

ASYMPTOTIC PROPERTIES OF THE SAMPLE MEAN IN ADAPTIVE STRATIFIED SEQUENTIAL SAMPLING WITH MULTIPLE SELECTION CRITERIA

A. PARVARDEH*, B. PANAHEBHAGH, M. SALEHI M.,
J. BROWN AND D. R. SMITH

Communicated by Hamid Pezeshk

ABSTRACT. We extend the method of adaptive two-stage sequential sampling to include designs where there is more than one criteria used in deciding on the allocation of additional sampling effort. These criteria, or conditions, can be a measure of the target population, or a measure of some related population. We develop Murthy estimator for the design that is unbiased estimators for the population mean, and propose another, more efficient, estimator. We investigate asymptotic properties of this estimator. We use a simulation study to investigate design properties of the multi-criteria adaptive stratified sequential sampling scheme and also some estimator properties under the design.

1. Introduction

Asymptotic normality of estimators in finite populations has been the focus of attention for many years for researchers in sampling and probability. Some examples of work over the last 50 years on asymptotic normality include research on estimators for the sample mean for simple random sampling without replacement (SRSWOR) (Erdos and Renyi [8] and Hajek [12]), the Horvitz-Thomson estimator (HTE) for sampling

MSC(2010): Primary: 62D05; Secondary: 62D99.

Keywords: Stratified sampling, adaptive two-stage sequential sampling, asymptotic properties in finite population, consistency, rare population.

Received: 4 February 2012, Accepted: 13 May 2012

*Corresponding author

© 2013 Iranian Mathematical Society.

without replacement and with different probabilities design (Rosen [20]), ratio and the regression estimators for SRSWOR (Scott and Wu [22]) and two-stage sampling estimators (Ohlsson [19]). Berger [1] analyzed the rate of convergence to the normal distribution for the HTE under any unequal probability sampling without replacement design. Brewer [2], Sarndal [21], Isaki and Fuller [15] have shown the consistency of some regression type estimators under different sampling design with unequal probabilities, and more recently, Martin [17] has investigated asymptotic normality in adaptive cluster sampling.

Despite this research activity, asymptotic normality results for commonly used estimators are still not available for many sampling designs. In this article we consider adaptive sequential sampling where the decision to adapt the sample is based on more than one decision criteria. We develop Murthy estimator for these designs and explore its asymptotic properties.

The commonly used estimator in adaptive sequential designs (e.g., adaptive two-stage sequential sampling, Brown et al. [3]) is the unbiased Murthy estimator. In section 2, we derive Murthy estimator for multiple criteria designs, and, based on a derived form of it, propose a new estimator. This new form of the estimator was first introduced by Moradi and Salehi [18] for single criterion adaptive stratified sequential sampling. Moradi and Salehi [18] have shown that this new form is asymptotically unbiased and can be more efficient than Murthy estimator. In section 3, we investigate the asymptotic properties of the new estimator, drawing on the method of Chen and Rao [5]. We prove this new estimator is consistent for the population mean and is asymptotically normal under Multi-criteria Adaptive Stratified Sequential (MAS) sampling. In section 4, we conclude with two case studies to illustrate the use of multiple criteria for adaptive sequential sampling. We discuss the efficiency of these designs and of our new estimator and briefly the asymptotic normality of the estimator under the design.

2. Adaptive two-stage sequential sampling with multiple criteria

Two-stage sampling designs (Fattorini and Pisani [9]) are popular choices for sampling surveys. In conventional two-stage sampling, the population is divided into Primary Sampling Units (PSU) from which a sample is drawn. Within selected PSUs a sample of Secondary Sampling

Units (SSU) is drawn. PSUs are typically large units, such as geographic units while SSUs are smaller, such as plots, quadrats, or households.

Adaptive Two-Stage Sequential (ATS) sampling was initially proposed by Brown et al. [3] as a sample design for sampling rare and clustered populations. Allocation of second-stage effort among PSUs is based on a preliminary information from the sampled PSUs. Additional survey effort is directed to those PSUs where the SSUs in the initial sample have met a pre-specified criterion, or condition (e.g., an individual from the rare population is present). This design effectively over-samples PSUs with high values, compared with other PSUs, a method consistent with the approach recommended by Kalton and Anderson [16] for sampling rare populations.

We begin by introducing the notation of Brown et al. [3]. Suppose we have a total population of N units partitioned into H PSUs of size N_h units. Let $\{(h, i), h = 1, 2, \dots, H, i = 1, 2, \dots, N_h\}$ denote the i^{th} unit in the h^{th} primary unit with an associated measurement or count y_{hi} . Then, $\bar{Y}_{N_h} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$ is the mean of the y -value for the h^{th} PSU and $\bar{Y}_N = \frac{1}{N} \sum_{h=1}^H N_h \bar{Y}_{N_h}$ is the population mean.

According to Brown et al. [3], a simple random sample of size m is taken without replacement, say s , in the first stage and in the second stage an initial simple random sample of n_{h1} units without replacement is taken from PSU h in s . The total initial sample size is $n_1 = \sum_{h \in s} n_{h1}$. Let C be the condition that, if satisfied for l_h units in the initial sample set from PSU h , results in $l_h \times d$ number of additional units being selected at random from the remaining units in PSU h , where d is a predetermined value. The number of adaptively added units in the h^{th} PSU is $n_{h2} = l_h d$ and $n_2 = \sum_{h \in s} n_{h2}$ is the number of adaptively added units in the final sample.

We now extend the design to more than one condition. We begin considering a design with just one PSU. For simplicity assume that we have just one PSU of size N with measurement $y_i, i = 1, 2, \dots, N$. We begin with two mutually exclusive conditions, C_1 and C_2 such that we have three different types of $y_i, y_i \in U_{C_1}, y_i \in U_{C_2}, y_i \in U_N - (U_{C_1} \cup U_{C_2}) = U_{C_3}$, where $U_N = \{y_i, i = 1, 2, \dots, N\}$ and U_{C_k} is the set of y_i that satisfies C_k . With this notation, $U_{C_k}, k = 1, 2, 3$ is a partition for U_N . Note that each y_i can satisfy only one of the conditions. Assume we take a sample of size n_1 , denoted by s_1 , in the initial phase, and let l_{11}, l_{12} denote the number

of units in the initial sample that satisfy respectively C_1, C_2 . We then take a sample of size $n_2 = f(l_{11}, l_{12})$ where the number of adaptively added units, n_2 , is a function of l_{11} and l_{12} , and can be constructed according to the relative importance of C_1 and C_2 . We summaries the design in Table 1.

In the next subsection, we drive an unbiased Murthy estimator for two

TABLE 1. Summary of two conditions situations, where $l_1 = l_{11} + l_{12}, l_2 = l_{21} + l_{22}, l = l_{c_1} + l_{c_2} = l_1 + l_2$.

number	condition	Initial phase s_1	Second phase s_2	$s = s_1 \cup s_2$
N		n_1	$n_2 = n(l_{11}, l_{12})$	$n = n_1 + n_2$
N_{C_1}	C_1	l_{11}	l_{21}	$l_{c_1} = l_{11} + l_{21}$
N_{C_2}	C_2	l_{12}	l_{22}	$l_{c_2} = l_{12} + l_{22}$
$N_{C_3} = N - N_{C_1} - N_{C_2}$	$C_3 = (C_1 \cup C_2)'$	$n_1 - l_1$	$n_2 - l_2$	$n - l$

conditions in ATS design.

2.1. Murthy estimator for ATS. To estimate the mean of the PSU, we can use

$$\widehat{Y} = \sum_{i \in s} \frac{P(s | I_i = 1)}{NP(s)} y_i$$

where $P(s)$ and $P(s | I_i = 1)$ are the probability of getting the sample s and the conditional probability of getting the sample s given the i^{th} unit was selected in the first draw, s is the final sample and I_i is an indicator function which takes the value 1 when unit i is chosen as the first selected unit, and 0 otherwise. Now with

$$n(s) = \binom{l_{c_1}}{l_{11}} \binom{l_{c_2}}{l_{12}} \binom{n - l_{c_1} - l_{c_2}}{n_1 - l_{11} - l_{12}} n_1! n_2!$$

and

$$n(s, I_i = 1) = \begin{cases} \binom{l_{c_1} - 1}{l_{11} - 1} \binom{l_{c_2}}{l_{12}} \binom{n - l_{c_1} - l_{c_2}}{n_1 - l_{11} - l_{12}} (n_1 - 1)! n_2!; & \text{if } y_i \in U_{C_1} \\ \binom{l_{c_1}}{l_{11}} \binom{l_{c_2} - 1}{l_{12} - 1} \binom{n - l_{c_1} - l_{c_2}}{n_1 - l_{11} - l_{12}} (n_1 - 1)! n_2!; & \text{if } y_i \in U_{C_2} \\ \binom{l_{c_1}}{l_{11}} \binom{l_{c_2}}{l_{12}} \binom{n - l_{c_1} - l_{c_2} - 1}{n_1 - l_{11} - l_{12} - 1} (n_1 - 1)! n_2!; & \text{if } y_i \in U_{C_3} \end{cases}$$

where $n(s)$ and $n(s, I_i = 1)$ are the number of permutations giving rise to s and the number of permutations giving rise to $s, I_i = 1$, because $P(I_i = 1) = \frac{1}{N}$, we have

$$\frac{P(s | I_i = 1)}{NP(s)} = \frac{n(s, I_i = 1)}{Nn(s)P(I_i = 1)} = \begin{cases} \frac{l_{11}}{n_1 l_{c1}}; & \text{if } y_i \in U_{C_1} \\ \frac{l_{12}}{n_1 l_{c2}}; & \text{if } y_i \in U_{C_2} \\ \frac{n_1 - l_{11}}{n_1(n-l)}; & \text{if } y_i \in U_{C_3} \end{cases}$$

Therefore

$$\begin{aligned} \widehat{Y} &= \frac{l_{11}}{n_1} \bar{y}_{s_{c_1}} + \frac{l_{12}}{n_1} \bar{y}_{s_{c_2}} + \frac{l_{13}}{n_1} \bar{y}_{s_{c_3}} \\ &= \widehat{p}_{11} \bar{y}_{s_{c_1}} + \widehat{p}_{12} \bar{y}_{s_{c_2}} + (1 - \widehat{p}_{11} - \widehat{p}_{12}) \bar{y}_{s_{c_3}} \\ &= \widehat{p}_{11} \bar{y}_{s_{c_1}} + \widehat{p}_{12} \bar{y}_{s_{c_2}} + \widehat{p}_{13} \bar{y}_{s_{c_3}} \end{aligned}$$

where s_{c_k} is the total sample that are satisfied in condition C_k and $\bar{y}_{s_{c_k}} = \frac{1}{l_{c_k}} \sum_{i \in s_{c_k}} y_i$.

We introduce another estimator by changing the form of the estimator to allow us to investigate the asymptotic properties. This new estimator not only has desired asymptotic properties but also can be more efficient than Murthy estimator for ATS. In the next subsection we introduce this new estimator.

2.2. A new form of Murthy estimator for ATS. In an ATS design we have,

$$\begin{aligned} l_{11}, l_{12} &\sim HG(N, N_{C_1}, N_{C_2}; n_1), \\ l_{21}, l_{22} &| \quad l_{11}, l_{12} \sim HG(N - n_1, N_{C_1} - l_{11}, N_{C_2} - l_{12}; n_2). \end{aligned}$$

where "HG" denotes the multivariate hypergeometric distribution. As shown above, Murthy estimator, \widehat{Y} , is a weighted estimator of the mean of y_i 's from the three conditions, where the weights are a vector of the unbiased estimators ($\frac{l_{11}}{n_1} = \widehat{p}_{11}$, $\frac{l_{12}}{n_1} = \widehat{p}_{12}$, $\frac{l_{13}}{n_1} = \widehat{p}_{13}$) of the three subcomputations portions ($\frac{N_{C_1}}{N} = p_1$, $\frac{N_{C_2}}{N} = p_2$, $\frac{N_{C_3}}{N} = p_3$). This vector is constructed using just the information from the first phase of sampling. We propose a new estimator where all the information, from both phases in the design, is used to construct these weights. The new estimator

constructed using information from both phases is

$$\begin{aligned}\bar{y}_s &= \frac{l_{c_1}}{n} \bar{y}_{s_{c_1}} + \frac{l_{c_2}}{n} \bar{y}_{s_{c_2}} + \frac{n - l_{c_1} - l_{c_2}}{n} \bar{y}_{s_{c_3}} \\ &= \frac{1}{n} \left(\sum_{i \in s_{c_1}} y_i + \sum_{i \in s_{c_2}} y_i + \sum_{i \in s_{c_3}} y_i \right) = \frac{1}{n} \left(\sum_{i \in s} y_i \right).\end{aligned}$$

Although this estimator is not an unbiased estimator for the population mean, Moradi and Salehi [18] have shown that this estimator under ATS is asymptotically unbiased for one condition. Furthermore, we show that this estimator is consistent for the population mean and it is asymptotically normal.

3. Asymptotic properties of the sample mean under MAS design

To investigate the asymptotic properties of the estimator we extend our application to adaptive stratified sampling. First we define a new form of the population to address the two conditions (i.e., there are three kinds of individuals). Hereafter, we add the index v to every population or sample size. Setting this index to infinity means that all the populations and sample sizes tend to infinity under some condition.

Let $\{\mathbf{z}_i\}$ and $\{\mathbf{z}_i^*\}$ be sequences of real numbers, with $\mathbf{z}_i = (y_i, \mathbf{x}_i) = (y_i, x_{1i}, x_{2i}, x_{3i})$ and $\mathbf{z}_i^* = (z_{1i}^*, z_{2i}^*, z_{3i}^*) = (y_i x_{1i}, y_i x_{2i}, y_i x_{3i})$, where y_i is the value corresponding to the i^{th} individual and x_{ki} , $k = 1, 2, 3$ are indicator functions that take 1 if y_i satisfies condition k and 0 otherwise. We define a sequence of finite populations $\{\mathcal{F}_{N^v}\}$, where the N^{th} finite population is composed of the first N values of the sequence $\{\mathbf{z}_i\}$.

Assume that

$$\lim_{v \rightarrow \infty} \left\{ \frac{1}{N^v} \sum_{i \in U_{N^v}} (x_{ki} y_i, x_{ki} y_i x_{mi} y_i) \mid \mathcal{F}_{N^v} \right\} = (\theta_k, \theta_{km}); \quad m \leq k = 1, 2, 3,$$

where (θ_k, θ_{km}) , $m \leq k = 1, 2, 3$ are finite and we assume that (θ_k, θ_{km}) are satisfied in all conditions that are necessary (for example, $\theta_{kk} - \theta_k^2 > 0$). We call this assumption as a moment assumption. Hereafter, when we write " $i \in U_{N^v}$ " it means every index " i " that " $y_i \in U_{N^v}$ ".

We then have

$$\bar{\mathbf{z}}_{N^v}^* = \frac{1}{N^v} \sum_{i \in U_{N^v}} \mathbf{z}_i^* = \begin{pmatrix} \bar{Y}_{x_1 N^v} \\ \bar{Y}_{x_2 N^v} \\ \bar{Y}_{x_3 N^v} \end{pmatrix}',$$

where

$$\bar{Y}_{x_k N^v} = \frac{1}{N^v} \sum_{i \in U_{N^v}} y_i x_{ik} = \frac{N_{C_k}^v}{N^v} \frac{1}{N_{C_k}^v} \sum_{i \in U_{C_k}} y_i = \frac{N_{C_k}^v}{N^v} \bar{Y}_{N_{C_k}^v}.$$

Also we have

$$\begin{aligned} \mathbf{S}_{\mathbf{z}_{N^v}^*}^2 &= \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (\mathbf{z}_i^* - \bar{\mathbf{z}}_{N^v}^*)' (\mathbf{z}_i^* - \bar{\mathbf{z}}_{N^v}^*) \\ &= \frac{1}{N^v - 1} \begin{pmatrix} S_{yx_1 N^v}^2 & S_{yx_1 x_2 N^v} & S_{yx_1 x_3 N^v} \\ S_{yx_1 x_2 N^v} & S_{yx_2 N^v}^2 & S_{yx_2 x_3 N^v} \\ S_{yx_1 x_3 N^v} & S_{yx_2 x_3 N^v} & S_{yx_3 N^v}^2 \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} S_{yx_k N^v}^2 &= \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (y_i x_{ik} - \bar{Y}_{x_k N^v})^2 \\ S_{yx_k x_h N^v} &= \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (y_i x_{ik} - \bar{Y}_{x_k N^v})(y_i x_{ih} - \bar{Y}_{x_h N^v}). \end{aligned}$$

According to the moment assumption, the population covariance $\mathbf{S}_{\mathbf{z}_{N^v}^*}^2$ approaches to a positive definite constant matrix $\mathbf{\Gamma}$ as $v \rightarrow \infty$, with

$$\mathbf{\Gamma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix},$$

where $\sigma_k^2 = \theta_{kk} - \theta_k^2$ and $\sigma_{km} = \theta_{km} - \theta_k \theta_m$, $m \leq k = 1, 2, 3$.

Now note that we can decompose $\bar{y}_s - \bar{Y}_{N^v}$ as below:

$$\begin{aligned} (3.1) \quad \bar{y}_s - \bar{Y}_{N^v} &= \frac{n_1^v}{n^v} (\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + \frac{n_1^v}{n^v} (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v}) \\ &+ \frac{n_1^v}{n^v} (\bar{y}_{x_3 s_1} - \bar{Y}_{x_3 N^v}) + \frac{n_2^v}{n^v} (\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v}) \\ &+ \frac{n_2^v}{n^v} (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v}) + \frac{n_2^v}{n^v} (\bar{y}_{x_3 s_2} - \bar{Y}_{x_3 N^v}). \end{aligned}$$

And, because $\bar{y}_s - \bar{Y}_{N^v}$ is decomposed to the first and second phase of sampling, the following theorem will be helpful before the main theorem.

Theorem 3.1. (Chen and Rao [5]) Let W_{n^v}, V_{n^v} be two sequences of random variables and let \mathbb{B}_{n^v} be a σ -algebra. Assume that

(1) there exists $\sigma_{1n^v} > 0$ such that

$$\sigma_{1n^v}^{-1} V_{n^v} \xrightarrow{d} N(0, 1)$$

as $n^v \rightarrow \infty$, and V_{n^v} is \mathbb{B}_{n^v} measurable,

(2) $E\{W_{n^v} \mid \mathbb{B}_{n^v}\} = 0$ and $\text{Var}(W_{n^v} \mid \mathbb{B}_{n^v}) = \sigma_{2n^v}^2$ such that

$$\sup_t |P(\sigma_{2n^v}^{-1} W_{n^v} \leq t \mid \mathbb{B}_{n^v}) - \Phi(t)| = o_p(1),$$

where $\Phi(t)$ is the cumulative distribution function of the standard normal random variable,

(3) $\gamma_{n^v}^2 = \frac{\sigma_{1n^v}^2}{\sigma_{2n^v}^2} \rightarrow \gamma^2$ in probability as $n^v \rightarrow \infty$,

then as $n^v \rightarrow \infty$

$$\frac{W_{n^v} + V_{n^v}}{\sqrt{\sigma_{1n^v}^2 + \sigma_{2n^v}^2}} \xrightarrow{d} N(0, 1).$$

Now, with this background we are ready to formulate a theorem about asymptotic normality of $\sqrt{n^v}(\bar{Y}_s - \bar{Y}_{N^v})$ for just one condition with $n_2^v = d \times l_1 + 1$. We begin with one condition with completely specified situation.

Theorem 3.2. *In adaptive sequential sampling, let $\mathbb{B}_{n_1^v}$, the σ -algebra generated by s_1 , contains all the information from the first phase of sampling, and*

- C1) $N^v, N_{C_k}^v, n_1^v, n_2^v \mid \mathbb{B}_{n_1^v} \rightarrow \infty$ as $v \rightarrow \infty$, and $N_{C_k}^v < N^v, n_1^v < N^v, n_2^v < N^v - n_1^v$,
- C2a) $(\frac{N_{C_k}^v}{N^v} \mid \mathcal{F}_{N^v})$ is a fixed sequence that converges to f_k as $v \rightarrow \infty$,
- C2b) $(\frac{n_1^v}{N^v} \mid \mathcal{F}_{N^v})$ is a fixed sequence that converges to f_1^* as $v \rightarrow \infty$,
- C2c) $(\frac{n_2^v}{N^v - n_1^v} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v})$ is a fixed sequence that converges to f_2^* as $v \rightarrow \infty$ and
- C3) for all $\epsilon > 0$

$$\lim_{v \rightarrow \infty} \left\{ \frac{1}{N^v} \sum_{\{i \in U_{N^v}\} \cap \{i: \|\tilde{\mathbf{z}}_i^*\| > \epsilon N^v \frac{n_1^v}{N^v} (1 - \frac{n_1^v}{N^v})\}} \|\tilde{\mathbf{z}}_i^*\|^2 \mid \mathcal{F}_{N^v} \right\} = 0,$$

where $\tilde{\mathbf{z}}_i^* = \mathbf{z}_i^* - \bar{\mathbf{z}}_{N^v}^*$, and $\|\cdot\|$ is the Euclidean norm,

- C4a) $[(\bar{Y}_{x_1 N^v - s_1} - \bar{Y}_{x_1 N^v}) | \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}] = O(n_2^{v - (\frac{1}{2} + \delta)})$
and $[(\bar{Y}_{x_2 N^v - s_1} - \bar{Y}_{x_2 N^v}) | \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}] = O(n_2^{v - (\frac{1}{2} + \delta)})$ for some $\delta > 0$, where $N^v - s_1$ means whole population after excluding s_1 ,
- C4b) $\left[\begin{pmatrix} S_{y x_1 N^v - s_1}^2 & S_{y x_1 x_2 N^v - s_1} \\ S_{y x_1 x_2 N^v - s_1} & S_{y x_2 N^v - s_1}^2 \end{pmatrix} - \begin{pmatrix} S_{y x_1 N^v}^2 & S_{y x_1 x_2 N^v} \\ S_{y x_1 x_2 N^v} & S_{y x_2 N^v}^2 \end{pmatrix} \right] | \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \rightarrow 0$ as $v \rightarrow \infty$,
- C4c) for all $\epsilon > 0$

$$\lim_{v \rightarrow \infty} \left\{ \frac{1}{N^v - n_1^v} \sum_{\{i \in U_{N^v - s_1}\} \cap \{i: \|\tilde{\mathbf{z}}_i^*\| > \epsilon(N^v - n_1^v) \frac{n_2^v}{N^v - n_1^v} (1 - \frac{n_2^v}{N^v - n_1^v})\}} \|\tilde{\mathbf{z}}_i^*\|^2 \right\} | \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \} = 0.$$

Then as $v \rightarrow \infty$

$$\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v}) \xrightarrow{d} N(0, V_\infty),$$

where

$$V_\infty = \left(1 - \frac{f_1^* + df_1 f_2^*}{1 + df_1}\right)^{\frac{1}{2}} \sqrt{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})}.$$

and furthermore,

$$(\bar{y}_s - \bar{Y}_{N^v} | \mathcal{F}_{N^v}) = o_p(1)$$

The proof of Theorem 3.2 is in Appendix A.

In adaptive sequential sampling, n is a random variable. The exact variance of the \bar{y}_s is difficult to find, but the asymptotic form of the variance to standardize the estimator can be presented in an easy and simple form. There are some unknown parameters in the estimator. The next corollary helps us to use the available information in the sample.

Corollary 3.3. *Under the conditions of Theorem 3.2, we have*

$$\frac{\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})}{\hat{\sigma}} \xrightarrow{d} N(0, 1),$$

where

$$\hat{\sigma} = \left(1 - \frac{\hat{f}_1^* + d\hat{f}_1 \hat{f}_2^*}{1 + d\hat{f}_1}\right)^{\frac{1}{2}} \sqrt{(s_1^2 + s_2^2 + 2s_{12})}$$

and

$$\begin{aligned}\widehat{f}_1 &= \frac{l_1}{n_1^v}, \widehat{f}_1^* = \frac{n_1^v}{N^v}, \widehat{f}_2^* = \frac{n_2^v}{N^v - n_1^v}, \\ s_k^2 &= \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})^2, k = 1, 2, \\ s_{12} &= \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{1i} - \bar{Y}_{x_1 n_1^v})(y_i x_{2i} - \bar{Y}_{x_2 n_1^v}), \bar{Y}_{x_k n_1^v} = \frac{1}{n_1^v} \sum_{i \in s_1} y_i x_{ki}.\end{aligned}$$

Proof. Using the conditions in Theorem 3.2, it is easy to show that:

$$\begin{aligned}\left(1 - \frac{\widehat{f}_1^* + d\widehat{f}_1 \widehat{f}_2^*}{1 + d\widehat{f}_1}\right)^{\frac{1}{2}} \sqrt{(s_1^2 + s_2^2 + 2s_{12})} \mid \mathcal{F}_{N^v} \\ \xrightarrow{p} \left(1 - \frac{f_1^* + df_1 f_2^*}{1 + df_1}\right)^{\frac{1}{2}} \sqrt{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})}.\end{aligned}$$

Implementing Slutsky Theorem (see Ferguson [10], chapter 6) the proof is completed. \square

According to results from Theorem 3.2, we can easily extend a theorem for more than one condition. Below we represent a theorem for two conditions with a general sample size $n_2^v = f(l_{11}, l_{12})$. For more than two conditions the method will be the same.

According to (3.1) we can decompose $\sqrt{n^v}(\bar{Y}_s - \bar{Y}_{N^v})$ as below:

$$\begin{aligned}\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v}) &= \sqrt{n_1^v} \left[\sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + \sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v}) \right. \\ &\quad \left. + \sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_3 s_1} - \bar{Y}_{x_3 N^v}) \right] + \sqrt{n_2^v} \left[\sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v}) \right. \\ &\quad \left. + \sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v}) + \sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_3 s_2} - \bar{Y}_{x_3 N^v}) \right]\end{aligned}$$

We now have the following theorem.

Theorem 3.4. *In adaptive sequential sampling, let $\mathbb{B}_{n_1^v}$, the σ -algebra generated by s_1 , contains all the information in first phase of sampling, and*

- 1) $N^v, N_{C_k}^v, n_1^v, n_2^v \mid \mathbb{B}_{n_1^v} \rightarrow \infty$ as $v \rightarrow \infty$, and $N_{C_k}^v < N^v, n_1^v < N^v, n_2^v < N^v - n_1^v$,
- 2a) $(\frac{n_1^v}{N^v} \mid \mathcal{F}_{N^v})$ is a fixed sequence that converge to f_1^* as $v \rightarrow \infty$,

- 2b) $(\frac{n_2^v}{N^v - n_1^v} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v})$ is a fixed sequence that converge to f_2^* as $v \rightarrow \infty$,
- 2c) $(\frac{n_1^v}{n^v} - a \mid \mathcal{F}_{N^v}) \xrightarrow{p} 0$,
- 3) for all $\epsilon > 0$

$$\lim_{v \rightarrow \infty} \left\{ \frac{1}{N^v} \sum_{\{i \in U_{N^v}\} \cap \{i: \|\tilde{\mathbf{z}}_i^*\| > \epsilon N^v \frac{n_1^v}{N^v} (1 - \frac{n_1^v}{N^v})\}} \|\tilde{\mathbf{z}}_i^*\|^2 \mid \mathcal{F}_{N^v} \right\} = 0,$$

where $\tilde{\mathbf{z}}_i^* = \mathbf{z}_i^* - \bar{\mathbf{z}}_{N^v}^*$, and $\|\cdot\|$ is the Euclidean norm,

- 4a) $[(\bar{Y}_{x_k N^v - s_1} - \bar{Y}_{x_k N^v}) \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}] = O(n_2^{v - (\frac{1}{2} + \delta)})$, for $k = 1, 2, 3$ and some $\delta > 0$ where $N^v - s_1$ means the whole population after excluding s_1 ,
- 4b) $\left[\begin{pmatrix} S_{yx_1 N^v - s_1}^2 & S_{yx_1 x_2 N^v - s_1} & S_{yx_1 x_3 N^v - s_1} \\ S_{yx_1 x_2 N^v - s_1} & S_{yx_2 N^v - s_1}^2 & S_{yx_2 x_3 N^v - s_1} \\ S_{yx_1 x_3 N^v - s_1} & S_{yx_2 x_3 N^v - s_1} & S_{yx_3 N^v - s_1}^2 \end{pmatrix} - \begin{pmatrix} S_{yx_1 N^v}^2 & S_{yx_1 x_2 N^v} & S_{yx_1 x_3 N^v} \\ S_{yx_1 x_2 N^v} & S_{yx_2 N^v}^2 & S_{yx_2 x_3 N^v} \\ S_{yx_1 x_3 N^v} & S_{yx_2 x_3 N^v} & S_{yx_3 N^v}^2 \end{pmatrix} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \right] \rightarrow 0$ as $v \rightarrow \infty$,
- 4c) for all $\epsilon > 0$

$$\lim_{v \rightarrow \infty} \left\{ \frac{1}{N^v - n_1^v} \sum_{\{i \in U_{N^v - s_1}\} \cap \{i: \|\tilde{\mathbf{z}}_i^*\| > \epsilon (N^v - n_1^v) \frac{n_2^v}{N^v - n_1^v} (1 - \frac{n_2^v}{N^v - n_1^v})\}} \|\tilde{\mathbf{z}}_i^*\|^2 \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \right\} = 0.$$

Then as $v \rightarrow \infty$

$$\frac{\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})}{\hat{\sigma}} \xrightarrow{d} N(0, 1), \quad (\bar{y}_s - \bar{Y}_{N^v} \mid \mathcal{F}_{N^v}) = o_p(1)$$

where

$$\hat{\sigma} = [\hat{a}(1 - \hat{f}_1^*) + (1 - \hat{a})(1 - \hat{f}_2^*)]^{\frac{1}{2}} \sqrt{s_1^2 + s_2^2 + s_3^2 + 2s_{12} + 2s_{13} + 2s_{23}},$$

and

$$\begin{aligned}\hat{a} &= \frac{n_1^v}{n^v}, \hat{f}_1^* = \frac{n_1^v}{N^v}, \hat{f}_2^* = \frac{n_2^v}{N^v - n_1^v}, \\ s_k^2 &= \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})^2, k = 1, 2, 3 \\ s_{km} &= \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})(y_i x_{mi} - \bar{Y}_{x_m n_1^v}), m < k = 1, 2, 3 \\ \bar{Y}_{x_k n_1^v} &= \frac{1}{n_1^v} \sum_{i \in s_1} y_i x_{ki}, k = 1, 2, 3.\end{aligned}$$

The proof of this follows from the results for one condition.

It is notable that in these theorems, we have established a set of conditions which guarantees the asymptotic normality and consistency for the estimator in the design. However, in practice we have a finite population with fixed size and in a real sampling situation, it will generally be impossible to verify this set of conditions. Therefore, the set of conditions should be used only as a general guideline. For example, here the conditions in the theorems suggest that in situations where the size of population and sample are large and also there are not so many outlying data in the population, the estimator will be normally distributed.

3.1. Design estimator. We now extend these results to MAS sampling. It is necessary to add an index h to all above result and then construct the final estimator. From the previous results (Theorem 3.4) we have, for $h = 1, 2, \dots, H$

$$\sqrt{n_h^v} \frac{(\bar{y}_{sh} - \bar{Y}_{N_h^v})}{\sigma_h} \mid \mathcal{F}_{N_h^v} \xrightarrow{d} N(0, 1),$$

where

$$\sigma_h^2 = [a_h(1 - f_{1h}^*) + (1 - a_h)(1 - f_{2h}^*)](\sigma_{1h}^2 + \sigma_{2h}^2 + \sigma_{3h}^2 + 2\sigma_{12h} + 2\sigma_{13h} + 2\sigma_{23h}).$$

We use stratified sampling setup where a PSU is considered a stratum, to construct conventional estimator, i.e.,

$$\bar{y}_s = \sum_{h=1}^H W_h^v \bar{y}_{sh},$$

where

$$W_h^v = \frac{N_h^v}{N^v},$$

and

$$\bar{y}_s - \bar{Y}_{N^v} = \sum_{h=1}^H W_h^v (\bar{y}_{sh} - \bar{Y}_{N_h^v}).$$

According to previous section we can re-state the final theorem as below:

Theorem 3.5. *In adaptive stratified sequential sampling, under the conditions in Theorem 3.4, with fixed H , if*

$$(1) \left(\frac{n_h^v}{n^v} - f_h^{**} \mid \mathcal{F}_{N^v} \right) \xrightarrow{p} 0, \quad h = 1, 2, \dots, H \text{ as } v \rightarrow \infty, \text{ where } n^v = \sum_{h=1}^H n_h^v \text{ and } n_h^v = n_{1h}^v + n_{2h}^v,$$

$$(2) (W_h^v \mid \mathcal{F}_{N^v}) \text{ is a fixed sequence that converge to } F_h; h = 1, 2, \dots, H, \text{ as } v \rightarrow \infty.$$

then

$$\sqrt{n^v} \frac{(\bar{y}_s - \bar{Y}_{N^v})}{\hat{\sigma}} \xrightarrow{d} N(0, 1), \quad \bar{y}_s - \bar{Y}_{N^v} \mid \mathcal{F}_{N^v} = o_p(1)$$

where

$$\hat{\sigma}^2 = \sum_{h=1}^H \frac{\hat{F}_h^2}{\hat{f}_h^{**}} \hat{\sigma}_h^2,$$

$$\hat{\sigma}_h = [\hat{a}_h(1 - \hat{f}_{1h}^*) + (1 - \hat{a}_h)(1 - \hat{f}_{2h}^*)]^{1/2} \sqrt{\sum_{k=1}^3 s_{kh}^2 + 2 \sum_{k=1}^3 \sum_{k' > k} s_{kk'h}},$$

$$\hat{f}_{1h}^* = \frac{n_{1h}^v}{N_h^v}, \hat{f}_{2h}^* = \frac{n_{2h}^v}{N_h^v - n_{1h}^v}, \hat{a}_h = \frac{n_{1h}^v}{n_h^v}, \hat{F}_h = W_h^v, \hat{f}_h^{**} = \frac{n_h^v}{n^v}$$

and $s_{kh}^2, k = 1, 2, 3$ and $s_{kmh}, m < k = 1, 2, 3$ are defined as before (in Theorem 3.4) but in the h^{th} PSU.

See Appendix C for the proof of Theorem 3.5.

The results of the above theorems give us some information about the distribution of the estimator. This allows us to use the standard inference methods like large sample and population size, and calculate a $(1 - \alpha)100\%$ confidence interval using normal distribution as below

$$\left(\bar{y}_s - z_{\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}}, \bar{y}_s + z_{\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} \right)$$

where $z_{\frac{\alpha}{2}}$ is the right tail of the normal distribution containing $\frac{\alpha}{2}$ probability ($0 < \alpha < 1$).

4. Case Study

In this section, we illustrate the use of multiple criteria adaptive sampling and we investigate asymptotic normality of the estimator under MAS design. We use an artificial population of mussels and the standard poisson cluster process population. We investigate sample efficiency and, in the second part, we briefly investigate normality of the estimator under the new design.

4.1. Multiple criteria adaptive sampling. In this part we try to evaluate multiple criteria adaptive sampling design using two data series.

4.1.1. Mussels populations. One of the standard data sets used in adaptive sampling research is a freshwater mussel population from the Cacapon River, West Virginia (Smith et al. [23]). The mussels in the population occur in clusters and at low density, what we refer to as a rare and clustered population.

In ecological systems there are usually multiple species present along with the species of interest. In this simulation study we have a target mussel population and treat the other species as auxiliary populations. We generated artificial populations of mussels and of some other species. The counts of the mussels were correlated with the auxiliary species' counts. The auxiliary species could represent a different mussel species, or a population of another animal or plant.

Each sample unit, or quadrat, contains two variables, y_{hi} , the count of the target mussels in the i^{th} quadrat in the h^{th} PSU and x_{hi} , the count of the other species in the respective quadrat. For the multiple criteria sampling, the first condition is based on the target mussel count. The second condition is based on the auxiliary variable.

The data set was divided into 400 quadrats. The quadrats are partitioned evenly into 4 PSU. We simulated three generated mussel populations whose parameters were set to be similar to real life studies by Hornbach et al. [13, 14]. They could span a realistic range, with mean count/quadrat of 0.785, 0.500 and 0.240 and variance of 3.973, 2.265 and 1.170 respectively for target variable. We call these three synthetic populations "popu1, popu2" and "popu3".

We used one auxiliary population, with mean count of 0.675 and variance of 5.212 (figure 1). The correlation between the target and the auxiliary populations were 0.312, 0.375 and 0.322, respectively, for the three simulations. We compared the sample design with one condition

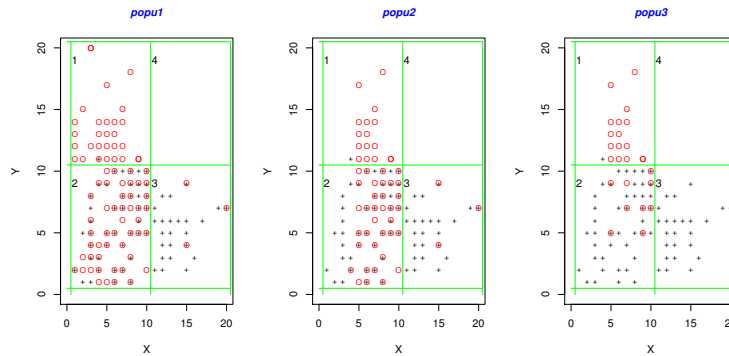


FIGURE 1. Three artificial mussel populations. The red circles and black pluses indicate where counts of the target mussel and auxiliary population were one, or more

and the two conditions design. For each design we calculated the proportion of proportions (PP) of the design. We define this as the proportion of the quadrats occupied by the target species that were detected in the design, divided by proportion of the quadrats occupied by the target species in the whole population. Also we calculated the expectation and variance of \bar{y}_s and Murthy estimator for the two designs and using them to compute the relative efficiencies of the designs. We calculated relative efficiency as the ratio of the simulated variances. We used a range of sample design parameters:

- (1) ATS_{2c} , ATS with two conditions, where both the target and auxiliary populations were used in the condition definitions. In this design $n_1 = 3, 6, 9, 12, n_2 = d_1 * l_{11} + d_2 * l_{12} + 1$ with $d_1 = 6, d_2 = 3$ (for popu3 we set $d_1 = 7$). We define C_1 as the condition when the quadrat has at least one mussel (target population) in it ($y_{hi} > 0$), and C_2 as the condition when the quadrat does not satisfy C_1 and the quadrat has at least 1 individual from the auxiliary population (y_{hi} does not satisfy C_1 and $x_{hi} > 0$). Therefore, C_3 is the condition when neither condition 1 nor 2 are met, the complement of $C_1 \cup C_2$.
- (2) ATS_{1c} , ATS with one condition. In this design $n_1 = 5, n_2 = d * l_1 + 1$ with $d = 5$. We define C as the condition when the quadrat has at least one individual in it from either the target or auxiliary population.

The values of n_1 , d , d_1 and d_2 were chosen such that the final sample sizes in both design were the same to have a fair comparison between the two.

It is notable that according to the plots of the populations, the target variable can help us more than the auxiliary variable to find more target variable and that's why we gave more weight to the target variable ($d_1 = 6$) relative to the auxiliary variable ($d_2 = 3$) in ATS_{c2} .

The sample results are shown in tables (2,3 and 4). In these tables when four PSU are selected ($m = 4$), we have adaptive stratified sequential sampling design.

TABLE 2. PP for ATS in 3 artificial mussels populations. The numbers of primary sample units sampled (m) was 1, 2, 3 and 4. Results from the two condition sample are in the column, 2c, and one condition results in column 1c.

	m=4		m=3		m=2		m=1	
	2c	1c	2c	1c	2c	1c	2c	1c
popu1								
$n_1=3$	1.289	1.231	1.211	1.169	1.080	1.070	0.836	0.855
$n_1=6$	1.344	1.273	1.284	1.236	1.164	1.136	0.884	0.893
$n_1=9$	1.365	1.287	1.304	1.245	1.179	1.142	0.917	0.925
$n_1=12$	1.375	1.299	1.321	1.254	1.204	1.160	0.937	0.936
popu2								
$n_1=3$	1.263	1.247	1.213	1.206	1.047	1.059	0.815	0.832
$n_1=6$	1.345	1.306	1.271	1.247	1.145	1.138	0.893	0.911
$n_1=9$	1.357	1.316	1.296	1.267	1.145	1.142	0.928	0.933
$n_1=12$	1.371	1.327	1.286	1.256	1.184	1.162	0.928	0.930
popu3								
$n_1=3$	1.014	0.997	0.992	0.985	0.893	0.920	0.705	0.769
$n_1=6$	1.079	1.036	1.039	1.018	0.944	0.949	0.795	0.836
$n_1=9$	1.095	1.039	1.089	1.034	1.029	1.007	0.853	0.866
$n_1=12$	1.126	1.062	1.094	1.046	1.027	1.002	0.866	0.894

The ATS_{2c} design detected the highest proportion of quadrats occupied by the target species (Table 2), for almost all simulations except for $m = 1$. The two-condition design uses information from both the

TABLE 3. Efficiency of ATS_{2c} relative to ATS_{1c} for \bar{y}_s for 3 artificial mussels population

	m=4	m=3	m=2	m=1
popu1				
$n_1=3$	1.033	1.016	1.045	0.993
$n_1=6$	1.000	1.031	1.019	1.043
$n_1=9$	1.043	1.009	1.000	1.016
$n_1=12$	1.067	1.019	0.996	1.007
popu2				
$n_1=3$	1.000	1.000	1.033	1.033
$n_1=6$	0.960	0.970	1.007	1.025
$n_1=9$	1.000	1.038	1.015	0.987
$n_1=12$	1.000	1.042	1.024	0.992
popu3				
$n_1=3$	0.981	1.003	1.082	1.203
$n_1=6$	0.934	0.975	1.032	1.114
$n_1=9$	0.909	0.907	0.965	1.029
$n_1=12$	0.896	0.924	0.958	1.065

target and auxiliary population separately in allocating second phase effort. With the two conditions, finding either the target mussels or the auxiliary species will initiate second phase effort. In the simulation with $m = 1$ the two condition design detected fewer occupied quadrats which may be a result of the low sample size. The relative improvement of the two condition design over the one-condition design was consistent over the range of densities of three populations.

The two conditions design showed a slight improvement in efficiency compared to the one condition design for populations 1 and 2 (Table 3). There was little, or no, gain in efficiency for population 3, the population with the lowest mussel abundance.

The comparison of new estimator, \bar{y}_s with Murthy estimator for ATS_{2c} (Table 4) showed remarkable gains in efficiency especially at the higher sample sizes ($m = 4$). The new estimator was more than twice as efficient as Murthy estimator for 20 of the 48 simulations, and for all simulations relative efficiencies were greater than 1.

TABLE 4. Efficiency of \bar{y}_s relative to Murthy estimator for ATS_{2c} for 3 artificial mussels population

	m=4	m=3	m=2	m=1
popu1				
$n_1=3$	2.035	1.700	1.619	1.563
$n_1=6$	2.141	1.555	1.370	1.348
$n_1=9$	2.317	1.409	1.279	1.205
$n_1=12$	2.840	1.349	1.221	1.157
popu2				
$n_1=3$	2.322	2.054	1.880	1.820
$n_1=6$	2.323	1.758	1.563	1.481
$n_1=9$	2.320	1.589	1.375	1.281
$n_1=12$	2.544	1.463	1.303	1.208
popu3				
$n_1=3$	2.626	2.666	2.927	2.741
$n_1=6$	2.213	2.194	2.145	2.209
$n_1=9$	2.134	2.160	1.883	1.840
$n_1=12$	2.180	1.938	1.739	1.677

4.1.2. Poisson cluster population. To illustrate another example of multiple criteria adaptive sampling we use the standard Poisson cluster process data (Diggle 1983). This dataset is another example of a clustered population, generated by a linked pairs process (Diggle 1983, pp.93-94) together with a bivariate Poisson cluster process, which is a generalization of Poisson cluster process (Diggle [7], pp.55-57), (Figure 2). The dataset has been used by both Chao [4] and Chutiman [6] to evaluate their results in using auxiliary variable in adaptive cluster sampling design. We partitioned the population into 8 PSUs, each of them of size 50 quadrats. The mean and variance of the target variable are 0.64, and 0.48, respectively. It is not necessary to have auxiliary information for using multiple conditions. Instead the sample design uses information based on the value of the target population count. For ATS_{2c} we set C_1 as " $0 < y_{hi} \leq 2$ " and C_2 as " $3 \leq y_{hi}$ ", with $d_1 = 1$, $d_2 = 8$. For ATS_{1c} we set " $y > 0$ " as C with $d = 1$ and for both of them $n_1 = 3, 5, 7$. Again the parameters were chosen such that the final sample size for both designs were the same.

The results are summarized in Tables 5, 6 and 7.

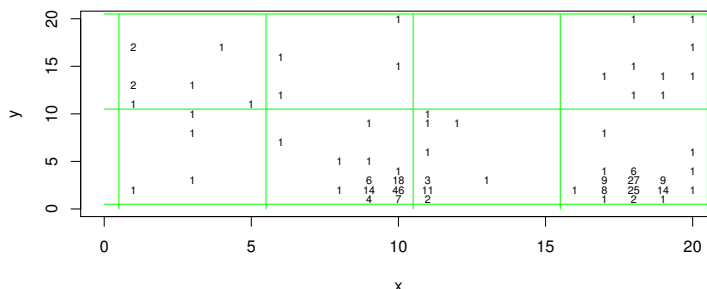


FIGURE 2. The values indicate the counts greater than 0 of the target population in each quadrat from the Poisson cluster process

TABLE 5. PP for ATS for Poisson cluster process population

	m=8		m=6		m=4	
	2c	1c	2c	1c	2c	1c
$n_1=3$	1.081	1.036	1.066	1.022	1.039	1.004
$n_1=5$	1.110	1.063	1.102	1.057	1.067	1.029
$n_1=7$	1.133	1.070	1.118	1.061	1.092	1.0491

TABLE 6. Efficiency of ATS_{2c} relative to ATS_{1c} for \bar{y}_s in the Poisson cluster process population

	m=8	m=6	m=4
$n_1=3$	1.316	1.366	1.299
$n_1=5$	1.302	1.312	1.233
$n_1=7$	1.285	1.265	1.306

As with the mussel population, the ATS_{2c} design detected the highest proportion of quadrats occupied by the target species (Table 5), for all simulations. The two conditions design was more efficient than the one condition design (Table 6). The sample variance for ATS_{2c} was

TABLE 7. Efficiency of \bar{y}_s relative to Murthy estimator for ATS_{2c} for the Poisson cluster process population

	m=8	m=6	m=4
$n_1=3$	2.201	2.226	2.560
$n_1=5$	1.975	2.094	2.030
$n_1=7$	1.914	1.904	1.888

smaller than the variance for the one-condition design for all simulations. The comparison of new estimator, \bar{y}_s with Murthy estimator for ATS_{2c} (Table 7) showed remarkable gains in efficiencies for all the situations. Here the ATS_{2c} design help us to spend the additional samples in more important PSUs.

4.2. Investigating normality of the estimator under MAS. To investigate the normality of the new estimator, \bar{y}_s , we revisited the mussel and the Poisson cluster populations. We used adaptive stratified sequential sampling and constructed a histograms from the simulation results (figures 3, 4, and 5). We calculated the sample fraction, the proportion of the quadrats that were include in the final sample. Normality was tested by Shapiro Normality test with 500 iteration and the p-values are included in figure 3, 4 and 5.

We used $n_1 = 2, 3, 4, 6, 8$ and 10 for popu1 of the mussels and $n_1 = 4, 6, 8, 10, 13$ and 15 for popu3 of the mussels. For both populations we used $d_1 = 3$ and $d_2 = 2$. For the Poisson cluster population we used $n_1 = 3, 5, 7, 9, 11$ and 13 and $d_1 = 1$ and $d_2 = 6$. The conditions for the adaptive selection were the same as in the previous sections.

The shape of the histograms of the new estimators were normal with the higher sample fractions. This was more evident with the lower density populations, the Poisson cluster population, and popu3 of the mussel population. For popu1 of the mussels, when the sample fraction was larger than 0.08 the distribution of the estimator was normal. For popu3 and the Poisson the sample fraction needed to be larger than 0.25 for normality.

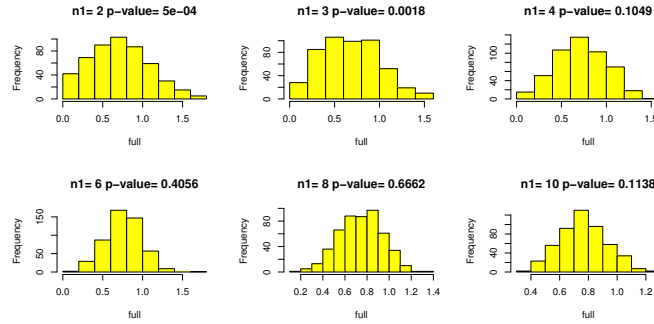


FIGURE 3. Normality of the estimator for popu1, with ” $y_{hi} > 0$ ” as C_1 and ” y_{hi} does not satisfy C_1 and $x_{hi} > 0$ ” as C_2 , and $d_1 = 3, d_2 = 2$. Sample fractions were 0.04, 0.06, 0.08, 0.11, 0.15 and 0.18 respectively.

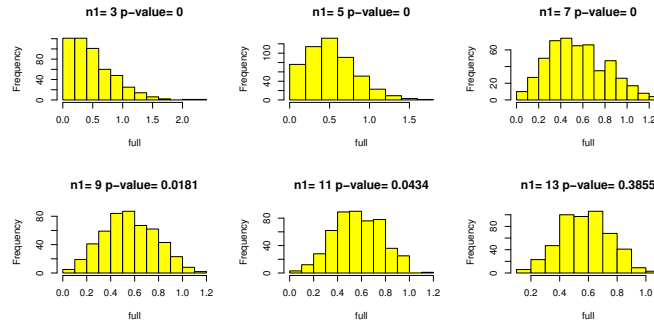


FIGURE 4. Normality of the estimator for the Poisson cluster process, with ” $0 < y_{hi} \leq 2$ ” as C_1 and ” $3 \leq y_{hi}$ ” as C_2 , and $d_1 = 1, d_2 = 6$. Sample fractions were 0.08, 0.12, 0.16, 0.21, 0.25 and 0.29, respectively.

5. Conclusion

We have extended the design for adaptive sequential sampling by including more than one condition. Considerable effort is often spent searching for rare individuals in sampling rare populations. Any design that allows the survey effort to be better focused to where the individuals are likely to occur is a welcome addition to the survey toolbox. In

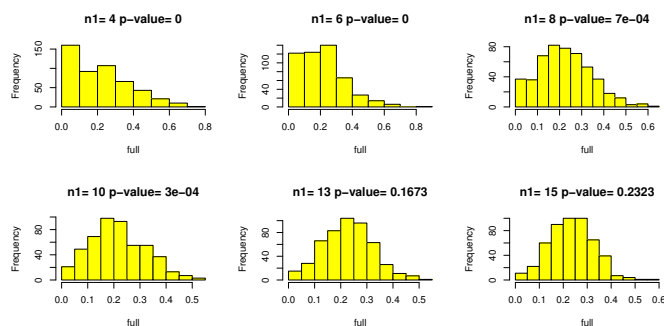


FIGURE 5. Normality of the estimator for popu3, with " $y_{hi} > 0$ " as C_1 and " y_{hi} does not satisfy C_1 and $x_{hi} > 0$ " as C_2 , and $d_1 = 3, d_2 = 2$. Sample fractions were 0.08, 0.11, 0.15, 0.18, 0.23 and 0.27 respectively.

this study the simulation study showed that when there was auxiliary information correlated to the target population, it can be used to help focus, or, adapt the survey effort and improve efficiency.

We illustrated multiple criteria sampling with two examples. In the first example, auxiliary information (e.g., a related species) was used for a second criteria and the adaptive allocation of the sample effort. In the second example, information from the target population was used as the second criteria whether the sample unit had a "high" count ($y_{hi} > 3$). In both examples the two criteria designs had high efficiency especially in finding rare units.

We developed a new estimator and showed that it is consistent and asymptotically normal. We also provided its asymptotic variance estimator to standardize the estimator. For large-population sampling, where the size of the final sample is large, we recommend the use of this new estimator instead of Murthy estimator. Asymptotic normality of the estimator allows standard, well known, inference techniques to be used.

Appendix A: Proof of Theorem 3.2: First assume we have one condition C_1 and $n_2^v = d * l_1 + 1$, where d is fixed before sampling and $N_{C_2}^v = N^v - N_{C_1}^v$ (see Table 8). According to (3.1) (but for one

TABLE 8. Attributes of the design for just one condition

number	condition	Initial phase s_1	Second phase s_2	$s = s_1 \cup s_2$
N^v		n_1^v	$n_2^v = n^v(l_1) = d * l_1 + 1$	$n^v = n_1^v + n_2^v$
$N_{C_1}^v$	C_1	l_1	l_2	$l = l_1 + l_2$
$N_{C_2}^v = N^v - N_{C_1}^v$	$C_2 = C_1'$	$n_1^v - l_1$	$n_2^v - l_2$	$n^v - l$

condition) we have:

$$\begin{aligned}
& \sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v}) \\
&= \underbrace{\sqrt{n_1^v} \left(\sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + \sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v}) \right)}_{\text{from first-phase}} \\
&+ \underbrace{\sqrt{n_2^v} \left(\sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v}) + \sqrt{\frac{n_2^v}{n^v}} (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v}) \right)}_{\text{from second-phase}}.
\end{aligned}$$

Now we need first to state and prove the below lemma.

Lemma 5.1. *In adaptive sequential sampling, under conditions given in Theorem 3.2,*

- (I) $\left(\frac{l_1}{n_1^v} - f_1 \mid \mathcal{F}_{N^v}\right) \xrightarrow{p} 0$, as $v \rightarrow \infty$
- (II) $\left(\frac{n_1^v}{n^v} - \frac{N^v}{N^v + dN_{C_1}^v} \mid \mathcal{F}_{N^v}\right) = o_p(1)$,
- (III) $\left(\frac{n_2^v}{n^v} - \frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} \mid \mathcal{F}_{N^v}\right) = o_p(1)$,
- (IV) $\left(\bar{y}_{x_k s_1} - \bar{Y}_{x_k N^v} \mid \mathcal{F}_{N^v}\right) = O_p(n_1^v)^{-\frac{1}{2}}$.

For proving Lemma 5.1 see Appendix B.
Now for the first phase of $\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})$ according to the above results

$$\begin{aligned}
 & \sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})_{\text{first-phase}} \mid \mathcal{F}_{N^v} \\
 &= \sqrt{n_1^v} \left(\sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + \sqrt{\frac{n_1^v}{n^v}} (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v}) \right) \\
 &= \sqrt{n_1^v} \sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} [(\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v})] \\
 &+ \underbrace{\sqrt{n_1^v} \left(\sqrt{\frac{n_1^v}{n^v}} - \sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} \right)}_{o_p(1)} \underbrace{[(\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v})]}_{O_p(n_1^v)^{-\frac{1}{2}}} + \underbrace{(\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v})}_{O_p(n_1^v)^{-\frac{1}{2}}}.
 \end{aligned}$$

Then

$$\begin{aligned}
 & \sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})_{\text{first-phase}} \mid \mathcal{F}_{N^v} \\
 &= \sqrt{n_1^v} \left(\sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} [(\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v})] \right) \\
 &+ o_p(1).
 \end{aligned}$$

From (C3) we have (see Thompson [24], page 60)

$$\sqrt{n_1^v} \begin{pmatrix} \bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v} \\ \bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v} \end{pmatrix} \mid \mathcal{F}_{N^v} \xrightarrow{d} N(0, (1 - f_1^*)\Gamma).$$

Then for the first phase of sampling using Slutsky Theorem we have:

$$\begin{aligned}
 & \frac{\sqrt{n_1^v} \left(\sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} (\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + \sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v}) \right)}{\sqrt{(1 - f_1^*) \left(\frac{N^v}{N^v + dN_{C_1}^v} \sigma_1^2 + \frac{N^v}{N^v + dN_{C_1}^v} \sigma_2^2 + 2 \frac{N^v}{N^v + dN_{C_1}^v} \sigma_{12} \right)}} \\
 & \mid \mathcal{F}_{N^v}^* \xrightarrow{d} N(0, 1).
 \end{aligned}$$

Now for matching to Theorem 3.1 we can set:

$$\begin{aligned}
 V_{n^v} &= \sqrt{n_1^v} \left(\sqrt{\frac{N^v}{N^v + dN_{C_1}^v}} [(\bar{y}_{x_1 s_1} - \bar{Y}_{x_1 N^v}) + (\bar{y}_{x_2 s_1} - \bar{Y}_{x_2 N^v})] \right), \\
 \sigma_{1n^v}^2 &= (1 - f_1^*) \frac{N^v}{N^v + dN_{C_1}^v} (\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}).
 \end{aligned}$$

Let $\mathbb{B}_{n_1^v}$ be the σ -algebra that contains information from the first phase. Conditional on $\mathbb{B}_{n_1^v}$, the second phase sample is a simple random sample without replacement of the populations \mathcal{F}_{N^v} after excluding s_1

with constant sample size $n_2^v = d \times l_1 + 1$.

let $\begin{pmatrix} \bar{y}_{x_1 N^v - s_1} \\ \bar{y}_{x_2 N^v - s_1} \end{pmatrix}$, $\begin{pmatrix} S_{yx_1 N^v - s_1}^2 & S_{yx_1 x_2 N^v - s_1} \\ S_{yx_1 x_2 N^v - s_1} & S_{yx_2 N^v - s_1}^2 \end{pmatrix}$ are the vectors of mean and the matrix of variance of the remainder population.

Also $\bar{y}_{x_1 s_2}$ and $\bar{y}_{x_2 s_2}$ conditional on $\mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}$, are the mean of a simple random sample without replacement with constant sample size $n_2^v = d \times l_1 + 1$ of the population $\mathcal{F}_{N^v} - s_1$ (\mathcal{F}_{N^v} after excluding s_1). Then we have

$$E(\bar{y}_{x_k s_2} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}) = \bar{Y}_{x_k N^v - s_1}, \quad k = 1, 2$$

For the second phase of $\sqrt{n^v}(\bar{Y}_s - \bar{Y}_{N^v})$ according to the above results and (C4a), it is easy to show that

$$\begin{aligned} & \sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})_{\text{second-phase}} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} = \\ & \sqrt{n_2^v} \sqrt{\frac{dN_{C_1}^v}{N^v + dN_{C_1}^v}} ((\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v - s_1}) + (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v - s_1})) + o_p(1). \end{aligned}$$

Then from moment assumption and (C4b) we have

$$\left[\begin{pmatrix} S_{yx_1 N^v - s_1}^2 & S_{yx_1 x_2 N^v - s_1} \\ S_{yx_1 x_2 N^v - s_1} & S_{yx_2 N^v - s_1}^2 \end{pmatrix} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \right] \longrightarrow \Gamma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \text{ as } v \rightarrow \infty, \text{ and then from (C4c) we have}$$

$$\sqrt{n_2^v} \begin{pmatrix} \bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v - s_1} \\ \bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v - s_1} \end{pmatrix} \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \xrightarrow{d} N(0, (1 - f_2^*)\Gamma).$$

Now for the second phase of sampling according to all above results we have, as $v \rightarrow \infty$

$$\begin{aligned} & \frac{\sqrt{n_2^v} \left(\sqrt{\frac{dN_{C_1}^v}{N^v + dN_{C_1}^v}} (\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v - s_1}) + \sqrt{\frac{dN_{C_1}^v}{N^v + dN_{C_1}^v}} (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v - s_1}) \right)}{\sqrt{(1 - f_2^*) \left(\frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} \sigma_1^2 + \frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} \sigma_2^2 + 2 \frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} \sigma_{12} \right)}} \\ & \mid \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \xrightarrow{d} N(0, 1). \end{aligned}$$

Again for matching to Theorem 3.1 we can set:

$$\begin{aligned} W_{n^v} &= \sqrt{n_2^v} \sqrt{\frac{dN_{C_1}^v}{N^v + dN_{C_1}^v}} ((\bar{y}_{x_1 s_2} - \bar{Y}_{x_1 N^v - s_1}) + (\bar{y}_{x_2 s_2} - \bar{Y}_{x_2 N^v - s_1})), \\ \sigma_{2n^v}^2 &= (1 - f_2^*) \frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} (\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}). \end{aligned}$$

We have

$$\begin{aligned}\sigma_{1n^v}^{-1} V_{n^v} &| \mathcal{F}_{N^v} \xrightarrow{d} N(0, 1), \\ \sigma_{2n^v}^{-1} W_{n^v} &| \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v} \xrightarrow{d} N(0, 1),\end{aligned}$$

and because the normal distribution has a continuous F we have (see Ferguson [10] chapter 1):

$$\sup_t |P(\sigma_{2n^v}^{-1} W_{n^v} \leq t | \mathbb{B}_{n_1^v}, \mathcal{F}_{N^v}) - \Phi(t)| \rightarrow 0.$$

Also we have

$$\begin{aligned}\gamma_{n^v}^2 &= \frac{(1 - f_1^*) \frac{N^v}{N^v + dN_{C_1}^v} ((\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}))}{(1 - f_2^*) \frac{dN_{C_1}^v}{N^v + dN_{C_1}^v} ((\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}))} \\ &= \frac{(1 - f_1^*) N^v}{(1 - f_2^*) dN_{C_1}^v} \rightarrow \frac{(1 - f_1^*)}{(1 - f_2^*) df_1}\end{aligned}$$

and with gathering all condition in Theorem 3.1 we have

$$\frac{V_{n^v} + W_{n^v}}{\sqrt{\sigma_{1n^v}^2 + \sigma_{2n^v}^2}} | \mathcal{F}_{N^v} \xrightarrow{d} N(0, 1),$$

but $\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v}) = V_{n^v} + W_{n^v} + o_p(1)$ and $\sqrt{\sigma_{1n^v}^2 + \sigma_{2n^v}^2} = [(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})(1 - \frac{N^v f_1^* + dN_{C_1}^v f_2^*}{N^v + dN_{C_1}^v})]^{1/2} = O(1)$, Therefore,

$$\frac{\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})}{(1 - \frac{N^v f_1^* + dN_{C_1}^v f_2^*}{N^v + dN_{C_1}^v})^{1/2} \sqrt{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})}} = \frac{V_{n^v} + W_{n^v}}{\sqrt{\sigma_{1n^v}^2 + \sigma_{2n^v}^2}} + o_p(1).$$

Then from Slutsky theorem we have

$$\frac{\sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v})}{(1 - \frac{N^v f_1^* + dN_{C_1}^v f_2^*}{N^v + dN_{C_1}^v})^{1/2} \sqrt{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})}} | \mathcal{F}_{N^v} \xrightarrow{d} N(0, 1).$$

Also for consistency, because of $\frac{V_\infty}{\sqrt{n^v}} \rightarrow^{a.s.} 0$ we can deduce that

$$\bar{y}_s - \bar{Y}_{N^v} | \mathcal{F}_{N^v} \xrightarrow{d} 0$$

and therefore

$$\bar{y}_s - \bar{Y}_{N^v} | \mathcal{F}_{N^v} \xrightarrow{p} 0$$

The proof of Theorem 3.2 is completed.

Appendix B: Proof of Lemma 5.1:

from

$$l_1 \mid \mathcal{F}_{N^v} \sim HG(N^v, N_{C_1}^v, n_1^v), \frac{l_1}{n_1^v} = \frac{1}{n_1^v} \sum_{i \in s_1} x_{1i}$$

it is easy to show that

$$\left(\frac{l_1}{n_1^v} - \frac{N_{C_1}^v}{N^v} \mid \mathcal{F}_{N^v} \right) = O_p(n_1^{v-\frac{1}{2}})$$

and then with (C2a) in Theorem 3.2

$$\left(\frac{l_1}{n_1^v} - f_1 \mid \mathcal{F}_{N^v} \right) = \left(\frac{l_1}{n_1^v} - \frac{N_{C_1}^v}{N^v} \mid \mathcal{F}_{N^v} \right) + \left(\frac{N_{C_1}^v}{N^v} - f_1 \mid \mathcal{F}_{N^v} \right) \xrightarrow{p} 0.$$

Proof of (II) and (III) are easy using (I) and

$$\frac{n_1^v}{n^v} = \frac{1}{1 + d \frac{l_1}{n_1^v} + \frac{1}{n_1^v}}, \frac{n_2^v}{n^v} = 1 - \frac{n_1^v}{n^v}.$$

For (IV), since from moment assumption $S_{y_{x_k}}^2 = O(1)$, we have

$$\begin{aligned} P(n_1^{v\frac{1}{2}} \mid \bar{y}_{x_k s_1} - \bar{Y}_{x_k N^v} \mid > M_\epsilon \mid \mathcal{F}_{N^v}) &= \\ \frac{n^v \text{Var}(\bar{y}_{x_k s_1} - \bar{Y}_{x_k N^v} \mid \mathcal{F}_{N^v})}{M_\epsilon^2} &= \frac{n_1^v \frac{(1 - \frac{n_1^v}{N^v})}{n_1^v} S_{y_{x_k}}^2}{M_\epsilon^2} = \frac{(1 - \frac{n_1^v}{N^v}) O(1)}{M_\epsilon^2} < \epsilon \end{aligned}$$

when we set $M_\epsilon^2 > \frac{M'}{\epsilon}$ and M' is a bound for $(1 - \frac{n_1^v}{N^v}) O(1)$.

Appendix C: Proof of Theorem 3.5:

Because $\sqrt{n_h^v}(\bar{y}_{sh} - \bar{Y}_{N_h^v})$ converges in distribution with finite constant variance, it is easy to show that

$$\begin{aligned} \sqrt{n^v}(\bar{y}_s - \bar{Y}_{N^v}) &= \\ \sum_{h=1}^H \frac{F_h}{\sqrt{f_h^{**}}} \sqrt{n_h^v}(\bar{y}_{sh} - \bar{Y}_{N_h^v}) &+ \sum_{h=1}^H \left(\frac{\sqrt{n^v}}{\sqrt{n_h^v}} W_h^v - \frac{F_h}{\sqrt{f_h^{**}}} \right) \sqrt{n_h^v}(\bar{y}_{sh} - \bar{Y}_{N_h^v}) \\ &= \sum_{h=1}^H \frac{F_h}{\sqrt{f_h^{**}}} \sqrt{n_h^v}(\bar{y}_{sh} - \bar{Y}_{N_h^v}) + o_p(1). \end{aligned}$$

and using Theorem 3.4, we have

$$\left[\sum_{h=1}^H \frac{F_h}{\sqrt{f_h^{**}}} \sqrt{n_h^v}(\bar{y}_{sh} - \bar{Y}_{N_h^v}) \mid (\mathcal{F}_{N_h^v}, h = 1, 2, \dots, H) \right] \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = \sum_{h=1}^H \frac{F_h^2}{f_h^{**}} \sigma_h^2.$$

The rest of the proof follows from $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$ and Slutsky Theorem. The proof of the consistency is obvious according to the previous results.

REFERENCES

- [1] Y. G. Berger, Rate of convergence to normal distribution for the Horvitz-Thompson estimator, *J. Statist. Plann. Inference* **67** (1998), no. 2, 209–226.
- [2] K. R. W. Brewer, A class of robust sampling designs for large-scale surveys, *J. Amer. Statist. Assoc.* **74** (1979), no. 368, 911–915.
- [3] J. A. Brown, M. M. Salehi, M. Moradi, G. Bell and D. R. Smith, An adaptive two-stage sequential sampling, *Popul. Ecol.* **50** (2008) 239–245.
- [4] C. T. Chao, Ratio estimation on adaptive cluster sampling, *J. Chin. Statist. Assoc.* **42** (2004) 307–327.
- [5] J. Chen and J. N. K. Rao, Asymptotic normality under two-phase sampling design, *Statist. Sinica* **17** (2007), no. 3, 1047–1064.
- [6] N. Chutiman, A new ratio estimator in stratified adaptive cluster sampling, *Thailand Statistician* **8** (2010) 223–233.
- [7] P. J. Diggle, *Statistical Analysis of Spatial Point Patterns*, Academic Press, Inc., London, 1983.
- [8] P. Erdos and A. Reni, On the central limit theorem for samples from a finite population, *Magyar Tud. Akad. Mat. Kutató Int. Közl* **4** (1959) 49–57.
- [9] L. Fattorini and C. Pisani, Variance decomposition in two-stage plot sampling: theoretical and empirical results, *Environ. Ecol. Stat.* **11** (2004), no. 4, 385–396.
- [10] T. S. Ferguson, *A Course in Large Sample Theory*, Chapman & Hall, London, 1996.
- [11] W. A. Fuller, *Sampling Statistics*, John Wiley & Sons, Inc., New Jersey, 2009.
- [12] J. Hajek, Limiting distribution in simple random sampling from a finite population, *Magyar Tud. Akad. Mat. Kutató Int. Közl* **5** (1960) 361–374.
- [13] D. J. Hornbach, M. C. Hove, B. D. Dickinson, K. R. MacGregor and J. R. Medland, Estimating population size and habitat associations of two federally endangered mussels in St. Croix River, Minnesota and Wisconsin, *Aquatic Conservation: Marine and Freshwater Ecosystems*. **20** (2010) 250–260.
- [14] D. J. Hornbach, J. G. March, T. Deneka, N. H. Troelstrup and J. A. Perry, Factors influencing the distribution and abundance of the endangered winged mapleleaf *Quadrula fragosa* in the St. Croix River, Minnesota and Wisconsin, *Am. Midl. Nat.* **136** (1996) 278–286.
- [15] L. Isaki and W. A. Fuller, Survey design under the regression superpopulation model, *J. Amer. Statist. Assoc.* **77** (1982), no. 377, 89–96.
- [16] G. Kalton and D. W. Anderson, Sampling rare populations, *J. R. Stat. Soc. Ser. A* **149** (1986) 65–82.
- [17] H. F. Martin, Asymptotics in adaptive cluster sampling, *Environ. Ecol. Stat.* **10** (2003), no. 1, 61–82.

- [18] M. Moradi and M. Salehi, An adaptive allocation sampling design for which the conventional stratified estimator is an appropriate estimator, *J. Statist. Plann. Inference* **140** (2010), no. 4, 1030–1037.
- [19] E. Ohlsson, Asymptotic normality for two-stage sampling from a finite population, *Probab. Theory Related Fields* **81** (1989), no. 3, 341–352.
- [20] B. Rosen, Asymptotic theory for successive sampling with varying probabilities without replacement, I, II, *Ann. Math. Statist.* **43** (1972) 373–397.
- [21] C. E. Särndal, On π -inverse weighting versus best linear unbiased weighting in probability sampling, *Biometrika* **67** (1980), no. 3, 639–650.
- [22] A. J. Scott and C. F. Wu, On the asymptotic distribution of ratio and regression estimators, *J. Amer. Statist. Assoc.* **76** (1981), no. 373, 98–102.
- [23] D. R. Smith, R. F. Vilella and D. P. Lemarié, Application of adaptive cluster sampling to low-density populations of freshwater mussels, *Environ. Ecol. Stat.* **10** (2003), no. 1, 7–15.
- [24] M. E. Thompson, *Theory of Sample Surveys*, Chapman & Hall, London, 1997.

Afshin Parvardeh

Department of Statistics, University of Isfahan, Isfahan, Iran

Email: a.parvardeh@sci.ui.ac.ir

Bardia Panahbehagh

Department of Mathematical Science, Isfahan University of Technology, Isfahan, Iran

Email: bardia_p@math.iut.ac.ir

Mohammad Salehi M.

Department of Mathematics, Qatar University, Doha, Qatar

Email: salehi@qu.edu.qa

and

Department of Mathematical Science, Isfahan University of Technology, Isfahan, Iran

Jennifer Brown

Department of Statistics, University of Canterbury, Christchurch, New Zealand

Email: jennifer.brown@canterbury.ac.nz

David R. Smith

Department of Statistics, U.S. Geological Survey, Kearneysville, West Virginia, USA

Email: drsmith@usgs.gov