STANFORD UNIVERSITY

DEPARTMENT OF STATISTICS

DEPARTMENT SEMINAR

4:15 p.m., Tuesday, June 25, 2002
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

James J. Cochran
Louisiana Tech University

Statistical Characteristics of Sample-based Coverage Optimization

The Maximal Coverage Problem (or MCP) is a problem for which the objective is to select a limited number of collectively exhaustive subsets that provide optimal population coverage (i.e., maximize the weighted cardinality of the selected subsets union). This remarkably common combinatorial optimization problem can be formulated as Integer Program (IP), and rather large instances can be solved using branch-and-bound based algorithms. Although the MCP is deterministic in nature, a random sample is often used to formulate the IP model. The resulting formulation is then used to select the subsets that provide the estimated optimal population coverage (though this solution is generally treated as though it were generated by a deterministic formulation).

We discuss various statistical characteristics of this estimate, prove that this approach yields an estimate of optimal coverage that is liberally biased, and provide some insight into how this bias behaves as various characteristics of the problem change. We then develop posterior and predictive models under both conjugate and vague priors, and provide both classical and empirical Bayes interpretations. The predictive approach is used to validate a common marketing exposure model that has previously been justified empirically. Finally, we examine an application of this work to the product design problem.

This is a fusion of joint work with Jeffery D. Camm, David J. Curry, and Martin S. Levy of the University of Cincinnati and Sriram Kannan of SABR Technologies.