propriate loss function) across the range of small areas. Such studies depend on "target values" for the parameter of interest for each small area, and generally accepted values of these target values are rarely, if ever, available (if they were, then there would be no need for indirect estimates). Thus, evaluation studies tend to produce conflicting and ambiguous results and leave all concerned less than completely satisfied. A good case in point are the many problems associated with use of a synthetic estimator to adjuste for state population undercounts in the 1990 census.

Having emphasized some of the problems associated with applications of indirect estimators, we should also mention the obvious fact that these estimation methods provide practitioners with many useful tools. Challenging research issues concerning the estimation of meaningful measures of error remain; without such measures, we must be cautious regarding inferences and actions based on these estimators. Nevertheless, in many applications, these methods provide us with an attractive alternative to the use of high variance direct estimates or, in some cases, no estimates at all.

Comment

Avinash C. Singh

The review paper of Ghosh and Rao fills a very important gap by giving a comprehensive and coherent picture of various developments in small area estimation over the last twenty years. This area is fascinating for at least three reasons: (1) there is a great demand for small area statistics by both government and private sectors for purposes of planning and policy analysis; (2) the small area problem provides a fertile ground for theoretical and applied research; and (3) the problem has attracted the attention of both Bayesians and frequentists because both approaches arise naturally and often seem to give similar results.

The main theme of my discussion is to compare and contrast the Bayesian and frequentist solutions to the problem of small area estimation. Why is it that for this problem the two approaches to statistical inference seem to converge in many practical examples including the one considered by Ghosh and Rao; that is, they provide similar results for both point estimates and the corresponding measures of uncertainty? Can we make some general statements about the similarity between the two approaches for small area estimation? How do their frequentist properties compare? Questions about the frequentist properties of some empirical Bayes methods are also raised by Ghosh and Rao in Section 5.2. Although the task of making exact comparisons is a difficult one, it is possible to make asymptotic comparisons for large m—the number of small areas. This will be the focus of my discussion.

Avinash C. Singh is Senior Methodologist, Methods Development and Analysis Section, Social Survey Methods Division, Statistics Canada, Ottawa K1A 0T6. He is also Adjunct Research Professor, Department of Mathematics and Statistics, Carleton University, Ottawa K1S 5B6.

1. MODEL REFORMULATION

As discussed in the review paper of Robinson (1991), understanding of procedures for estimating fixed and random effects helps to bridge the apparent gulf between the Bayesian and frequentist schools of thought. The present discussion will also strengthen this point. First, it will be convenient for our purposes to reformulate the model with fixed and random effects for small area estimation. Now, the general mixed linear model is given by

$$y = X\beta + Z\nu + \epsilon \tag{1}$$

where y is the n-vector of element-level data; X and Z are known matrices of orders $n \times p$ and $n \times m$, respectively, with rank $(X) = p$; β is a p-vector of fixed effects; ν is a m-vector of small area specific random effects and ϵ is a n-vector of random errors independent of ν such that $\nu \sim \mathrm{WS}\,(0, G)$, $\epsilon \sim \mathrm{WS}\,(0, R)$. The abbreviation "WS" stands for "wide sense"; that is, the distribution is specified only up to the first two moments. The covariance matrices G and R depend on some parameters λ called variance components. For the reformulation of (1), we will regard the fixed effects β as random with mean 0 and covariance matrix $\sigma_\beta^2 I$ where $\sigma_\beta^2 \to \infty$. Thus, the limiting prior distribution of β is uniform (improper) which is commonly assumed in the Bayesian approach. The reformulation is useful for computational convenience as well as for making connections