EPA/600/A-96/104

Adaptive Sampling in Environmental Studies

Steven K. Thompson
Department of Statistics
326 Classroom Building
Pennsylvania State University
University Park, PA 16802-2111

Abstract

Adaptive sampling refers to designs in which the procedure for selecting units or sites at which to make observations may depend on observations made during the survey. For example, in pollution assessment studies, additional observations may be made in the vicinity of observed hot spots. In surveys of animal and plant populations, sampling intensity may be increased in the vicinity of high observed abundance. In this paper, examples of adaptive sampling procedures in survey situations involving whales, shrimp, moose, waterfowl, forest birds, fish, hardwood tree species, and environmental restoration are described.

Introduction

Adaptive sampling refers to designs in which the procedure for selecting units or sites on which to make observations may depend on observations made during the survey. Thus, the sampling plan has the flexibility to change during the course of the survey in response to observed patterns in the population. For example, in a survey of a rare, clustered animal population, whenever unusually high abundance is observed at a sample site, neighboring sites may be added to the sample. In a survey of an environmental pollutant, additional observations may be added in the vicinity of observed hot spots. Descriptions of a variety of adaptive sampling designs are found in Thompson (1992) and Thompson and Seber (1996).

Some types of adaptive designs include adaptive cluster sampling, adaptive allocation, ordinary sequential designs, and optimal Bayes designs. In adaptive cluster sampling, an initial sample is selected by some conventional design such as simple random sampling, stratified sampling, unequal probability sampling, systematic sampling, or conventional cluster sampling. Then, whenever a unit in the sample satisfies a specified condition, such as having a high value, neighboring units are added to the sample. Still more units may then be added because some of the neighboring units also satisfy the condition. With adaptive allocation designs the allocation of sampling effort among strata is determined sequentially during the survey, usually based on observed sample means or variances (Seber and Thompson 1994, Thompson 1990, 1991a,b, 1993, 1994, 1996).

Examples

An example of a situation in which adaptive allocation is used is the annual moose survey conducted by the Alaska Department of Fish and Game in interior Alaska (Gasaway et al. 1986). Moose are surveyed from aircraft flying over selected sample plots within a large study region. The study region is stratified based on the quality of habitat. At the end of each day during the survey, sample variances are computed for each stratum and the sampling effort for the next day is allocated to approximate the optimal Neyman allocation.

Adaptive allocation designs have also been used and investigated for surveys of commercial fish species including mackerel and orange roughy (Francis 1991), anchovy (Jolly 1993, Jolly and Hampton 1990, 1991), and shrimp (Thompson, Ramsey, and Seber 1992). The inherent problem in such surveys is that the mobility of the populations makes the patterns of abundance unpredictable prior to the survey so that classical allocation formulas are of limited usefulness. In the adaptive version, the pattern is assessed in the initial part of the survey, for example as the research vessel travels westward through the study region, and sampling effort reallocated for the second part of the survey, as the vessel returns westward in the example. With adaptive allocation the conventional stratified sampling estimator is not in general unbiased. Unbiased estimation is possible, however, through a variety of methods.

Adaptive cluster sampling or similar designs have been investigated for ecological populations such as waterfowl (Smith, Conroy, and Brakhage 1995), and trees (Roesch, 1993). A related procedure has sometimes been used in whale surveys (Kishino and Kasamatsu 1987, Schweder and Øien 1993). Other adaptive sampling designs have been examined for salmon (Geiger 1994) and for soil remediation (Englund and Herari 1995). The populations of interest tend to be rare and unevenly distributed spatially, with aggregation or social linkage tendencies. By adding neighboring units whenever high or interesting values are observed in the sample, the sample total of the variable of interest may be substantially increase—increasing for example the number of animals of the species that are observed—so that unbiased estimation needs to take the adaptive selection procedure into account.

The possibilities with adaptive designs are very wide. Much work is needed to determine effective designs for specific populations to compare the efficiency of adaptive designs with conventional designs and to determine which type of design will be most effective for the population of interest. For rare, clustered populations, adaptive cluster sampling can produce substantial gains in efficiency relative to conventional designs of equivalent sample size or cost. Factors influencing the relative efficiency include within-network variation, rarity within the study region, and cost issues. Adaptive allocation designs have also been shown to give gains in efficiency for unpredictably distributed populations including schooling species of fish and shrimp.

Nonsampling errors such as imperfect detection of animals, variable catchability of nets for fish, and incorrect self reporting of drug use must be taken into account with adaptive as well as with conventional designs. Generally, it is straightforward to adjust estimates for imperfect detectability by dividing observed values by detection probabilities. Imperfect detectability also adds components to the variance which must be estimated (Thompson and Seber 1994).

Conclusions

Motivation for adaptive sampling is provided both by real world situations such as those described above and by results in sampling theory showing that the optimal sampling strategy in many cases will be an adaptive one. The practical motivations for using adaptive sampling procedures for rare, clustered populations, from whales to rare plant species, have led field researchers to suggest adaptive procedures or to use them on an improvised basis. The theoretically optimal strategies tend to be too complex or require too much prior knowledge for practical implementation. Even so the theoretical results are suggestive of practical adaptive procedures with the potential to improve efficiency in surveys of real populations. Consideration of adaptive along with conventional sampling procedures greatly increases the possibilities in survey sampling.

References

Brown, J.A. (1994). The application of adaptive cluster sampling to ecological studies. *In* D.J. Fletcher and B.F.J. Manly (Eds), *Statistics in Ecology and Environmental Monitoring*, pp. 86–97. Otago Conference Series No. 2. Dunedin, New Zealand: University of Otago Press.

Englund, E.J., and Herari, N. (1995). Phased sampling for soil remediation. *Environmental and Ecological Statistics*, to appear.

Francis, R.I.C.C. (1984). An adaptive strategy for stratified random trawl surveys. *New Zealand Journal of Marine and Freshwater Research* **18**, 59–71.

Francis, R.I.C.C. (1991). Statistical properties of two-phase surveys: comment. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 1228.

Gasaway, W.C., DuBois, S.D., Reed, D.J., and Harbo, S.J. (1986). *Estimating Moose Population Parameters from Aerial Surveys.* Biological Papers of the University of Alaska (Institute of Arctic Biology) Number 22. Fairbanks: University of Alaska.

Geiger, H.J. (1994). A Bayesian approach for estimating hatchery contribution in a series of salmon fisheries. *Alaska Fishery Research Bulletin* **1**, 66–75.

Jolly, G.M. (1993). Bias in two-phase stratified random sampling for optimum allocation. Unpublished manuscript.

Jolly, G.M., and Hampton, I. (1990). A stratified random transect design for acoustic surveys of fish stocks. *Canadian Journal of Fisheries and Aquatic Science* **47**, 1282–1291.

Jolly, G.M., and Hampton, I. (1991). Reply to comment by R.I.C.C. Francis. *Canadian Journal of Fisheries and Aquatic Science* **48**, 1228–1229.

Kishino, H., and Kasamatsu, F. (1987). Comparison of the closing and passing mode procedures used in sighting surveys. *Report of the International Whaling Commission* **37**, 253–258.

Munholland, P.L., and Borkowski, J.J. (1993b). Adaptive Latin square sampling + 1 designs. Technical Report No. 3-23-93, Department of Mathematical Sciences, Montana State University, Bozeman.

Roesch, F.A. Jr. (1993). Adaptive cluster sampling for forest inventories. *Forest Science* **39**, 655–669.

Schweder, T., Øien, N., and Høst, G. (1993). Estimates of abundance of northeastern Atlantic minke whales in 1989. *Report of the International Whaling Commission* **43**, 323–331.

Seber, G.A.F., and Thompson, S.K. (1994). Environmental adaptive sampling. *In* G.P. Patil and C.R. Rao (Eds), *Handbook of Statistics, Vol. 12 (Environmental Sampling)*, pp. 201–220. New York: North Holland/Elsevier Science Publishers.

Smith, D.R., Conroy, M.J., and Brakhage, D.H. (1995). Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics*, to appear.

Thompson, S.K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association* **85**, 1050–1059.

Thompson, S.K. (1991a). Adaptive cluster sampling: Designs with
primary and secondary units. *Biometrics* **47**, 1103-1115.

Thompson, S.K. (1991b). Stratified adaptive cluster sampling. *Biometrika* **78**, 389-397.

Thompson, S.K. (1993). Multivariate aspects of adaptive cluster sampling. *In* G.P. Patil and C.R. Rao (Eds), *Multivariate Environmental Statistics*, pp.561–572. New York: North Holland/Elsevier Science Publishers.

Thompson, S.K. (1994). Factors influencing the efficiency of adaptive cluster sampling. Technical Report 94-0301, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, University Park.

Thompson, S.K. (1996). Adaptive cluster sampling based on order statistics. *Environmetrics* **7** 123-133.

Thompson, S.K., Ramsey, F.L., and Seber, G.A.F. (1992). An adaptive procedure for sampling animal populations. *Biometrics* **48**, 1195–11.

Thompson, S.K., and Seber, G.A.F. (1994). Detectability in conventional and adaptive sampling. *Biometrics* **50**, 712–724.

Thompson, S.K., and Seber, G.A.F. (1996). *Adaptive Sampling*. New York: Wiley. In press.

Computing Science and Statistics

Volume 27

Interface '95

PITTSBURGH, PENNSYLVANIA

Statistics and Manufacturing with Subthemes in Environmental Statistics, Graphics and Imaging

Proceedings of the
27th Symposium on the Interface
Pittsburgh, PA, June 21-24, 1995

Editors

Michael M. Meyer

James L. Rosenberger

INTERFACE
FOUNDATION
OF NORTH AMERICA

NHEERL-COR-2051A

TECHNICAL REPORT DATA
(Please read instructions on the reverse before com-'

1. REPORT NO. EPA/600/A-96/104	2.	:
4. TITLE AND SUBTITLE Adaptive sampling in environmental studies		5. REPORT DATE 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) S.K. Thompson		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics 326 Classroom Building Pennsylvania State University University Park, PA 16802-2111		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS US EPA ENVIRONMENTAL RESEARCH LABORATORY 200 SW 35th Street Corvallis, OR 97333		13. TYPE OF REPORT AND PERIOD COVERED Symposium paper 14. SPONSORING AGENCY CODE EPA/600/02

15. SUPPLEMENTARY NOTES
1995. Proceedings of the 27th Symposium on the Interface. Computer Science and Statistics, Statistics & Manufacturing with Subthemes in Environmental Statistics, Graphics, and Imaging, Pittsburgh, Pa, June 21-24, 1995.

16. ABSTRACT
Adaptive sampling refers to designs in which the procedure for selecting units or sites at which the procedure for selecting units or sites at which to make observations may depend on observations made during the survey. *For* example, in pollution assessment studies, additinal observatioons may be made in the viciniity of observed hot spots. In surveys of animal and plant populations, sampling intensity may be increased in the vicinitiy of observed abundance. In this paper, examples of adaptive sampling procedures in survey situations involving whales, shrimp, moose, waterfowl, forest birds, fish, hardwood tree species, and environmental restoration are described.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Adaptive sampling, environmental statistics.		

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES 4
	20. SECURITY CLASS *(This page)*	22. PRICE