

Two-Stage Sequential Sampling: A Neighborhood-Free Adaptive Sampling Procedure

Mohammad SALEHI M. and David R. SMITH

Designing an efficient sampling scheme for a rare and clustered population is a challenging area of research. Adaptive cluster sampling, which has been shown to be viable for such a population, is based on sampling a neighborhood of units around a unit that meets a specified condition. However, the edge units produced by sampling neighborhoods have proven to limit the efficiency and applicability of adaptive cluster sampling. We propose a sampling design that is adaptive in the sense that the final sample depends on observed values, but it avoids the use of neighborhoods and the sampling of edge units. Unbiased estimators of population total and its variance are derived using Murthy's estimator. The modified two-stage sampling design is easy to implement and can be applied to a wider range of populations than adaptive cluster sampling. We evaluate the proposed sampling design by simulating sampling of two real biological populations and an artificial population for which the variable of interest took the value either 0 or 1 (e.g., indicating presence and absence of a rare event). We show that the proposed sampling design is more efficient than conventional sampling in nearly all cases. The approach used to derive estimators (Murthy's estimator) opens the door for unbiased estimators to be found for similar sequential sampling designs.

Key Words: Clustered population; Freshwater mussels; Murthy's estimator; Neyman allocation; Rare events; Waterfowl.

1. INTRODUCTION

Designing an efficient sampling scheme for a rare and clustered population is a challenging area of research. Recently, adaptive sampling designs have been identified as leading candidates for this application (Thompson and Seber 1996; Christman 2000; Smith, Brown, and Lo 2004). A sampling design is "adaptive" if the procedure for selecting the sample depends on the values of the variable of interest observed in the sample. One of the simplest and perhaps oldest of the adaptive sampling designs is inverse sampling. Haldane (1945)

Mohammad Salehi M. is Associate Professor, School of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran (E-mail: salehi_m@cc.iut.ac.ir). David R. Smith is Biostatistician, U.S. Geological Survey, Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430 (E-mail: david_r.smith@usgs.gov).

©2005 American Statistical Association and the International Biometric Society
Journal of Agricultural, Biological, and Environmental Statistics, Volume 10, Number 1, Pages 84–103
DOI: 10.1198/108571105X28183

used inverse sampling to estimate the population proportion of a rare event in a finite population context. In this design, sampling continues until a predetermined number of rare events are found.

Adaptive cluster sampling, introduced by Thompson (1990), is a recent example of an adaptive sampling design. This sampling design has drawn the attention of practitioners and researchers during the last decade because it has been shown to be useful for estimating parameters of highly clustered and rare populations (Smith, Brown, and Lo 2004). In general, if a rare event is found when applying adaptive cluster sampling, then sampling continues in the vicinity of that location with the hope of observing more rare events. The essential ingredients of the adaptive cluster sampling method consist of an initial selection of units, a condition C that determines when one adaptively samples neighboring units, and a description of the neighborhood for each unit.

Defining the neighborhood, which is a crucial task, can be problematic for several reasons. First, an inappropriate definition of the neighborhood will lead to selecting excessive numbers of “edge units.” Edge units do not meet the condition C , are in the neighborhood of unit(s) that do meet the condition, and contribute little or nothing to the precision of estimators (Salehi 1999). Sampling freshwater mussels, which we discuss in the next section, provides a practical example of edge units’ negative effects (Smith, Vilella, and Lemarie 2003).

Second, navigating among all units in a neighborhood can be logistically difficult. For example, in fisheries surveys, conducted over large bodies of water, it is not practical to follow the adaptive cluster sampling protocol by sailing to all neighboring units until the condition is no longer met. Lo, Griffith, and Hunter (1997) and Hanselman et al. (2003) applied adaptive cluster sampling to Pacific hake larvae (*Merluccius productus*) and Gulf of Alaska rockfish (*Sebastes* sp.), respectively. However, in both studies the investigators restricted adaptive sampling by imposing an arbitrary stopping rule, which resulted in biased estimates. Aerial surveys of wildlife can suffer from similar navigational limitations.

Third, defining a neighborhood can be impossible in some cases. Consider the problem of estimating the number of teenagers who participate in a rare activity (e.g., using illegal drugs or cleaning their bedrooms). Teenagers’ tendency to follow the behavior of other teenagers living in the same area could be used in adaptive cluster sampling by defining the neighborhood as all of a teenager’s friends who live in the same area. However, adaptive cluster sampling requires neighborhoods to be “symmetric” and the teenagers’ friendships would have to be mutual. That is, if teenager A considers teenager B to be a friend, would B consider A to be a friend? If the answer could be “no,” then adaptive cluster sampling would not work in this case.

This article introduces an adaptive sampling design that does not require a neighborhood and does not generate “edge units” in the sample, but does exploit clustering in the population to find rare events. Importantly, through the application of Murthy’s estimator (Murthy 1957) the design is unbiased. Our approach follows the two-stage sampling approach in which one selects primary sample units (PSUs) by some conventional sampling design and then selects a subsample of secondary sample units within each of the selected primary units.

Throughout this article, the term “unit” will mean secondary unit. Whenever we come across a unit satisfying a condition, say C , a predetermined number of extra units are selected at random within its primary unit with the hope of finding more units satisfying condition C . We call this method the two-stage sequential sampling design.

When sampling rare populations, Kalton and Anderson (1986) recommended stratified sampling (a special case of two-stage sampling) with unequal allocation of the sample size in which (1) strata with high concentrations of the rare events are oversampled and (2) a small proportion of strata contain a substantial proportion of the rare events. Thus, based on their recommendations and assuming the rare events are clustered, we endeavor to confine the rare clusters in some PSUs such that those PSUs contain a substantial proportion of the rare events.

The proposed sampling design differs from other two-stage or stratified designs that incorporate adaptive or sequential sampling (Francis 1984; Jolly and Hampton 1990; Salehi and Seber 1997; Christman 2003). An important difference is that the other designs, with the exception of Salehi and Seber (1997), are presented with biased estimators. In addition, we suggest that our design offers advantages in ease and flexibility of implementation. Francis (1984) introduced an adaptive strategy for stratified sampling, which was developed for a fishery trawl survey. In Francis’s design, strata are sampled in the first phase and then extra effort is allocated in the second phase proportional to a variance reduction criterion. Sampling in the second phase, which should have a low sample size, must wait until completion of the first phase, which must have a large enough sample to yield good estimates of strata-specific variance. Jolly and Hampton (1990) introduced and used a two-phase procedure similar to Francis (1984). In our proposed design, sequential sampling within a PSU can be completed immediately after the initial sample is taken and before leaving the PSU, which in our opinion would ease the implementation over the preceding designs. The two-stage adaptive cluster sampling design introduced by Salehi and Seber (1997) is quite different from the proposed two-stage sequential sampling design in terms of methodology and theory. Christman (2003) introduced adaptive two-stage one-per-stratum sampling and applied it to three small and artificial populations for which adaptive sampling was known to be appropriate. Our proposed design is efficient for a wide range of populations, and little or no efficiency is lost when applying it to populations that lack clustering. Thus, the two-stage sequential sampling design presented in this article is easier to implement and more flexible than those introduced before, and the approach used to derive estimators (Murthy’s estimator) opens the door for unbiased estimators to be found for similar sequential sampling designs.

To investigate the properties of the estimator, we simulated sampling of two real biological populations and one artificial population. Section 2 introduces the populations: a blue-winged teal population that is highly clustered, three freshwater mussel populations that are less clustered, and an artificial population chosen to mimic presence/absence data. Section 3 develops the methodology and presents notation and estimators along with example calculations. Section 4 describes the sampling simulations and presents the results. Section 5 discusses the results and makes suggestions for application of the proposed design.

Blue-winged Teal Population

								60	
				1			122	114	3
1				7144	6339		14		
			103	150	6				
			10						
					2				2
2									
				3					
		12							
		2			2				
3		4							
	5	20							
	3								
4									

Figure 1. Numbers of blue-winged teal as given by Smith et al. (1995) and a demonstration of two-stage sequential sampling. Units are partitioned into eight primary sampling units (PSUs), which are labeled along the left and right margins. Shaded units show results from two-stage sequential sampling. A simple random sample of four primary sampling units (PSUs) is selected from eight numbered PSUs. Two units are initially selected without replacement from each selected PSU (units shaded light gray). Because the criteria is met, four additional units are selected without replacement from PSU's one and five (units shaded dark gray).

2. EXAMPLE POPULATIONS

Smith, Conroy, and Brakhage (1995) used a population of blue-winged teal and populations of two other waterfowl species to evaluate adaptive cluster sampling. The populations came from comprehensive counts, which were made from helicopters during December 13–15, 1992, in central Florida. The blue-winged teal population is extremely clustered, with a total of $N = 200$ units (Figure 1). Adaptive cluster sampling was found by a simulation study to be efficient for the population, in the sense of having smaller variance of the estimator than that of simple random sampling (Smith et al. 1995).

Density of freshwater mussels is difficult to estimate well because of their tendency to be rare and clustered at some spatial scales (Strayer and Smith 2003). Smith et al. (2003) applied adaptive cluster sampling to low-density populations of freshwater mussels and found that adaptive cluster sampling increased observations of individuals and rare species. However, they also found that sampling of edge units greatly increased effort with little or no gain in efficiency in density estimates. Thus, an adaptive sampling procedure that reduces or eliminates sampling of edge units would be of great interest for freshwater mussel population assessments. An exhaustive search for freshwater mussels at the substrate surface was conducted in a section of the Cacapon River, West Virginia, during June 1994.

Individual mussels were counted and mapped to the nearest $.25 \text{ m}^2$ (Figure 2). The three species, *Elliptio complanata*, *E. fisheriana*, and *Lampsilis cariosa*, exhibited different spatial distributions. The *E. complanata* population was not rare, but it was relatively clustered. The *E. fisheriana* population was relatively rare and clustered. The *L. cariosa* population was rare, but not clustered. Application of two-stage sequential sampling to these populations will demonstrate design performance over a range of spatial distributions.

The artificial population is one for which we assume that defining a neighborhood is impossible. In the artificial population, the variable of interest is an indicator function

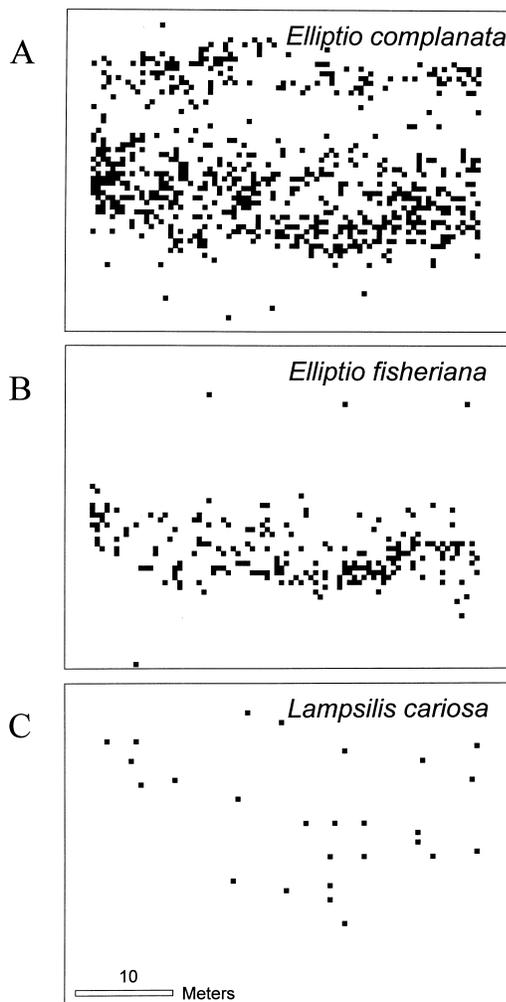


Figure 2. Populations of three species of freshwater mussels in a 40m section of the Cacapon River, West Virginia. An exhaustive search for mussels at the substrate surface was conducted during June 1994, and counts of individuals were recorded at a resolution of $.25 \text{ m}^2$. Panel A shows the distribution of *Elliptio complanata*. Panel B shows the distribution of *E. fisheriana*. Panel C shows the distribution of *Lampsilis cariosa*. River flow is from right to left.

(i.e., the variable takes either 0 or 1), which could represent observations of the presence and absence of a rare species, for example. Thompson and Seber (1996, p. 120 and 158) discussed that when the variable of interest is an indicator function adaptive cluster sampling is not an efficient sampling design. We will show that two-stage sequential sampling would be appropriate for such a population.

3. NOTATION AND ESTIMATOR

Suppose that we have a total population of N units, which are partitioned into M primary units of size N_i units ($i = 1, 2, \dots, M$). Based on our information about the rare clusters, their size and natural restrictions we choose sizes for the PSUs. We should note that precise information about their size is often not available. Let the unit (i, j) denote the j th unit in the i th primary unit with an associated measurement or count y_{ij} . Let $\tau_i = \sum_{j=1}^{N_i} y_{ij}$ be the sum of the y values in the i th primary unit, and let $\tau = \sum_{i=1}^M \tau_i$ be the total for the whole population.

In the first stage of the sampling, we choose a sample of m from the M primary units without replacement by some design with inclusion probability π_i for primary unit i and the joint inclusion probability $\pi_{ii'}$ for primary units i and i' . At the second stage, we take an initial simple random sample of n_{i1} units without replacement from primary unit i ($i = 1, 2, \dots, m$) so that $n_1 = \sum_{i=1}^m n_{i1}$ is the total initial sample size. Let C be the condition that if satisfied for at least one unit in primary unit i causes a predetermined number of additional units, say n_{i2} , to be selected at random from primary unit i . As a result, $n_2 = \sum_{i=1}^m n_{i2}$ is the number of adaptively added units and is a random variable. Let l_i be the number of units satisfying condition C in the final sample from primary unit i . When $m = M$, we have a stratified sequential sampling scheme.

Here we use Murthy's estimator (1957) to devise unbiased estimators for this design. Murthy's estimator is originally a Rao-Blackwell improvement of Raj's estimator (Raj 1956). Salehi and Seber (2001) showed that Murthy's estimator is also a Rao-Blackwell improvement of a trivial unbiased estimator, which can be used for sequential sampling designs. Let I_{ij} be an indicator function, which takes the values 1 (with probability p_{ij}) when unit j is chosen as the very first selected unit in PSU i and 0 otherwise.

$$\hat{t}_i = \sum_{j=1}^{N_i} \frac{y_{ij}}{p_{ij}} I_{ij}$$

is a trivial unbiased estimator of τ_i provided that $p_{ij} > 0$ for $j = 1, \dots, N_i$. Let s_i be the final sample set in the primary unit i . Using Rao-Blackwell theorem, we have Murthy's estimator

$$\hat{\tau}_i = E[\hat{t}_i | s_i] = \sum_{j \in s_i} \frac{P(s_i | j)}{P(s_i)} y_{ij}, \quad (3.1)$$

where $P(s_i)$ is the probability of obtaining the sample s_i in primary unit i and $P(s_i|j)$ is the conditional probability of getting the sample s_i given the j th unit was selected in the first draw in primary unit i . Two important measures of the quality of an estimator are the bias and the mean square error (Särndal, Swensson, and Wretman 1992, pp. 40). The latter becomes the variance when the bias is zero. Hájek (1971) resolutely expressed that greatly biased estimators are poor no matter what other properties they have. According to the Rao-Blackwell theorem estimator $\hat{\tau}_i$ is unbiased for τ_i if $p_{ij} > 0$ for $j = 1, \dots, N_i$. The variance of $\hat{\tau}_i$ is given by

$$\text{var}[\hat{\tau}_i] = \sum_{j=1}^{N_i} \sum_{j' < j}^{N_i} \left(1 - \sum_{s_i \ni j, j'} \frac{P(s_i|j)P(s_i|j')}{P(s_i)} \right) \left(\frac{y_{ij}}{p_{ij}} - \frac{y_{ij'}}{p_{ij'}} \right)^2 p_{ij} p_{ij'}.$$

Because we have $p_{ij} = n_{i1}/N_i$ for all $j = 1, 2, \dots, N_i$,

$$\text{var}[\hat{\tau}_i] = \sum_{j=1}^{N_i} \sum_{j' < j}^{N_i} \left(1 - \sum_{s_i \ni j, j'} \frac{P(s_i|j)P(s_i|j')}{P(s_i)} \right) (y_{ij} - y_{ij'})^2, \quad (3.2)$$

and its unbiased estimator is

$$\widehat{\text{var}}[\hat{\tau}_i] = \sum_{j \in s_i} \sum_{j' < j} \left(\frac{P(s_i|j, j')}{P(s_i)} - \frac{P(s_i|j)P(s_i|j')}{P(s_i)^2} \right) (y_{ij} - y_{ij'})^2, \quad (3.3)$$

where $P(s_i|j, j')$ is the probability of the sample s_i given that the units j and j' were selected (in either order) in the first two draws in primary unit i . A relatively simple proof of unbiasedness of (3.3) was given by Salehi and Seber (2001). It is assumed that $P(s_i|j, j')$ is well-defined. For two-stage sequential sampling Murthy's estimator provides an unbiased estimator for $\hat{\tau}_i$ since $p_{ij} > 0$ for $j = 1, \dots, N_i$.

For evaluating (3.1) we need to compute $P(s_i|j)/P(s_i)$ which is given by

$$\frac{P(s_i|j)}{P(s_i)} = \begin{cases} \frac{N_i}{n_{i1}} & n_{i2} = 0 \\ \frac{N_i}{n_{i1} + n_{i2}} & n_{i2} > 0 \text{ and } l_i > n_{i2} \\ \frac{N_i(n_{i1} + n_{i2} - 1)!}{(n_{i1} + n_{i2})! - n_{i2}!(n_{i1} + n_{i2} - l_i)! / (n_{i2} - l_i)!} & n_{i2} > 0 \text{ and } l_i \leq n_{i2} \\ & \text{and } j \text{ satisfies } C \\ \frac{N_i \{ (n_{i1} + n_{i2} - 1)! - n_{i2}!(n_{i1} + n_{i2} - 1 - l_i)! / (n_{i2} - l_i)! \}}{(n_{i1} + n_{i2})! - n_{i2}!(n_{i1} + n_{i2} - l_i)! / (n_{i2} - l_i)!} & n_{i2} > 0 \text{ and } l_i \leq n_{i2} \\ & \text{and } j \text{ not satisfy } C. \end{cases} \quad (3.4)$$

On substituting (3.4) into (3.1) we can compute $\hat{\tau}_i$. Note that when $n_{i2} = 0$ or $n_{i2} > 0$ and $l_i > n_{i2}$, $\hat{\tau}_i$ is N_i times the sample mean, which is the population total estimator for simple random sampling design.

For evaluating (3.3) we need to compute $P(s_i|j, j')/P(s_i)$ which is given by

$$\frac{P(s_i|j, j')}{P(s_i)} = \begin{cases} \frac{N_i(N_i-1)}{n_{i1}(n_{i1}-1)} & n_{i2} = 0 \\ \frac{N_i(N_i-1)}{(n_{i1}+n_{i2})(n_{i1}+n_{i2}-1)} & n_{i2} > 0 \text{ and } l_i > n_{i2} \\ \frac{N_i(N_i-1)(n_{i1}+n_{i2}-2)!}{(n_{i1}+n_{i2})!-n_{i2}!\frac{(n_{i1}+n_{i2}-l_i)!}{(n_{i2}-l_i)!}} & n_{i2} > 0 \text{ and } l_i \leq n_{i2} \text{ and} \\ & \text{either } j \text{ or } j' \text{ satisfies } C \\ \frac{N_i(N_i-1)\{(n_{i1}+n_{i2}-2)!-n_{i2}!\frac{(n_{i1}+n_{i2}-l_i-2)!}{(n_{i2}-l_i)!}\}}{(n_{i1}+n_{i2})!-n_{i2}!\frac{(n_{i1}+n_{i2}-l_i)!}{(n_{i2}-l_i)!}} & n_{i2} > 0 \text{ and } l_i \leq n_{i2} \text{ and} \\ & \text{neither } j \text{ nor } j' \text{ satisfy } C. \end{cases} \quad (3.5)$$

On substituting (3.4) and (3.5) into (3.3) we can compute $\widehat{\text{var}}[\hat{\tau}_i]$. Note that when $n_{i2} = 0$ or $n_{i2} > 0$ and $l_i > n_{i2}$, $\widehat{\text{var}}[\hat{\tau}_i]$ is an unbiased variance estimator of the population total for simple random sampling design. Proofs for (3.4) and (3.5) are given in the Appendix.

Because for two-stage sequential sampling the same sampling design is used to subsample each primary unit and subsampling is carried out independently, we can apply the theory of Horvitz-Thompson estimator (1952) for two-stage sampling to estimate population total. Thus, we have

$$\hat{\tau} = \sum_{i=1}^m \frac{\hat{\tau}_i}{\pi_i}, \quad (3.6)$$

which is unbiased for τ_i . Its variance is

$$\text{var}[\hat{\tau}] = \sum_{i=1}^M \sum_{i'=1}^M \left(\frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_i \pi_{i'}} \right) \tau_i \tau_{i'} + \sum_{i=1}^M \frac{\text{var}[\hat{\tau}_i]}{\pi_i}, \quad (3.7)$$

where π_{ii} is interpreted as π_i . An unbiased estimator of the above is

$$\widehat{\text{var}}[\hat{\tau}] = \sum_{i=1}^m \sum_{i'=1}^m \left(\frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_i \pi_{i'}} \right) \frac{\hat{\tau}_i \hat{\tau}_{i'}}{\pi_{ii'}} + \sum_{i=1}^m \frac{\widehat{\text{var}}[\hat{\tau}_i]}{\pi_i}, \quad (3.8)$$

where $\hat{\tau}_i$, $\text{var}[\hat{\tau}_i]$ and $\widehat{\text{var}}[\hat{\tau}_i]$ are, respectively, substituted by (3.1), (3.2), and (3.3). Because $\hat{\tau}_i$'s are unbiased for τ_i 's using the unbiasedness property of the Horvitz-Thompson estimator, the $\widehat{\text{var}}[\hat{\tau}]$ are also unbiased (Särndal, Swensson, and Wretman 1992, p. 137).

4. SIMULATIONS

We simulated two-stage sequential sampling on two real populations and one artificial population. The real populations were based on biological populations: blue-winged teal (Figure 1) and freshwater mussels (Figure 2). The artificial population was based on an indicator function and could represent the presence or absence of a rare event.

We used simple random sampling in the first stage to sample PSUs. In practice, if the sizes of PSUs are similar and auxiliary variables are not available, then use of simple

random sampling would be acceptable. Otherwise, the sizes of PSUs or other auxiliary variable could be incorporated in the design to increase efficiency, although that would affect the inclusion probabilities π_i 's. Using simple random sampling in the first stage, we have

$$\hat{\tau} = M \frac{1}{m} \sum_{i=1}^m \hat{\tau}_i = M \bar{\tau}, \quad (4.1)$$

its variance is

$$\text{var}[\hat{\tau}] = M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{M-1} \sum_{i=1}^M (\tau_i - \tau)^2 + \frac{M}{m} \sum_{i=1}^M \text{var}[\hat{\tau}_i], \quad (4.2)$$

with unbiased estimator

$$\widehat{\text{var}}[\hat{\tau}] = M^2 \left(1 - \frac{m}{M}\right) \frac{1}{m} \frac{1}{m-1} \sum_{i=1}^m (\hat{\tau}_i - \bar{\tau})^2 + \frac{M}{m} \sum_{i=1}^m \widehat{\text{var}}[\hat{\tau}_i]. \quad (4.3)$$

We computed relative efficiencies of two-stage sequential sampling over simple random sampling and traditional two-stage sampling. To compute variance of simple random sampling, say $\text{var}[\hat{\tau}_s]$, and of conventional two-stage sampling, say $\text{var}[\hat{\tau}_t]$, we set sample size equal to the effective sample size, namely $E[\nu]$; ν is final sample size for two-stage sequential sampling. Expected final sample size is

$$E[\nu] = \frac{m}{M} \sum_{i=1}^M \left\{ n_{i1} + n_{i2} \left(1 - \frac{(N_i - L_i)}{\binom{N_i}{n_{i1}}} \right) \right\},$$

where L_i is the number of units satisfying condition C in the primary unit i . We simulated 10,000 samples for given m , n_1 , and n_2 . If ν_k denotes the final sample size for replication k , we then have

$$E[\tau] = \bar{\tau} = \frac{1}{r} \sum_{k=1}^r \hat{\tau}_k \quad \text{and} \quad E[\nu] = \frac{1}{r} \sum_{k=1}^r \nu_k,$$

where r ($=10,000$) is the number of replications. We then calculated

$$\text{var}(\hat{\tau}) = \frac{1}{r-1} \sum_{k=1}^r (\hat{\tau}_k - \bar{\tau})^2.$$

For all cases Monte Carlo errors, $(\bar{\tau} - \tau)$, was very close to zero.

4.1 BLUE-WINGED TEAL POPULATION

We partitioned the blue-winged teal population into $M = 8$ PSUs (Figure 1) to compare results from Salehi and Seber (1997). In Section 5, we discuss the rationale for choice of M . As it is shown in Figure 1, each PSU contained $N_i = 25$ units. For one set of cases

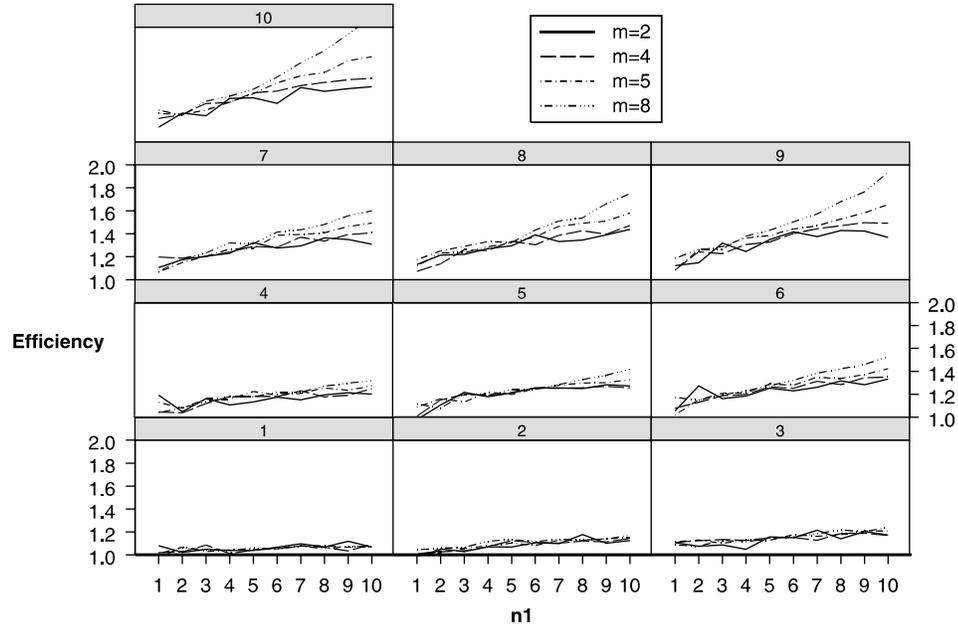


Figure 3. The relative efficiency of estimator $\hat{\tau}$ for two-stage sequential sampling over estimator $\hat{\tau}_t$ for conventional two-stage sampling with the same m and sample size $E[\nu]/m$ in the selected PSUs. Each panel represents a different value for n_{i2} . The target population was the blue-winged teal population presented in Figure 1.

we let $m = 2, 4, 6, 8$; $n_{i1} = 1, 2, \dots, 10$; and $n_{i2} = 1, 2, \dots, 10$. In another set of cases where $m = 2, 4, 6, 8$; $n_{i1} = 1, 2, 3, 4, 5$; and $n_{i2} = 13, 14, \dots, 20$, we computed $E[\nu]$ and $\text{var}[\hat{\tau}]$ using 10,000 replications for all 560 cases. The condition for sequential sampling was $y_{ij} > 10$. We computed $\text{var}[\hat{\tau}_s]$ with sample size equal to the effective sample size, $E[\nu]$, and computed $\text{var}[\hat{\tau}_t]$ with a sample size m of PSU's and a sample size of $E[\nu]/m$ units in each selected PSU. We defined the efficiency of two-stage sequential sampling relative to simple random sampling and conventional two-stage sampling as the following

$$\text{eff}_s[\hat{\tau}] = \frac{\text{var}[\hat{\tau}_s]}{\text{var}[\hat{\tau}]} \quad \text{eff}_t[\hat{\tau}] = \frac{\text{var}[\hat{\tau}_t]}{\text{var}[\hat{\tau}]}$$

The plot of $\text{eff}_t[\hat{\tau}]$ for different values of n_{i1} and n_{i2} is given in Figure 3. The range of $\text{eff}_t[\hat{\tau}]$ was from .93 to 2.1 or equivalently a 7% loss to 110% gain in efficiency over conventional two-stage sampling. In only 5 of the 560 cases was $\text{eff}_t[\hat{\tau}]$ less than one. It was .93 for $m = 2, n_{i1} = 1, n_{i2} = 2$; .97 for $m = 2, n_{i1} = 1, n_{i2} = 5$; .98 for $m = 4, n_{i1} = 8, n_{i2} = 1$; .95 for $m = 6, n_{i1} = 1, n_{i2} = 2$; and .98 for $m = 8, n_{i1} = 1, n_{i2} = 1$. Efficiency was an increasing function of m, n_{i1} , and n_{i2} .

The plot of $\text{eff}_s[\hat{\tau}]$ for different values of $E[\nu]$ is given in Figure 4. The range of $\text{eff}_s[\hat{\tau}]$ was from .93 to 2.09 which means a 7% loss to 109% gain in efficiency over simple random sample. In only four of the cases was $\text{eff}_s[\hat{\tau}]$ less than one; all cases had $n_{i1} = 1$. It was .93 for $m = 2, n_{i1} = 1, n_{i2} = 2$; .98 for $m = 2, n_{i1} = 1, n_{i2} = 5$; .95 for $m = 6, n_{i1} = 1,$

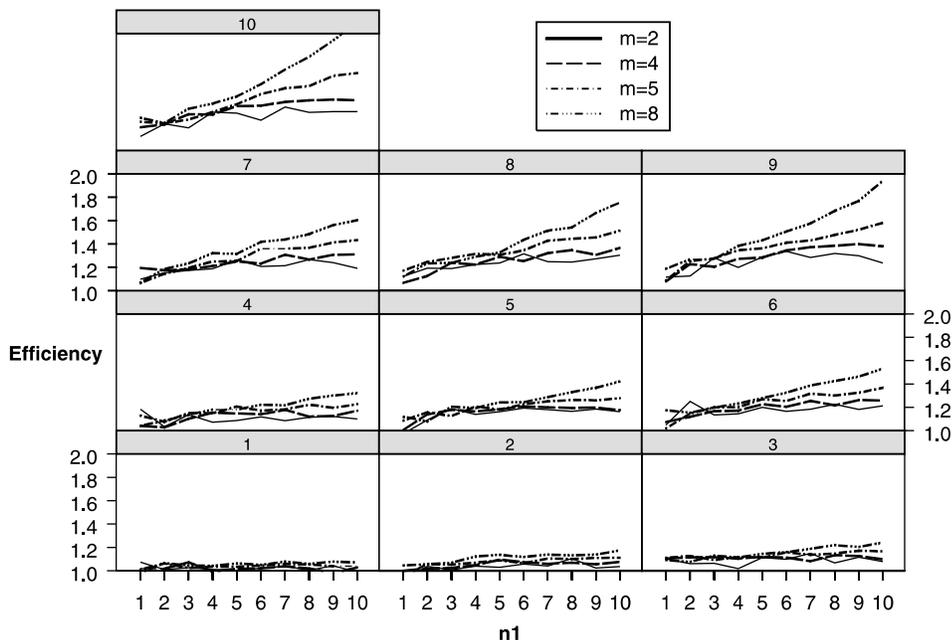


Figure 4. The relative efficiency of estimator τ for two-stage sequential sampling over estimator $\hat{\tau}_s$ for simple random sampling with the same effective sample size $E[\nu]$. Each panel represents a different value for n_{i2} . The target population was the blue-winged teal population presented in Figure 1.

$n_{i2} = 2$; and .98 for $m = 8$, $n_{i1} = 1$, $n_{i2} = 1$. Efficiency was an increasing function of m , n_{i1} , and n_{i2} .

To compare two-stage adaptive cluster sampling (Salehi and Seber 1997) and two-stage sequential sampling, we computed variances of the modified Hansen-Hurwitz estimator, say $\text{var}[\hat{\tau}_{HH}]$, the modified Horvitz-Thompson estimator, say $\text{var}[\hat{\tau}_{HT}]$, and the effective sample size for two-stage adaptive cluster sampling, say $E[\nu^*]$, for different initial sample sizes. Because sequentially added units, n_{i2} , are confined in the selected primary units it is more appropriate to compare $\text{var}[\hat{\tau}]$ with variances of estimators for the nonoverlapping scheme of the two-stage adaptive design. For $m = 4, 8$, $n_{0i} = 1, 2, \dots, 10$ (i.e., the initial sample size from each primary sample unit), and with the condition $y_{ij} > 10$, we computed the variances of estimators and the effective sample sizes for two-stage adaptive cluster sampling, $E[\nu^*]$. We defined

$$\text{eff.}[\hat{\tau}_{HH}] = \frac{\text{var}[\hat{\tau}]}{\text{var}[\hat{\tau}_{HH}]} \quad \text{eff.}[\hat{\tau}_{HT}] = \frac{\text{var}[\hat{\tau}]}{\text{var}[\hat{\tau}_{HT}]}$$

and in place of “.” we used s to denote simple random sampling and t to denote conventional two-stage sampling.

For specific m and $E[\nu^*]$, we found the closest $E[\nu]$ with the same m from our study, and the results are given in Table 1. As expected the efficiency of the HT estimator is greater than that of the HH estimator (Salehi 2003). For 16 out of the 20 cases, the $\text{eff.}[\hat{\tau}]$'s were greater than $\text{eff.}[\hat{\tau}_{HT}]$'s (Table 1).

Table 1. The Relative Efficiencies of Estimators for Population of Blue-Winged Teal in Figure 1. Estimators are HH and HT for two-stage adaptive cluster sampling design with nonoverlapping scheme. Efficiencies for $m = 4, 8$ and initial sample size $n_{j0} = 1, 2, \dots, 10$ are presented in columns 7, 8, 9, and 10. The effective sample sizes, $E[\nu^*]$, are given in column 3. From our study given in Figure 3 and 4, we choose the closest effective sample sizes, $E[\nu]$, with the same m for two-stage adaptive sampling, and their relative efficiencies are given in columns 11 and 12.

m	n_{0i}	$E[\nu^*]$	n_{1i}	n_{2i}	$E[\nu]$	$eff_t[\hat{\tau}_{HT}]$	$eff_t[\hat{\tau}_{HH}]$	$eff_s[\hat{\tau}_{HT}]$	$eff_s[\hat{\tau}_{HH}]$	$eff_t[\hat{\tau}]$	$eff_s[\hat{\tau}]$
4	1	5.44	1	7	5.42	1.26	1.26	1.26	1.26	1.19	1.20
4	2	10.64	2	7	10.62	1.28	1.26	1.28	1.26	1.17	1.19
4	3	15.62	2	20	15.63	1.30	1.26	1.30	1.26	1.53	1.56
4	4	20.41	3	16	20.57	1.31	1.26	1.32	1.26	1.46	1.50
4	5	25.02	4	14	25.39	1.33	1.26	1.34	1.27	1.45	1.49
4	6	29.48	4	20	29.51	1.34	1.25	1.36	1.27	1.62	1.68
4	7	33.80	5	17	33.53	1.35	1.25	1.37	1.27	1.60	1.67
4	8	37.99	7	10	37.81	1.36	1.25	1.38	1.26	1.42	1.49
4	9	42.08	8	10	42.70	1.36	1.25	1.38	1.26	1.43	1.52
4	10	46.07	9	9	46.24	1.36	1.24	1.39	1.26	1.40	1.50
8	1	10.88	1	7	10.84	1.27	1.27	1.26	1.26	1.18	1.17
8	2	21.28	2	7	21.29	1.30	1.28	1.29	1.27	1.19	1.18
8	3	31.25	2	20	31.22	1.33	1.29	1.33	1.28	1.62	1.62
8	4	40.82	3	16	41.26	1.37	1.30	1.36	1.30	1.57	1.56
8	5	50.05	4	13	49.40	1.40	1.31	1.40	1.31	1.59	1.59
8	6	58.75	4	20	58.83	1.44	1.32	1.44	1.32	1.99	1.98
8	7	67.60	7	7	68.02	1.48	1.32	1.47	1.33	2.08	2.08
8	8	75.98	7	10	75.66	1.52	1.34	1.51	1.34	1.69	1.69
8	9	84.15	8	9	83.23	1.55	1.35	1.55	1.35	1.68	1.68
8	10	92.13	9	9	92.44	1.59	1.36	1.59	1.36	1.77	1.76

4.2 FRESHWATER MUSSEL POPULATIONS

The region for freshwater mussel populations was partitioned into 4,800 quadrats (Figure 2). Simulations for 24, 48, and 96 PSUs were conducted. Because results were similar in each case, we show results for 48 PSUs of size 100 quadrats each. We computed the efficiency of two-stage sequential sampling compared to conventional two-stage sampling, $eff_t[\tau_{HT}]$, for $m = 35, 48$; $n_{i1} = 4, 6, 8, 10, 12, 14$; and $n_{i2} = 2, 4, 6$. These cases were selected to arrive at sample sizes typical of freshwater mussel surveys.

For the clustered but not rare population of *E. complanata*, two-stage sequential sampling was efficient for 32 out of the 36 cases (Figure 5). Losses in efficiency occurred when $n_{i1} = 4$ and $n_{i2} = 6$.

For the clustered and rare population of *E. fisheriana*, two-stage sequential sampling was efficient in all cases (Figure 5). Efficiency ranged from 1.02 to 1.18. There did not appear to be much gain in efficiency for $n_{i1} > 10$. Efficiency tended to increase as m and n_{i2} increased.

For the rare but not clustered population of *L. cariosa*, two-stage sequential sampling did not result in a major gain or loss in efficiency (Figure 5). Efficiency ranged from .97 to 1.03. Because the population is not clustered, extra sampling in the vicinity of observed mussels does not increase the likelihood of finding additional mussels. Thus, we did not

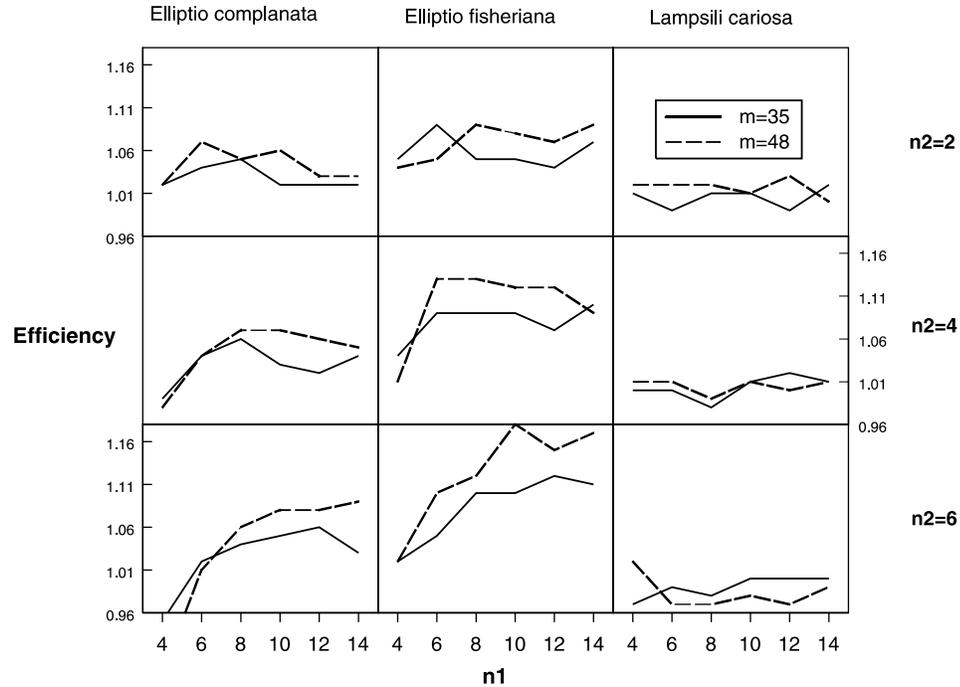


Figure 5. The relative efficiency of estimator $\hat{\tau}$ for two-stage sequential sampling over estimator $\hat{\tau}_t$ for conventional two-stage sampling with the same $m = 35, 48$ and sample size $E[\nu]/m$ in the selected PSUs. The target populations were *Elliptio fisheriana*, *Elliptio complanata*, and *Lampsilis cariosa* populations presented in Figure 2.

expect a gain in efficiency. However, there was not a loss in efficiency either due to the application of two-stage sequential sampling.

4.3 INDICATOR FUNCTION (PRESENCE/ABSENCE) POPULATION

To demonstrate properties of two-stage sequential sampling for an indicator function population (e.g., population comprised of the presence or absence of a rare event), we considered a population of size 5,000, which was partitioned into 50 PSUs of size 100. There were no rare events in 46 PSU's and four remaining units contained 1, 2, 47, and 55 rare events. We considered the objective of estimating population total, $\tau = 105$.

We computed $E[\nu]$, $eff_s[\hat{\tau}]$ and $eff_t[\hat{\tau}]$, for $m = 10, 20, 30, 40, 50$, $n_{i1} = 1, 2, 3, 4$, and $n_{i2} = 1, 2, 3, 4, 5, 10, 20, 30$. The results for $m = 40, 50$, $n_{i1} = 1, 2, 3, 4$, and $n_{i2} = 2, 4, 10, 20$ are presented in Table 2. In all cases the $eff_s[\hat{\tau}]$ were greater than one. Efficiency ranged from 8% to 189% when the comparison was to conventional two-stage sampling and from 5% to 459% when the comparison was to simple random sampling. The result in the last row of Table 2 was for a two-stage sequential sampling with $m = 50$, $n_{i1} = 5$, and $n_{i2} = 20$ for which the effective sample size was 291.8 (i.e., effective sample fraction was less than 6%) and the $CV[\hat{\tau}] = \sqrt{\text{var}[\hat{\tau}]} / \hat{\tau}$ was less than 16.5%.

Table 2. Relative Efficiency for the Indicator Function Population of Size 5,000, Which is Partitioned into 50 Primary Units of Size 100. There is no rare event in 46 primary units and four remaining units contain 1, 2, 47, 55 rare events. For different values of m , $n_{i,1}$, and $n_{i,2}$, the variances of population total τ were computed using simulation, $\text{var}[\hat{\tau}]$. The variances of simple random samples, $\text{var}[\hat{\tau}_s]$, and two-stage sampling, $\text{var}[\hat{\tau}_t]$, were computed with the same effective sample size $E[\nu]$. The relative efficiencies, $\text{eff}_t[\hat{\tau}] = \text{var}[\hat{\tau}]/\text{var}[\hat{\tau}_t]$ and $\text{eff}_s[\hat{\tau}] = \text{var}[\hat{\tau}]/\text{var}[\hat{\tau}_s]$, are given in column 6 and 7.

$E[\nu]$	m	$n_{i,1}$	$n_{i,2}$	$\text{eff}_t[\hat{\tau}]$	$\text{eff}_s[\hat{\tau}]$
82.54	40	2	2	1.16	1.60
85.04	40	2	4	1.20	1.64
92.68	40	2	10	1.14	1.52
105.20	40	2	20	1.07	1.37
123.01	40	3	2	1.17	1.42
125.97	40	3	4	1.25	1.50
134.84	40	3	10	1.32	1.54
149.66	40	3	20	1.28	1.42
163.22	40	4	2	1.18	1.27
166.40	40	4	4	1.24	1.32
176.06	40	4	10	1.35	1.39
191.97	40	4	20	1.42	1.41
203.33	40	5	2	1.14	1.10
206.71	40	5	4	1.23	1.17
216.82	40	5	10	1.33	1.24
233.46	40	5	20	1.43	1.27
103.15	50	2	2	1.22	2.37
106.35	50	2	4	1.27	2.45
115.78	50	2	10	1.21	2.35
131.78	50	2	20	1.11	2.14
153.73	50	3	2	1.35	2.61
157.39	50	3	4	1.48	2.86
168.55	50	3	10	1.71	3.31
187.22	50	3	20	1.63	3.16
204.02	50	4	2	1.34	2.59
208.04	50	4	4	1.64	3.18
220.00	50	4	10	2.02	3.91
240.26	50	4	20	2.28	4.41
254.19	50	5	2	1.34	2.60
258.36	50	5	4	1.63	3.15
270.98	50	5	10	2.21	4.27
291.80	50	5	20	2.89	5.59

5. DISCUSSION

The two-stage sequential sampling design is adaptive because the final sample depends on observed values, but it is free of the constraints imposed by neighborhood-based adaptive sampling. As a result, two-stage sequential sampling avoids edge units, which have limited the efficiency of adaptive cluster sampling in some cases. In addition, the design restrains the final sample size because selection of adaptive units is restricted to be no more than $n_{i,2}$ in each PSU. In a sense, two-stage sequential sampling is a “restricted” adaptive sampling design (see Brown and Manly 1998). Unlike restricted adaptive cluster sampling, however, the estimators for two-stage sequential sampling are unbiased due to our use of Murthy’s estimator.

We first intended to develop a sampling design for which an additional sample is selected whenever at least a predetermined number of selected units in primary sample unit i , say n_{ij}^* , satisfy condition C . We found that by setting $n_{ij}^* = 1$ and changing the condition C and n_{i2} , we achieved sampling designs with similar properties to two-stage sequential sampling with n_{ij}^* greater than 1. Fixing $n_{ij}^* = 1$ greatly simplified the formulation.

A wide range of sampling designs are possible through the choice of a method to select PSUs, condition C , n_{i1} , and n_{i2} . We evaluated many more choices for n_{i1} and n_{i2} than those given in this article under a simple design for PSU selection and found that nearly all choices lead to greater efficiency than either two-stage adaptive cluster sampling or conventional two-stage sampling. Improvement over neighborhood-based adaptive cluster sampling designs was most evident when effective sample sizes were moderate to large, for example, sampling fraction $> .05$. When effective sample sizes were relatively small, two-stage adaptive cluster sampling was more efficient than two-stage sequential sampling (see Table 1). Further development of the two-stage sequential sampling design could focus on designs where n_{i1} and n_{i2} vary among PSUs and on designs with unequal sized PSUs.

The most important step when applying sequential two-stage sampling is to partition the population into PSUs in such a way that confines rare events to a few PSUs. For these few PSUs a large proportion of their units will contain the rare events. The “optimal” design when the rare events are clusters is to have PSU’s that are the same size as the clusters so that there is one cluster per PSU. We then select relatively small n_{i1} and relatively large n_{i2} . For example, we partitioned the blue-winged teal population into PSUs ranging from 50 PSUs of 4 units to 2 PSUs of 100 units and found two-stage sequential sampling to be an efficient sampling design for all cases. However, the efficiency of two-stage sequential sampling relative to simple random sampling, $\text{eff}_s[\hat{\tau}]$, ranged from 1.005 to a stunning 2,051 and the efficiency of two-stage sequential sampling relative to two-stage sampling, $\text{eff}_t[\hat{\tau}]$, ranged from 1.006 to an equally stunning 1,350. The most efficient sampling design corresponded to 50 PSU’s of size 4 units with condition $c > 0$, $n_{i1} = 1$, and $n_{i2} = 3$. In this design, one PSU contains units with counts of 7,144, 6,339, 150, and 6. This PSU has 13,639 out of the 14,121 blue-winged teals in the population. All the units in this PSU will be in the sample set with probability one in this sampling design. We note that identifying the *most* efficient design requires more knowledge about the population than is often available in practice. Efficient designs are attainable with little prior information. However, when some information about the rare clusters is available, then we recommend choosing PSUs so that there is one cluster per PSU and the PSU is roughly the size of the cluster.

We can provide general guidelines for choosing appropriate values for C , n_{i1} , and n_{i2} . Suppose that the condition has the form of $y_{ij} > c$ and consider the two extreme cases. If we choose c too small such that almost all y_{ij} ’s become greater than c , then two-stage sequential sampling would tend to be more like conventional two-stage sampling with $n_i = n_{i1} + n_{i2}$. If we choose c too big such that almost all y_{ij} ’s are smaller than c , then two-stage sequential sampling would tend to be more like conventional two-stage sampling with $n_i = n_{i1}$. If the variable of interest is an indicator function (i.e., takes either 0 or 1), there is no problem

choosing the condition C . However, there should be a set of optimal (or nearly optimal) values for different population models, and simulation may be required to find these values.

Based on our study, larger value of n_{i1} 's and smaller value of n_{i2} 's resulted in small improvement in efficiency, but the improvement was observed for almost all of the populations under study. For some cases, small values of n_{i1} 's and large values of n_{i2} 's resulted in a great improvement in efficiency, but there were some cases where such a design resulted in a loss of efficiency. Thus, if we do not know how the degree of rarity and clustering in the population, then we recommend moderate or larger values of n_{i1} 's and relatively small values of n_{i2} 's. If we know the population is highly clustered and these clusters are confined in a few PSUs we recommend setting n_{i2} to be a large proportion of N_i . It is important to mention that if we fail to make optimal choices for PSUs, C , n_{i1} , and n_{i2} efficiency of sequential two-stage sampling would be similar to conventional two-stage sampling.

When a rare and clustered population is partitioned into PSUs such that there is no more than one cluster per PSU and the PSU is roughly the size of the cluster, most PSUs will be empty and a few will contain rare events. Hence, the variance within PSUs will be less than the variance between PSUs. For such a population, stratified sampling is an efficient design (two-stage sampling with $m = M$). This was true for the *E. fisheriana* population where greatest efficiency was found when $m = M$. We conclude that two-stage sequential sampling with larger m can therefore be more efficient. Our simulation study confirms this fact. For the indicator-function population when m was small ($m = 5, 10, 20$) even though two-stage sequential sampling was more efficient than the conventional two-stage sampling, it was less efficient than simple random sampling.

For $m = M$, two-stage sequential sampling provides an optimal-type allocation even when a priori information on PSU-specific variance is unavailable. When condition C has the form of $y_{ij} > c$ or $y_{ij} < c$, as it is frequently, and those units satisfying condition C have very large or very small values, the rare units will have y_{ij} far from the mean of the i th PSU. Two-stage sequential sampling therefore allocates more samples to the primary sample units (strata) with larger standard deviations. It is in line with Neyman allocation for stratified sampling (Cochran 1977). We can therefore expect that sample variance from two-stage sequential sampling will be similar to stratified sampling with optimal allocation. To implement stratified sampling with optimal allocation we would need reliable estimates of strata-specific standard deviations, which might not be available. However, optimal-type allocation is a natural product of the two-stage sequential sampling design.

An obvious alternative to the two-stage sequential sampling with $m = M$ is stratified sampling in which a pilot survey is conducted using proportional allocation, then pilot survey results are used to calculate optimal allocation and strata are sampled accordingly. Environmental and ecological surveys are usually time consuming and logistically expensive. A stratified sampling design that includes a pilot survey would require two separate field trips. We must have the data from all strata taken during the first trip to calculate optimal allocation. We would then return to the strata to conduct the optimal stratified sample. It could be quite prohibitive in term of time or logistics to conduct such a survey. Using

the two-stage sequential sampling with $m = M$, a near-optimal stratified sample could be taken in one field trip and we just travel inside PSUs for selecting the adaptively selected n_{i2} units. The decision to take extra samples could be made prior to leaving a stratum. A potential inefficiency would be due to having to traverse the same area when sampling the n_{i2} 's after already sampling the n_{i1} 's.

Finally, it is often desirable and natural to partition the population under study into PSUs regardless of the issue of sampling rare and clustered populations. For example, if a rare event is related to socio-economic status, then districts or blocks are appropriate partitions of a city. Sampling freshwater mussels is expensive and labor intensive especially in deep water where SCUBA diving is required (Strayer and Smith 2003). In rivers, SCUBA divers must work close to an anchored boat for safety, and repositioning the boat is time consuming. Thus, it is practical to sample in stages by first selecting a primary unit to position the boat and then selecting secondary units for collection of freshwater mussels. If we want to estimate population size of a rare species, then the amount of habitat is an appropriate way to gauge the partitioning of the study area into PSUs. If species distribution is known to be restricted to a small fraction of a site and auxiliary information is available, a stratified sampling could be designed to confine rare units to a few strata (Kalton and Anderson 1986; Christman 2000). However, if species distribution is unknown or relevant auxiliary information is not in the sampling frame, then two-stage sequential sampling can still be an efficient sampling design by partitioning the population in a haphazard fashion for rare and highly clustered populations.

APPENDIX

Evaluation of $\frac{P(s_i|j)}{P(s_i)}$. If $n_{i2} = 0$, $P(s_i)$ would be $1/\binom{N_i}{n_{i1}}$ and $P(s_i|j)$ would be $1/\binom{N_i-1}{n_{i1}-1}$. Thus

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i}{n_{i1}}.$$

If $n_{i2} > 0$ and $l_i > n_{i2}$ there would be at least one sample unit satisfying C which is placed in the n_{i1} first samples for all possible permutation of $n_{i1} + n_{i2}$ sample units. This means that all possible initial samples of size n_{i1} give rise to select another sample of size n_{i2} . Hence, $P(s_i) = 1/\binom{N_i}{n_{i1}+n_{i2}}$ and $P(s_i|j) = 1/\binom{N_i-1}{n_{i1}+n_{i2}-1}$. Thus

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i}{n_{i1} + n_{i2}}.$$

If $n_{i2} > 0$, $l_i \leq n_{i2}$ and j satisfies C , the probability of choosing an ordered sample giving rise to s_i is $1/(N_i \times N_i - 1 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)$. Those permutations without a sample unit satisfying C in the n_{i1} first samples do not give rise to s_i . To find the number of permutation not giving rise to s_i , we first allocate all the l_i sample units satisfying C to the n_{i2} last samples which can be done by $n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1$

and we allocate the $n_{i1} + n_{i2} - l_i$ remaining sample units to the rest of places which can be done by $(n_{i1} + n_{i2} - l_i)!$. Hence the number of permutations giving rise to s_i is $(n_{i1} + n_{i2})! - (n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - l_i)!$. Thus $P(s_i) = \{(n_{i1} + n_{i2})! - (n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - l_i)!\} / \{(N_i \times N_i - 1 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)\}$. Because unit j satisfies C all permutations of the remaining sample units give rise to s_i , $P(s_i|j) = 1 / \binom{N_i - 1}{n_{i1} + n_{i2} - 1}$. Thus

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i(n_{i1} + n_{i2} - 1)!}{(n_{i1} + n_{i2})! - n_{i2}!(n_{i1} + n_{i2} - l_i)!(n_{i2} - l_i)!}$$

When $n_{i2} > 0$, $l_i \leq n_{i2}$ and j does not satisfy C , $P(s_i)$ is the same as the previous case. The probability of choosing an ordered sample giving rise to $s_i|j$ is $1 / (N_i - 1 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)$. Because j does not satisfy C those permutations without a sample unit satisfying C in the remaining $n_{i1} - 1$ first samples do not give rise to $s_i|j$. To find the number of permutations not giving rise to $s_i|j$, we first allocate all the l_i sample units satisfying C to the n_{i2} last samples which can be done by $n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1$ and we allocate the $n_{i1} + n_{i2} - l_i - 1$ remaining sample units to the rest of places which can be done by $(n_{i1} + n_{i2} - 1 - l_i)!$. Hence the number of permutations giving rise to s_i is $(n_{i1} + n_{i2} - 1)! - (n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - 1 - l_i)!$. Thus, $P(s_i|j) = \{(n_{i1} + n_{i2} - 1)! - (n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - 1 - l_i)!\} / \{(N_i - 1 \times N_i - 2 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)\}$. Thus

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i \{(n_{i1} + n_{i2} - 1)! - n_{i2}!(n_{i1} + n_{i2} - 1 - l_i)!(n_{i2} - l_i)!\}}{(n_{i1} + n_{i2})! - n_{i2}!(n_{i1} + n_{i2} - l_i)!(n_{i2} - l_i)!}$$

Evaluation of $\frac{P(s_i|j,j')}{P(s_i)}$. If $n_{i2} = 0$, $P(s_i)$ would be $1 / \binom{N_i}{n_{i1}}$ and $P(s_i|j, j')$ would be $1 / \binom{N_i - 2}{n_{i1} - 2}$. Thus,

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i(N_i - 1)}{n_{i1}(n_{i1} - 1)}$$

If $n_{i2} > 0$ and $l_i > n_{i2}$, $P(s_i) = 1 / \binom{N_i - 2}{n_{i1} + n_{i2}}$ and $P(s_i|j, j') = 1 / \binom{N_i - 2}{n_{i1} + n_{i2} - 2}$. Thus,

$$\frac{P(s_i|j)}{P(s_i)} = \frac{N_i(N_i - 1)}{(n_{i1} + n_{i2})(n_{i1} + n_{i2} - 1)}$$

If $n_{i2} > 0$, $l_i \leq n_{i2}$ and either j or j' satisfies C , $P(s_i) = \{(n_{i1} + n_{i2})! - (n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - l_i)!\} / \{(N_i \times N_i - 1 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)\}$. Since either j or j' satisfies C all permutations of the remaining sample units give rise to s_i , $P(s_i|j, j') = 1 / \binom{N_i - 2}{n_{i1} + n_{i2} - 2}$. Thus,

$$\frac{P(s_i|j, j')}{P(s_i)} = \frac{N_i(N_i - 1)(n_{i1} + n_{i2} - 2)!}{(n_{i1} + n_{i2})! - n_{i2}!(n_{i1} + n_{i2} - l_i)!(n_{i2} - l_i)!}$$

If $n_{i2} > 0$, $l_i \leq n_{i2}$ and neither j nor j' satisfy C , $P(s_i)$ is the same as the previous case. The probability of choosing an ordered sample giving rise to $s_i|j, j'$ is $1 / (N_i - 2 \times \cdots \times$

$N_i - n_{i1} - n_{i2} + 1$). Because neither j nor j' satisfy C those permutations without a sample unit satisfying C in the remaining $n_{i1} - 2$ first samples do not give rise to $s_i|j, j'$. To find the number of permutations not giving rise to $s_i|j, j'$, we first allocate all the l_i sample units satisfying C to the n_{i2} last samples which can be done by $n_{i2} \times n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1$ and we allocate the $n_{i1} + n_{i2} - l_i - 2$ remaining sample units to the rest of places which can be done by $(n_{i1} + n_{i2} - 2 - l_i)!$. Hence the number of permutations giving rise to s_i is $(n_{i1} + n_{i2} - 1)! - (n_{i2} - 1 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - 1 - l_i)!$. Thus $P(s_i|j, j') = \{(n_{i1} + n_{i2} - 2)! - (n_{i2} - 2 \times n_{i2} - 3 \times \cdots \times n_{i2} - l_i + 1)(n_{i1} + n_{i2} - 2 - l_i)!\} / \{(N_i - 1 \times N_i - 2 \times \cdots \times N_i - n_{i1} - n_{i2} + 1)\}$. Thus,

$$\frac{P(s_i|j, j')}{P(s_i)} = \frac{N_i(N_i - 1)\{(n_{i1} + n_{i2} - 2)! - n_{i2}! \frac{(n_{i1} + n_{i2} - l_i - 2)!}{(n_{i2} - l_i)!}\}}{(n_{i1} + n_{i2})! - n_{i2}! \frac{(n_{i1} + n_{i2} - l_i)!}{(n_{i2} - l_i)!}}.$$

ACKNOWLEDGMENTS

M. Salehi's research is partially supported by the Research Council of Isfahan University of Technology. D.R. Smith appreciates the research support provided by the Aquatic Ecology Branch of the U.S. Geological Survey, Leetown Science Center. The authors thank the editors and referees for their helpful comments.

[Received July 2003. Revised July 2004.]

REFERENCES

- Brown, J. A., and Manly, B. F. J. (1998), "Restricted Adaptive Cluster Sampling," *Environmental and Ecological Statistics*, 5, 49–63.
- Christman, M. C. (2000), "A Review of Quadrat-Based Sampling of Rare, Geographically Clustered Populations," *Journal of Agricultural, Biological and Environmental Statistics*, 5, 168–201.
- (2003), "Adaptive Two-Stage One-per-Stratum Sampling," *Environmental and Ecological Statistics*, 10, 43–60.
- Cochran, W. G. (1977), *Sampling Techniques* (3rd ed.), New York: Wiley.
- 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.
- Haldane, J. B. S. (1945), "On a Method of Estimating Frequencies," *Biometrika*, 33, 222–225.
- Hájek, J. (1971), Comment on Basu, in *Foundation of Statistical Inference*, eds. V. P. Godambe and D. A. Sprott, Toronto: Holt, Rinehart and Winston.
- Hanselman, D. H., Quinn II, T. J., Lunsford, C., Heifetz, J., and Clausen, D. (2003), "Applications in Adaptive Cluster Sampling of Gulf of Alaska Rockfish," *Fisheries Bulletin*, 101, 501–513.
- Horvitz, D. G., and Thompson, D. J. (1952), "A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, 47, 663–685.
- Jolly, G. M., and Hampton, I. (1990), "A Stratified Random Transect Design for Acoustic Surveys of Fish Stocks," *Canadian Journal of Fisheries and Aquatic Sciences*, 47, 1282–1291.
- Kalton, G., and Anderson, D. W. (1986), "Sampling Rare Populations," *Journal of the Royal Statistical Society, Ser. A*, 149, 65–82.

- Lo, N. C. H., Griffith, D., and Hunter, J. R. (1997), "Using Restricted Adaptive Cluster Sampling to Estimate Pacific Hake Larval Abundance," *CalCOFI Report*, 37, 160–174.
- Murthy, M. N. (1957), "Ordered and Unordered Estimators in Sampling Without Replacement," *Sankhyā*, 18, 379–390.
- Raj, D. (1956), "Some Estimators in Sampling With Varying Probabilities Without Replacement," *Journal of the American Statistical Association*, 51, 269–284.
- Salehi, M. M. (1999), "Rao-Blackwell Versions of the Horvitz-Thompson and Hansen-Hurwitz in Adaptive Cluster Sampling," *Environmental and Ecological Statistics*, 6, 183–195.
- (2003), "Comparison Between Hansen-Hurwitz and Horvitz-Thompson for Adaptive Cluster Sampling," *Environmental and Ecological Statistics*, 10, 115–127.
- Salehi M. M., and Seber, G. A. F. (1997), "Two-Stage Adaptive Cluster Sampling," *Biometrics*, 53, 959–970.
- (2001), "A New Proof of Murthy's Estimator Which Applies to Sequential Sampling," *Australian & New Zealand Journal of Statistics*, 43, 901–906.
- Särndal, C. E., Swensson, B., and Wretman, J. (1992), *Model Assisted Survey Sampling*, New York: Springer-Verlag.
- Smith, D. R., Conroy, M. J., and Brakhage, D. H. (1995), "Efficiency of Adaptive Cluster Sampling for Estimating Density of Wintering Waterfowl," *Biometrics*, 51, 777–788.
- Smith, D. R., Vilella, R. F., and Lemarie, D. P. (2003), "Application of Adaptive Cluster Sampling to Low-Density Populations of Freshwater Mussels," *Environmental and Ecological Statistics*, 10, 7–15.
- Smith, D. R., Brown, J. A., and Lo, N. C. H. (2004), "Application of Adaptive Cluster Sampling to Biological Populations," in *Sampling Rare or Elusive Species*, ed. W. L. Thompson, Covelo, CA: Island Press, pp. 93–152.
- Strayer, D. L., and Smith, D. R. (2003), *A Guide to Sampling Freshwater Mussel Populations*, Bethesda, MD: American Fisheries Society.
- Thompson, S. K. (1990), "Adaptive Cluster Sampling," *Journal of the American Statistical Association*, 85, 1050–1059.
- Thompson, S. K., and Seber, G. A. F. (1996), *Adaptive Sampling*, New York: Wiley.