty. For small-area inference, its MSE needs to be estimated with high accuracy since it generally depends on an unknown parameter even if the true MSE can be derived in a closed form. Given a consistent estimator of an unknown model variance parameter, the MSE of EBLUP is always larger than that of the best linear unbiased prediction (BLUP) estimator under certain conditions (Kackar and Harville, 1984). In most small-area applications, sufficient accuracy cannot be achieved by ignoring this difference, which is of the order of $O(m^{-1})$ for large m (the number of areas). Moreover, the naive MSE estimator, a consistent estimator substituted for the model variance parameter A in the MSE of BLUP, lacks second-order unbiasedness for sufficient asymptotic accuracy in small-area estimation with large m. Therefore, several second-order unbiased MSE estimators, with some bias correction, are suggested in place of the naive estimator. A pioneer work by Prasad and Rao (1990) proposed such a second-order unbiased MSE estimator based on Taylor linearization, which used a method of moment estimator of the model variance parameter A. Subsequently, MSE estimators have been developed adopting other estimators of unknown model parameter via the Taylor linearization method. Datta and Lahiri (2000) and Das et al. (2004) proposed MSE estimators of EBLUP using the standard likelihood method to estimate the model variance estimator A. Datta et al. (2005) also suggested an MSE estimator using another method for the moment estimator of A, proposed in Fay and Herriot (1979). MSE estimators of EBLUP have been constructed through not only the Taylor linearization method but also resampling methods—a bootstrap and jackknife method. The bootstrap MSE estimator was first introduced in a small-area context by Butar and Lahiri (2003) and extended by Hall and Maiti (2006). Another resampling method, the Jackknife-type MSE estimator, was developed by Jiang et al. (2002, 2016) and Chen and Lahiri (2008).

However, the estimation methods for A mentioned above could cause zero estimates although such estimates are unreasonable in the context of small-area estimation. Such estimates of A could cause additional estimation problems: an over-shrinking problem in estimating the shrinkage factor B_i and occurrence of unreasonable MSE estimates as if the MSE of EBLUP is only caused by variabilities of the estimators of β and A. Incidentally, Bell (1999) has reported zero estimates of PML and REML for four consecutive years for the Fay and Train (1997) model of 5- to 17-year-old poverty state rates in U.S. To avoid such zero estimates, Li and Lahiri (2010) proposed a specific adjusted maximum likelihood method based on the frequentist approach. Lahiri and Li (2009) put forward the concept of generalized maximum likelihood of unknown variance