

Seminar
ACTIVE SET ADAPTIVE SAMPLING¹
Steven K. Thompson²
Pennsylvania State University
Los Alamos National Laboratory

ABSTRACT

This seminar will describe adaptive sampling designs in which, at any point in the sampling, the next unit or set of units is with high probability selected from a distribution that depends on the values of variables of interest in an *active set* of units already selected. With lower probability, the next selection is made from a distribution not dependent on those values. The active set may consist of the entire current sample, or only the most recently selected unit, or a wide range of other possibilities. Design-unbiased estimation with such designs is based on a combination of initial and conditional selection probabilities, and these preliminary estimators are improved using the Rao-Blackwell method. Markov chain resampling estimators are used for larger sample sizes. Network- and spatially-based applications of the designs to a hidden human population at risk for HIV/AIDS and a wintering waterfowl survey are evaluated. The new designs can give efficiency gains over comparable conventional designs in some situations and, in comparison with other adaptive and link-tracing sampling methods, the present class of strategies has advantages in flexibility regarding adaptive criteria and breadth and depth of sample coverage, ease of implementation, control of sample sizes, and the availability of robust if computationally intense design-based estimators.

Key words: Adaptive sampling, link-tracing designs, Markov chain Monte Carlo, network sampling, Rao-Blackwell, spatial sampling.

¹ Los Alamos National Laboratory Paper LA-UR-04-0852

² Department of Statistics, 314 Thomas Building, Pennsylvania State University, University Park, PA 16801 USA. Current address: Statistical Sciences Group, D-1, MS F-600, Los Alamos National Laboratory, Los Alamos, NM 87544, skt@stat.psu.edu. Partial support for this work has been provided by the National Center for Health Statistics. The Statistical Sciences Group of Los Alamos National Laboratory provided time and resources for this work through their Visiting Faculty Program.

DATE: April 19, 2004

TIME: 4:10 pm; Refreshments will be served at 3:45 p.m. in Room 008 of the Statistics Building

PLACE: Engineering E202, CSU