

University of Helsinki
Research Seminar in Statistics
Wednesday 9.9. at 16-18, Unioninkatu 35, room 114

Risto Lehtonen (HY): Developments in generalized regression and calibration estimation for small areas

Abstract: Reliable statistical information for population subgroups or regional areas is needed on important societal phenomena such as well-being, employment and income. This information is often produced using sample surveys. Small area estimation (SAE) refers to methods that can be used in situations where the area sample sizes are too small for reliable estimation. In SAE methods, the available auxiliary information of the population is incorporated in the estimation procedure by statistical models. Small area methods can be roughly divided into two groups. The traditional small area methods use area-level auxiliary information and area-level models (e.g. the Fay-Herriot model). The more recent SAE methods use unit-level models and unit-level auxiliary data are incorporated in the estimation. We focus on the family of generalized regression (GREG) estimators and calibration estimators in the estimation of parameters (totals, means and proportions such as regional poverty rates) for population subgroups (domains) and small areas. In GREG estimation we discuss the logistic GREG estimator that is assisted by a logistic mixed model. Calibration estimators considered include the traditional model-free (linear) calibration technique, which does not require an explicit model statement, and the more recent model calibration methods, where an explicit assisting model is postulated. We introduce a new calibration method called hybrid calibration, which aims at combining some of the favorable properties of model-free calibration and model calibration. The statistical properties (design bias and accuracy) of the estimators are assessed and compared by simulation methods using synthetic populations and register data maintained by Statistics Finland. For poverty rate we also compare the results with certain synthetic estimators (empirical best predictor type estimators).