

Computer-Assisted Probability Sampling and Statistical Inference in the Field

Pushpal Mukhopadhyay and Sarah Nusser

Department of Statistics and Center for Survey Statistics and Methodology
Iowa State University

pushpal@iastate.edu and nusser@iastate.edu

<http://dg.statlab.iastate.edu/dg>

Abstract

In the past, designing probability samples for biological field studies has been limited by the difficulty of locating random points in the field. The ability to receive precise GPS signals in many field settings has largely removed this constraint. We have been investigating statistical and computer-based technologies that enable field data collectors to use more sophisticated probability sampling designs. In particular, our research has focused on a scenario in which a field data gatherer can use a mobile computer to download geospatial data for the target region from the internet, and use the geospatial data layer as a basis for selecting a probability sample real-time in the field. We will discuss this research through an example with prototype software. The software allows stratified random sampling with geographic strata defined as regular rectangles on an arbitrary geographic image covering the target area, or as classes derived from a discrete-valued image such as a land cover map. Design-based estimators are included that generate valid statistical estimates in the field. We will discuss extensions of this research that include adaptive sampling designs for detecting rare events, using covariate data layers to determine selection probabilities, and generating accuracy metadata via field data collection.