Small Area Estimation Using Times Series Models Subject to Benchmarking Constraints

Danny Pfeffermann, The Hebrew University of Jerusalem

The problem of Small Area Estimation is how to produce reliable estimates of area (domain) characteristics, when the sample sizes within the areas are too small to warrant the use of traditional direct survey estimates. This presentation will focus on the use of time series models as a vehicle for borrowing strength from past surveys. In order to protect against possible model breakdowns and to satisfy arithmetic consistency in publication, it is often required to benchmark the model dependent estimates in the small areas to the corresponding direct survey estimate in a large area for which the survey estimate is sufficiently accurate. This benchmarking process defines implicitly a way of borrowing information across the areas, which can be further enhanced via the model equations.

The presentation will show how the benchmarking can be implemented within state-space time series modelling. The computation of the benchmarked estimators and their variances requires joint modelling of the direct estimators in several areas, which in the case of many areas requires the development of new filtering and smoothing algorithms for state-space models with correlated measurement errors. The application of the proposed procedure is illustrated using U.S. Employment and Unemployment series.

Small Area Estimation Under Informative Sampling

Danny Pfeffermann, The Hebrew University of Jerusalem

The problem of small area estimation (SAE) is how to produce reliable predictors for the true means or proportions in areas with very small or no samples. This can be done by basing the inference on statistical models that permit borrowing information across the areas or over time. In this talk we consider situations where the sampling of areas is with probabilities that are related to the true (unknown) area means, and the sampling of units within the selected areas is with probabilities that are related to the values of the study variable. The problem with this kind of sampling schemes is that the model holding for the sample data differs from the model holding for the population values, giving rise to *informative sampling*. Failure to account for the effects of an informative sampling scheme may result in severe bias of the small area predictors.

We use relationships between the *population distribution*, the *sample distribution* and the *sample-complement distribution* of a study variable in order to derive approximately unbiased predictors of the area means in sampled and nonsampled areas. Appropriate bootstrap MSE estimators of correct order are also developed. The results of a Monte-Carlo study that illustrates the performance of the proposed predictors and their MSE estimators will be shown.