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

A. PARVARDEH*, B. PANAHBEHAGH, 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
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 com-
monly used estimators are still not available for many sampling designs.
In this article we consider adaptive sequential sampling where the deci-
sion 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., adap-
tive two-stage sequential sampling, Brown et al. [3]) is the unbiased
Murthy estimator. In section 2, we derive Murthy estimator for mul-
tiple 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 estima-
tor. 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 asymptoti-
cally 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 $\{ (h, i), h = 1, 2, ..., H, i = 1, 2, ..., N_h \}$ 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 $y_i, i = 1, 2, ..., N$. 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 $U_N = \{ y_i, i = 1, 2, ..., N \}$ 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 conditions in ATS design.

### Table 1. Summary of two conditions situations, where $l_1 = l_{11} + l_{12}, l_2 = l_{21} + l_{22}, l = l_{c1} + l_{c2} = l_1 + l_2$.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Initial phase $s_1$</th>
<th>Second phase $s_2$</th>
<th>$s = s_1 \cup s_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{C_1}$</td>
<td>$C_1$</td>
<td>$n_1$</td>
<td>$n_2 = n(l_{11}, l_{12})$</td>
</tr>
<tr>
<td>$N_{C_2}$</td>
<td>$C_2$</td>
<td>$l_{12}$</td>
<td>$l_{22}$</td>
</tr>
<tr>
<td>$N_{C_3} = N - N_{C_1} - N_{C_2}$</td>
<td>$(C_1 \cup C_2)$</td>
<td>$n_1 - l_1$</td>
<td>$n_2 - l_2$</td>
</tr>
</tbody>
</table>

#### 2.1. Murthy estimator for ATS

To estimate the mean of the PSU, we can use

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

where $P(s)$ and $P(s \mid I_i = 1)$ 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

$$n(s) = \binom{l_{c1} - 1}{l_{11} - 1} \binom{l_{c2}}{l_{12}} \binom{n - l_{c1} - l_{c2}}{n_1 - l_{11} - l_{12}} n_1! n_2!$$

and

$$n(s, I_i = 1) =$$

$$\left\{\begin{array}{ll}
\binom{l_{c1} - 1}{l_{11} - 1} \binom{l_{c2}}{l_{12}} & \left(\binom{n - l_{c1} - l_{c2}}{n_1 - l_{11} - l_{12}}\right) (n_1 - 1)! n_2!; \text{ if } y_i \in U_{C_1} \\
\binom{l_{c1} - 1}{l_{11} - 1} \binom{l_{c2} - 1}{l_{12} - 1} & \left(\binom{n - l_{c1} - l_{c2}}{n_1 - l_{11} - l_{12}}\right) (n_1 - 1)! n_2!; \text{ if } y_i \in U_{C_2} \\
\binom{l_{c1} - 1}{l_{11} - 1} \binom{l_{c2} - 1}{l_{12} - 1} & \left(\binom{n - l_{c1} - l_{c2} - 1}{n_1 - l_{11} - l_{12} - 1}\right) (n_1 - 1)! n_2!; \text{ if } y_i \in U_{C_3}
\end{array}\right."
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

\[
P(s | I_i = 1) = \frac{n(s, I_i = 1)}{N n(s) P(I_i = 1)} = \begin{cases} \frac{l_{11}}{n_1}; & \text{if } y_i \in U_{C_1} \\ \frac{l_{12}}{n_1}; & \text{if } y_i \in U_{C_2} \\ \frac{n_1 - l_{11}}{n_1(n - l_i)}; & \text{if } y_i \in U_{C_3} \end{cases}
\]

Therefore

\[
\hat{Y} = \frac{l_{11}}{n_1} y_{s_{c_1}} + \frac{l_{12}}{n_1} y_{s_{c_2}} + \frac{l_{13}}{n_1} y_{s_{c_3}}
\]

\[
= \hat{p}_{11} y_{s_{c_1}} + \hat{p}_{12} y_{s_{c_2}} + (1 - \hat{p}_{11} - \hat{p}_{12}) y_{s_{c_3}}
\]

\[
= \hat{p}_{11} y_{s_{c_1}} + \hat{p}_{12} y_{s_{c_2}} + \hat{p}_{13} y_{s_{c_3}}
\]

where \(s_{c_k}\) is the total sample that are satisfied in condition \(C_k\) and \(y_{s_{c_k}} = \frac{1}{i_{c_k}} \sum_{i \in s_{c_k}} y_i\).

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,

\[
l_{11}, l_{12} \sim HG(N, N_{C_1}, N_{C_2}; n_1),
\]

\[
l_{21}, l_{22} \mid l_{11}, l_{12} \sim HG(N - n_1, N_{C_1} - l_{11}, N_{C_2} - l_{12}; n_2).
\]

where "HG" denotes the multivariate hypergeometric distribution. As shown above, Murthy estimator, \(\hat{Y}\), 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 \((\frac{l_{11}}{n_1} = \hat{p}_{11}, \frac{l_{12}}{n_1} = \hat{p}_{12}, \frac{l_{13}}{n_1} = \hat{p}_{13})\) 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
\[
\bar{y}_s = \frac{l_{c_1}}{n} \bar{y}_{s_{c_1}} + \frac{l_{c_2}}{n} \bar{y}_{s_{c_2}} + \frac{n - l_{c_1} - l_{c_2}}{n} \bar{y}_{s_{c_3}}
\]
\[
= \frac{1}{n} \left( \sum_{i \in s_{c_1}} y_i + \sum_{i \in s_{c_2}} y_i + \sum_{i \in s_{c_3}} y_i \right) = \frac{1}{n} \left( \sum_{i \in s} y_i \right).
\]

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 \( \{z_i\} \) and \( \{z_i^*\} \) be sequences of real numbers, with \( z_i = (y_i, x_{1i}, x_{2i}, x_{3i}) \) and \( z_i^* = (z_{1i}^*, z_{2i}^*, z_{3i}^*) = (y_i x_{1i}, y_i x_{2i}, y_i x_{3i}) \), 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 \( N^v \)th finite population is composed of the first \( N \) values of the sequence \( \{z_i\} \).

Assume that
\[
\lim_{v \to \infty} \sum_{i \in U_{N^v}} (x_{ki} y_i, x_{ki} y_i z_{mi} y_i) \mid \mathcal{F}_{N^v} = (\theta_k, \theta_{km}); \quad m \leq k = 1, 2, 3,
\]
where \( (\theta_k, \theta_{km}), m \leq k = 1, 2, 3 \) are finite and we assume that \( (\theta_k, \theta_{km}) \) 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 "\( i \in U_{N_{C_k}^v} \)" it means every index "\( i \)" that "\( y_i \in U_{N_{C_k}^v} \)."

We then have
\[
z_{N^v} = \frac{1}{N^v} \sum_{i \in U_{N^v}} z_i^* = \left( \frac{\sum_{x_{1i}^N} y_i^N}{\sum_{x_{2i}^N} y_i^N} \frac{\sum_{x_{3i}^N} y_i^N}{\sum_{x_{3i}^N} y_i^N} \right)^t,
\]
where
\[
\bar{Y}_{x_kN^v} = \frac{1}{N^v} \sum_{i \in U_{N^v}} y_i x_{ik} = \frac{N_{C_k}^v}{N^v} \frac{1}{C_k} \sum_{i \in U_{C_k}} y_i = \frac{N_{C_k}^v}{N^v} \bar{Y}_{N^v C_k}.
\]

Also we have
\[
S_{xN^v}^2 = \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (z_i^* - \bar{z}_N^v)'(z_i^* - \bar{z}_N^v)
\]
\[
= \frac{1}{N^v - 1} \begin{pmatrix}
S_{yx1N^v} & S_{yx1x2N^v} & S_{yx1x3N^v} \\
S_{yx1x2N^v} & S_{yx2N^v} & S_{yx2x3N^v} \\
S_{yx1x3N^v} & S_{yx2x3N^v} & S_{yx3N^v}
\end{pmatrix},
\]
where
\[
S_{yxkN^v} = \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (y_i x_{ik} - \bar{Y}_{x_kN^v})^2
\]
\[
S_{yxkxhN^v} = \frac{1}{N^v - 1} \sum_{i \in U_{N^v}} (y_i x_{ik} - \bar{Y}_{x_kN^v})(y_i x_{ih} - \bar{Y}_{x_hN^v}).
\]

According to the moment assumption, the population covariance \(S_{xN^v}^2\) approaches to a positive definite constant matrix \(\Gamma\) as \(v \to \infty\), with
\[
\Gamma = \begin{pmatrix}
\sigma_1^2 & \sigma_{12} & \sigma_{13} \\
\sigma_{12} & \sigma_2^2 & \sigma_{23} \\
\sigma_{13} & \sigma_{23} & \sigma_3^2
\end{pmatrix},
\]
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:
\[
(3.1) \quad \bar{y}_s - \bar{Y}_{N^v} = \frac{n_1^v}{n^v} (\bar{y}_{x1s1} - \bar{Y}_{x1N^v}) + \frac{n_1^v}{n^v} (\bar{y}_{x2s1} - \bar{Y}_{x2N^v})
\]
\[
+ \frac{n_2^v}{n^v} (\bar{y}_{x3s1} - \bar{Y}_{x3N^v}) + \frac{n_2^v}{n^v} (\bar{y}_{x1s2} - \bar{Y}_{x1N^v})
\]
\[
+ \frac{n_2^v}{n^v} (\bar{y}_{x2s2} - \bar{Y}_{x2N^v}) + \frac{n_2^v}{n^v} (\bar{y}_{x3s2} - \bar{Y}_{x3N^v}).
\]

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 \(\sigma\)-algebra. Assume that
(1) there exists $\sigma_{1n^\nu} > 0$ such that

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

as $n^\nu \to \infty$, and $V_n^\nu$ is $\mathcal{B}_n^\nu$ measurable,

(2) $E\{W_n^\nu \mid \mathcal{B}_n^\nu\} = 0$ and $\text{Var}(W_n^\nu \mid \mathcal{B}_n^\nu) = \sigma_{2n^\nu}^2$ such that

$$\sup_t |P(\sigma_{2n^\nu}^{-1} W_n^\nu \leq t \mid \mathcal{B}_n^\nu) - \Phi(t)| = o_p(1),$$

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

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

then as $n^\nu \to \infty$

$$\frac{W_n^\nu + V_n^\nu}{\sqrt{\sigma_{1n^\nu}^2 + \sigma_{2n^\nu}^2}} \xrightarrow{d} N(0, 1).$$

Now, with this background we are ready to formulate a theorem about asymptotic normality of $\sqrt{n^\nu} (Y_s - Y_{N^\nu})$ for just one condition with $n^\nu_2 = d \times l_1 + 1$. We begin with one condition with completely specified situation.

**Theorem 3.2.** In adaptive sequential sampling, let $\mathcal{B}_{n_1^\nu}$, the $\sigma$–algebra generated by $s_1$, contains all the information from the first phase of sampling, and

- C1) $N^\nu, N_{C_1}^\nu, n_1^\nu, n_2^\nu \mid \mathcal{B}_{n_1^\nu} \to \infty$ as $\nu \to \infty$, and $N_{C_1}^\nu < N^\nu, n_1^\nu < N^\nu, n_2^\nu < N^\nu - n_1^\nu$.

- C2a) $(N_{C_1}^\nu \mid \mathcal{F}_{N^\nu})$ is a fixed sequence that converges to $f_k$ as $\nu \to \infty$.

- C2b) $(n_1^\nu \mid \mathcal{F}_{N^\nu})$ is a fixed sequence that converges to $f_1^*$ as $\nu \to \infty$.

- C2c) $(N_{C_1}^\nu - n_1^\nu \mid \mathcal{B}_{n_1^\nu}, \mathcal{F}_{N^\nu})$ is a fixed sequence that converges to $f_2^*$ as $\nu \to \infty$ and

C3) for all $\epsilon > 0$

$$\lim_{\nu \to \infty} \{ \frac{1}{N^\nu} \sum_{i \in U_{N^\nu} \cap \{ i \mid \| \tilde{z}_i^* \| > \epsilon N^\nu n_1^\nu \}} \} = 0,$$

where $\tilde{z}_i^* = z_i^* - z_{N^\nu}^*$, and $\| . \|$ is the Euclidean norm,
Asymptotic properties of the sample mean under MAS design

- **C4a)** \( \left[ Y_{x_1, N^\upsilon} - s_1 - Y_{x_1, N^\upsilon} \right] | \mathbb{B}_{n^\upsilon_1}, F_{N^\upsilon} = O(n^\upsilon_2 \cdot \frac{1}{2 + \delta}) \) and \( \left[ Y_{x_2, N^\upsilon} - s_1 - Y_{x_2, N^\upsilon} \right] | \mathbb{B}_{n^\upsilon_1}, F_{N^\upsilon} = O(n^\upsilon_2 \cdot \frac{1}{2 + \delta}) \) for some \( \delta > 0 \), where \( N^\upsilon - s_1 \) means whole population after excluding \( s_1 \),

- **C4b)** \( \left[ S_{yx_1, N^\upsilon} - s_1 S_{yx_1, N^\upsilon} S_{yx_2, N^\upsilon} - s_1 S_{yx_2, N^\upsilon} \right] | \mathbb{B}_{n^\upsilon_1}, F_{N^\upsilon} | \rightarrow 0 \) as \( \upsilon \rightarrow \infty \),

- **C4c)** for all \( \epsilon > 0 \)

\[
\lim_{\upsilon \to \infty} \{ \frac{1}{N^\upsilon - n^\upsilon_1} \sum_{\{i \in U_{N^\upsilon} : \| \tilde{z}_i \| > \epsilon \}} \frac{\| \tilde{z}_i \|^2}{n^\upsilon - n^\upsilon_1} \} = 0.
\]

Then as \( \upsilon \rightarrow \infty \)

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

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,

\[
(\bar{y}_s - \bar{Y}_{N^\upsilon} | F_{N^\upsilon}) = o_p(1)
\]

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

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

where

\[
\hat{\sigma} = \left(1 - \frac{\hat{f}_1^* + df_1 \hat{f}_2^*}{1 + df_1} \right)^{\frac{1}{2}} \sqrt{(s_1^2 + s_2^2 + 2s_{12})}
\]
and
\[ \hat{f}_1 = \frac{l_1}{n_1^v}, \hat{f}_1^* = \frac{n_1^v}{N^v}, \hat{f}_2 = \frac{n_2^v}{N^v}, \]
\[ s_k^2 = \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})^2, k = 1, 2, \]
\[ s_{12} = \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{1i} - \bar{Y}_{x_1 n_1^v})(y_i x_{2i} - \bar{Y}_{x_2 n_1^v}), \bar{Y}_{x_k n_1^v} = \frac{1}{n_1^v} \sum_{i \in s_1} y_i x_{ki}. \]

Proof. Using the conditions in Theorem 3.2, it is easy to show that:
\[ (1 - \hat{f}_1^* + d \hat{f}_1 \hat{f}_2^*)^\frac{1}{2} \sqrt{(s_1^2 + s_2^2 + 2s_{12})} \mid F_{N^v} \]
\[ \quad \quad \quad \quad \quad \quad \quad \quad \rightarrow (1 - \hat{f}_1^* + d \hat{f}_1 \hat{f}_2^*)^\frac{1}{2} \sqrt{(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12})}. \]

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

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:
\[ \sqrt{n^v(\bar{Y}_s - \bar{Y}_{N^v})} = \sqrt{n_1^v} \left[ \sqrt{\frac{n_1^v}{n^v}(\bar{Y}_{x_1 s_1} - \bar{Y}_{x_1 N^v})} + \sqrt{\frac{n_1^v}{n^v}(\bar{Y}_{x_2 s_1} - \bar{Y}_{x_2 N^v})} \right. \]
\[ + \sqrt{\frac{n_2^v}{n^v}(\bar{Y}_{x_2 s_2} - \bar{Y}_{x_2 N^v})} + \sqrt{\frac{n_2^v}{n^v}(\bar{Y}_{x_3 s_2} - \bar{Y}_{x_3 N^v})} \]
\[ \left. + \sqrt{\frac{n_2^v}{n^v}(\bar{Y}_{x_3 s_2} - \bar{Y}_{x_3 N^v})} \right]. \]

We now have the following theorem.

**Theorem 3.4.** In adaptive sequential sampling, let \( B_{n_1^v} \), the \( \sigma \)-algebra generated by \( s_1 \), contains all the information in first phase of sampling, and

- 1) \( N^v, N_1^v, n_1^v, n_2^v \mid B_{n_1^v} \rightarrow \infty \) 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) \( (\frac{n_1^v}{N^v} \mid F_{N^v}) \) is a fixed sequence that converge to \( f_1^* \) as \( v \rightarrow \infty \),
• 2b) \( \left( \frac{n_{2v}}{N_{v} - n_{1v}} \right) = \mathbb{B}_{n_{1v}, F_{N_{v}}} \) is a fixed sequence that converge to \( f_{1}^2 \) as \( v \to \infty \),

• 2c) \( \frac{n_{1v}}{n} - a \mid F_{N_{v}} \) \( \xrightarrow{p} 0 \),

• 3) for all \( \epsilon > 0 \)

\[
\lim_{v \to \infty} \left\{ \frac{1}{N_{v} - n_{1v}} \sum_{\{i \in U_{N_{v}} \cap \{i : \|\tilde{z}_{i}^*\| > \epsilon (N_{v} - n_{1v})^{\frac{1}{2}} (1 - \frac{n_{1v}}{N_{v}})\}} \|\tilde{z}_{i}^*\|^2 \mid F_{N_{v}} \} = 0,
\]

where \( \tilde{z}_{i}^* = z_{i}^* - \hat{z}_{N_{v}}^* \), and \( \|\| \) is the Euclidean norm,

• 4a) \( \left[ (Y_{x_{k}N_{v} - s_{1}} - Y_{x_{k}N_{v}}) \mid \mathbb{B}_{n_{1v}, F_{N_{v}}} \right] = O(\delta^{1/2 + s}) \), for \( k = 1, 2, 3 \) and some \( \delta > 0 \) where \( N_{v} - s_{1} \) means the whole population after excluding \( s_{1} \),

• 4b) \( \left[ \begin{pmatrix} S_{y_{1x_{2}N_{v} - s_{1}}^2} & S_{y_{1x_{2}N_{v} - s_{1}}y_{1x_{2}N_{v} - s_{1}}} & S_{y_{1x_{2}N_{v} - s_{1}}y_{1x_{2}N_{v} - s_{1}}} \\ S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} & S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} & S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} \\ S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} & S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} & S_{y_{1x_{2}N_{v} - s_{1}}x_{2}N_{v} - s_{1}} \end{pmatrix} \right] \xrightarrow{\mathbb{B}_{n_{1v}, F_{N_{v}}}} 0 \) as \( v \to \infty \),

• 4c) for all \( \epsilon > 0 \)

\[
\lim_{v \to \infty} \left\{ \frac{1}{N_{v} - n_{1v}} \sum_{\{i \in U_{N_{v}} \cap \{i : \|\tilde{z}_{i}^*\| > \epsilon (N_{v} - n_{1v})^{\frac{1}{2}} (1 - \frac{n_{1v}}{N_{v}})\}} \|\tilde{z}_{i}^*\|^2 \mid \mathbb{B}_{n_{1v}, F_{N_{v}}} \} = 0.
\]

Then as \( v \to \infty \)

\[
\sqrt{\frac{n_{v}}{\sigma}} \left( \bar{y}_{s} - \overline{Y}_{N_{v}} \right) \xrightarrow{d} N(0, 1), \quad (\bar{y}_{s} - \overline{Y}_{N_{v}} \mid F_{N_{v}}) = o_{p}(1)
\]

where

\[
\hat{\sigma} = [\hat{a}(1 - \hat{f_{1}^2}) + (1 - \hat{a})(1 - \hat{f_{2}^2})]^{\frac{1}{2}} \sqrt{s_{1}^2 + s_{2}^2 + s_{3}^2 + 2s_{12} + 2s_{13} + 2s_{23}},
\]
and

\[
\hat{a} = \frac{n_1^v}{N^v}, \hat{f}_1 = \frac{n_1^v}{N^v}, \hat{f}_2 = \frac{n_2^v}{N^v},
\]

\[
s_k^2 = \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})^2, k = 1, 2, 3
\]

\[
s_{km} = \frac{1}{n_1^v - 1} \sum_{i \in s_1} (y_i x_{ki} - \bar{Y}_{x_k n_1^v})(y_i x_{mi} - \bar{Y}_{x_m n_1^v}), m < k = 1, 2, 3
\]

\[
\bar{Y}_{x_k n_1^v} = \frac{1}{n_1^v} \sum_{i \in s_1} y_i x_{ki}, k = 1, 2, 3.
\]

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 \( h = 1, 2, \ldots, H \)

\[
\sqrt{n_h^v} \frac{(\bar{y}_{sh} - \bar{Y}_{N_k^v})}{\sigma_h} \overset{d}{\rightarrow} N(0, 1),
\]

where

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

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

\[
W_h^v = \frac{N_h^v}{N^v},
\]
Asymptotic properties of the sample mean under MAS design

\[ y_s - Y_N^\upsilon = H \sum_{h=1}^H W_{\upsilon h} (y_{sh} - Y_N^\upsilon_{\upsilon h}). \]

According to previous section we can restate the final theorem as below:

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

\[ (1) \left( n_{\upsilon h}^u - f_{\ast \ast h}^u \mid F_N^{\upsilon h} \right) \xrightarrow{P} 0, \quad h = 1, 2, \ldots, H \]

as \( \upsilon \to \infty \), where

\[ n_{\upsilon} = \sum_{h=1}^H n_{\upsilon h}, \quad n_{\upsilon h} = n_{\upsilon 1h} + n_{\upsilon 2h}, \]

and \( W_{\upsilon h} = W_{\upsilon 1h} \), \( W_{\upsilon h} = W_{\upsilon 2h} \), \( W_{\upsilon} = W_{\upsilon 1} + W_{\upsilon 2} \),

then

\[ \sqrt{n_{\upsilon}} \left( y_s - Y_N^\upsilon \right) \xrightarrow{d} N(0, 1), \quad y_s - Y_N^\upsilon \mid F_N^{\upsilon h} = o_p(1) \]

where \( \left( y_s - Y_N^\upsilon \right) \) is the right tail of the normal distribution containing \( \frac{\alpha}{2} \) probability \((0 < \alpha < 1)\).

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

\[ \left( y_s - z_{\alpha/2} \hat{\sigma} \sqrt{n_{\upsilon}}, y_s + z_{\alpha/2} \hat{\sigma} \sqrt{n_{\upsilon}} \right) \]

where

\[ \hat{\sigma}^2 = \sum_{h=1}^H \hat{F}_{\upsilon h}^2, \]

and

\[ \hat{F}_{\upsilon h} = \frac{N_{\upsilon h}}{N_{\upsilon}}, \quad \hat{F}_{\upsilon 1h} = \frac{n_{\upsilon 1h}}{n_{\upsilon h}}, \quad \hat{F}_{\upsilon 2h} = \frac{n_{\upsilon 2h}}{n_{\upsilon h}} \]

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

See Appendix C for the proof of Theorem 3.5.

Theorem 3.5. In adaptive stratified sequential sampling, under the conditions in Theorem 3.4, with fixed \( H \), \( g \)

\[ (1) \left( \hat{a}_h(1 - \hat{f}_{\ast h}^u) + (1 - \hat{a}_h)(1 - \hat{f}_{\ast 1 h}) \right) \xrightarrow{P} 0, \quad h = 1, 2, \ldots, H \]

as \( \upsilon \to \infty \), where

\[ \hat{a}_h = \frac{n_{\upsilon 1h}}{n_{\upsilon h}} \]
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
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 $n_1 = 3, 6, 9, 12, n_2 = d_1 \cdot l_{11} + d_2 \cdot l_{12} + 1$ 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 $n_1 = 5, n_2 = d \cdot l_1 + 1$ 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 \( ATS_{2c} \).

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.

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

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

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

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

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

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.
Asymptotic properties of the sample mean under MAS design

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

<table>
<thead>
<tr>
<th></th>
<th>m=8</th>
<th>m=6</th>
<th>m=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_1=3$</td>
<td>2.201</td>
<td>2.226</td>
<td>2.560</td>
</tr>
<tr>
<td>$n_1=5$</td>
<td>1.975</td>
<td>2.094</td>
<td>2.030</td>
</tr>
<tr>
<td>$n_1=7$</td>
<td>1.914</td>
<td>1.904</td>
<td>1.888</td>
</tr>
</tbody>
</table>

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 $n_1 = 2, 3, 4, 6, 8, 10$ for popu1 of the mussels and $n_1 = 4, 6, 8, 10, 13, 15$ for popu3 of the mussels. For both populations we used $d_1 = 3$ and $d_2 = 2$. For the Poisson cluster population we used $n_1 = 3, 5, 7, 9, 11, 13$ 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.
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.

In 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 $n_2 = d * l_1 + 1$, where $d$ is fixed before sampling and $N_{C_2} = N - N_{C_1}$ (see Table 8). According to (3.1) (but for one
Table 8. Attributes of the design for just one condition

<table>
<thead>
<tr>
<th>number</th>
<th>condition</th>
<th>Initial phase $s_1$</th>
<th>Second phase $s_2$</th>
<th>$s = s_1 \cup s_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>$C_1$</td>
<td>$n_1$</td>
<td>$n_2 = n_1 + l_1 = d \cdot l_1 + 1$</td>
<td>$n = n_1 + n_2$</td>
</tr>
<tr>
<td>$N_{C_1}$</td>
<td>$C_1$</td>
<td>$n_1^* = l_1$</td>
<td>$n_2^* - l_2$</td>
<td>$n^* - l$</td>
</tr>
<tr>
<td>$N_{C_2} = N^* - N_{C_1}$</td>
<td>$C_2 = C_1'$</td>
<td>$n_1^* - l_1$</td>
<td>$n_2^* - l_2$</td>
<td>$n^* - l$</td>
</tr>
</tbody>
</table>

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) \( \left( \frac{l_1}{n_1^*} - f_1 \mid F_{N^*} \right) \xrightarrow{p} 0, \) as \( v \to \infty \)
- (II) \( \left( \frac{n_1^*}{n_1} \frac{N_{C_1}}{N^* + d N_{C_1}} \mid F_{N^*} \right) = o_p(1), \)
- (III) \( \left( \frac{n_2^*}{n_2} \frac{N_{C_1}}{N^* + d N_{C_1}} \mid F_{N^*} \right) = o_p(1) , \)
- (IV) \( \left( \frac{\bar{y}_{x_k s_1} - \bar{Y}_{x_k N^*}}{\bar{y}_{x_k N^*}} \mid F_{N^*} \right) = O_p(n_1^{-\frac{1}{2}}). \)

For proving Lemma 5.1 see Appendix B.

Now for the first phase of \( \sqrt{n^*} \left( \bar{y}_s - \bar{Y}_{N^*} \right) \) according to the above results
\[
\sqrt{n^v}(\bar{y}_s - \bar{Y}_N^v)_{\text{first-phase}} \mid \mathcal{F}_{N^v}
\]
\[
= \sqrt{n_1^v}(\sqrt{n_1^v}(\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + \sqrt{n_1^v}(\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v}))
\]
\[
= \sqrt{n_1^v}\left(\frac{N^v}{N^v + dN_C^1}\right)[(\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + (\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v})]
\]
\[
+ \sqrt{n_1^v}\left(\frac{n_1^v}{n^v - \sqrt{N^v + dN_C^1}}\left((\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + (\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v})\right)\right)_{op(1)}
\]

Then
\[
\sqrt{n^v}(\bar{y}_s - \bar{Y}_N^v)_{\text{first-phase}} \mid \mathcal{F}_{N^v}
\]
\[
= \sqrt{n_1^v}\left(\frac{N^v}{N^v + dN_C^1}\right)[(\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + (\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v})]_{op(n_1^v \frac{1}{2})}
\]
\[+ o_p(1).
\]

From (C3) we have (see Thompson [24], page 60)
\[
\sqrt{n_1^v}\left(\frac{\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}}{\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v}}\right) \mid \mathcal{F}_{N^v} \xrightarrow{d} N(0, (1 - f_1^*)\Gamma).
\]

Then for the first phase of sampling using Slutsky Theorem we have:
\[
\sqrt{n_1^v}\left(\frac{N^v}{N^v + dN_C^1}\right)[(\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + (\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v})]
\]
\[
\sqrt{(1 - f_1^*)((\frac{N^v}{N^v + dN_C^1}\sigma_1^2 + \frac{N^v}{N^v + dN_C^1}\sigma_2^2 + 2\frac{N^v}{N^v + dN_C^1}\sigma_{12})}
\]
\[\mid \mathcal{F}_{N^v}^* \xrightarrow{d} N(0, 1).
\]

Now for matching to Theorem 3.1 we can set:
\[
V_{N^v} = \sqrt{n_1^v}\left(\frac{N^v}{N^v + dN_C^1}\right)[(\bar{y}_{x_1s_1} - \bar{y}_{x_1N^v}) + (\bar{y}_{x_2s_1} - \bar{y}_{x_2N^v})],
\]
\[
\sigma_{1N^v}^2 = (1 - f_1^*)\frac{N^v}{N^v + dN_C^1}(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}).
\]

Let \( \mathbb{B}_{n_1^v} \) be the \( \sigma \) – 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 $n_2^\nu = d \times l_1 + 1$.

Let \( \left( \bar{y}_{x1 N^\nu - s_1} \right) \), \( S^2_{y x1 N^\nu - s_1} \), \( S^2_{y x1 x2 N^\nu - s_1} \) are the vectors of mean and the matrix of variance of the remainder population.

Also \( \bar{y}_{x1 s_2} \) and \( \bar{y}_{x2 s_2} \) conditional on \( B_{n_1^\nu}, F_{N^\nu} \), are the mean of a simple random sample without replacement with constant sample size $n_2^\nu = d \times l_1 + 1$ of the population \( F_{N^\nu - s_1}(F_{N^\nu} \text{ after excluding } s_1) \). Then we have

\[
E(\bar{y}_{x k s_2} \mid B_{n_1^\nu}, F_{N^\nu}) = Y_{x k N^\nu - s_1}, \; k = 1, 2
\]

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

\[
\sqrt{n^\nu}(\bar{y}_s - \bar{Y}_{N^\nu})_{\text{second-phase}} \mid B_{n_1^\nu}, F_{N^\nu} = \sqrt{n_2^\nu} \sqrt{\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}} (\bar{y}_{x1 s_2} - \bar{Y}_{x1 N^\nu - s_1}) + (\bar{y}_{x2 s_2} - \bar{Y}_{x2 N^\nu - s_1}) + o_p(1).
\]

Then from moment assumption and (C4b) we have

\[
\left[ \begin{array}{cc}
S^2_{y x1 N^\nu - s_1} & S^2_{y x1 x2 N^\nu - s_1} \\
S^2_{y x1 x2 N^\nu - s_1} & S^2_{y x2 N^\nu - s_1}
\end{array} \right] \mid B_{n_1^\nu}, F_{N^\nu} \rightarrow \Gamma = \left( \begin{array}{cc}
\sigma^2_1 & \sigma^2_{12} \\
\sigma^2_{12} & \sigma^2_2
\end{array} \right)
\]

\( u \rightarrow \infty \), and then from (C4c) we have

\[
\sqrt{n_2^\nu} \left( \bar{y}_{x1 s_2} - \bar{Y}_{x1 N^\nu - s_1}, \bar{y}_{x2 s_2} - \bar{Y}_{x2 N^\nu - s_1} \right) \mid B_{n_1^\nu}, F_{N^\nu} \rightarrow N(0, (1 - f^*_2) \Gamma).
\]

Now for the second phase of sampling according to all above results we have, as \( u \rightarrow \infty \)

\[
\sqrt{n_2^\nu} \sqrt{\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}} (\bar{y}_{x1 s_2} - \bar{Y}_{x1 N^\nu - s_1}) + \sqrt{\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}} (\bar{y}_{x2 s_2} - \bar{Y}_{x2 N^\nu - s_1})
\]

\[
\sqrt{1 - f^*_2}((\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}) \sigma^2_1 + (\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}) \sigma^2_2 + 2 (\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}) \sigma_{12})
\]

\( \mid B_{n_1^\nu}, F_{N^\nu} \rightarrow N(0, 1) \).

Again for matching to Theorem 3.1 we can set:

\[
W_{n^\nu} = \sqrt{n_2^\nu} \sqrt{\frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu}} ((\bar{y}_{x1 s_2} - \bar{Y}_{x1 N^\nu - s_1}) + (\bar{y}_{x2 s_2} - \bar{Y}_{x2 N^\nu - s_1}))
\]

\[
\sigma^2_{2 n^\nu} = (1 - f^*_2) \frac{d N_{C_1}^2}{N^\nu + d N_{C_1}^\nu} (\sigma^2_1 + \sigma^2_2 + 2 \sigma_{12}).
\]
We have
\[ \sigma_{1n^v}^{-1} V_{n^v} \mid F_{n^v} \xrightarrow{d} N(0, 1), \]
\[ \sigma_{2n^v}^{-1} W_{n^v} \mid B_{n^v}, F_{n^v} \xrightarrow{d} N(0, 1), \]
and because the normal distribution has a continuous \( F \) we have (see Ferguson [10] chapter 1):
\[ \sup_t \left| P(\sigma_{2n^v}^{-1} W_{n^v} \leq t \mid B_{n^v}, F_{n^v}) - \Phi(t) \right| \to 0. \]

Also we have
\[ \gamma_{n^v}^2 = (1 - f_1^*) \frac{\sigma_{1n^v}^2 + \sigma_{2n^v}^2 + 2\sigma_{12}}{\sigma_{1n^v}^2 + \sigma_{2n^v}^2} \]
and with gathering all condition in Theorem 3.1 we have
\[ \sqrt{n^v} (\bar{y}_s - \bar{Y}_{n^v}) \mid \sigma_{1n^v}^2 + \sigma_{2n^v}^2 \xrightarrow{d} N(0, 1), \]
but \[ \sqrt{n^v} (\bar{y}_s - \bar{Y}_{n^v}) = V_{n^v} + W_{n^v} + o_p(1) \]
and
\[ \sqrt{n^v} (\bar{y}_s - \bar{Y}_{n^v}) = \frac{(1 - f_1^*) \sigma_{1n^v}^2 + \sigma_{2n^v}^2 + 2\sigma_{12}}{(1 - f_1^*) \sigma_{1n^v}^2 + \sigma_{2n^v}^2} = O(1), \]
Therefore,
\[ \sqrt{n^v} (\bar{y}_s - \bar{Y}_{n^v}) = \frac{V_{n^v} + W_{n^v}}{\sqrt{\sigma_{1n^v}^2 + \sigma_{2n^v}^2}} + o_p(1). \]

Then from Slutsky theorem we have
\[ \sqrt{n^v} (\bar{y}_s - \bar{Y}_{n^v}) \mid \sigma_{1n^v}^2 + \sigma_{2n^v}^2 \xrightarrow{d} N(0, 1). \]

Also for consistency, because of \( \frac{V_{n^v}}{\sqrt{n^v}} \to a.s. 0 \) we can deduce that
\[ \bar{y}_s - \bar{Y}_{n^v} \mid F_{n^v} \xrightarrow{d} 0 \]
and therefore
\[ \bar{y}_s - \bar{Y}_{n^v} \mid F_{n^v} \xrightarrow{p} 0. \]
The proof of Theorem 3.2 is completed.

**Appendix B: Proof of Lemma 5.1:**

from

\[ l_1 \mid F_{N^v} \sim HG(N^v, N_{C^1}, n^v_{1}) \]

\[ \frac{l_1}{n^v_{1}} = \frac{1}{n^v_{1}} \sum_{i \in s_1} x_{1i} \]

it is easy to show that

\[ \left( \frac{l_1}{n^v_{1}} - N_{C^1} \mid F_{N^v} \right) = O_p(n^{-\frac{1}{2}}) \]

and then with (C2a) in Theorem 3.2

\[ \left( \frac{l_1}{n^v_{1}} - f_1 \mid F_{N^v} \right) = \left( \frac{l_1}{n^v_{1}} - N_{C^1} \mid F_{N^v} \right) + \left( N_{C^1} - f_1 \mid F_{N^v} \right) \rightarrow 0. \]

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

\[ \frac{n^v_{1}}{n^v} = \frac{1}{1 + d \frac{l_1}{n^v} + \frac{1}{n^v_1}}, \quad \frac{n^v_{2}}{n^v} = 1 - \frac{n^v_{1}}{n^v}. \]

For (IV), since from moment assumption \( S_{y_k} = O(1) \), we have

\[ P\left( n^v_{1} \gamma_{x_{s_1}} - \gamma_{N^v} > M_\epsilon \mid F_{N^v} \right) = \]

\[ \frac{n^v \text{Var}\left( \gamma_{x_{s_1}} - \gamma_{N^v} \mid F_{N^v} \right)}{M^2_\epsilon} = \frac{n^v(1 - \frac{n^v}{n_1^v})S_{y_k}^2}{M^2_\epsilon} = \frac{(1 - \frac{n^v}{n_1^v})O(1)}{M^2_\epsilon} < \epsilon \]

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

**Appendix C: Proof of Theorem 3.5:**

Because \( \sqrt{n^v_{h}}(\gamma_{sh} - \gamma_{N^v}) \) converges in distribution with finite constant variance, it is easy to show that

\[ \sqrt{n^v}(\gamma_s - \gamma_{N^v}) = \]

\[ \sum_{h=1}^{H} \frac{F_h}{\sqrt{f_h}} \sqrt{n^v_{h}}(\gamma_{sh} - \gamma_{N^v}) + \sum_{h=1}^{H} \left( \frac{\sqrt{n^v_{h}}}{\sqrt{n^v_{h}}} W_{h} - \frac{F_h}{\sqrt{f_h}} \right) \sqrt{n^v_{h}}(\gamma_{sh} - \gamma_{N^v}) \]

\[ = \sum_{h=1}^{H} \frac{F_h}{\sqrt{f_h}} \sqrt{n^v_{h}}(\gamma_{sh} - \gamma_{N^v}) + o_p(1). \]

and using Theorem 3.4, we have

\[ \sum_{h=1}^{H} \frac{F_h}{\sqrt{f_h}} \sqrt{n^v_{h}}(\gamma_{sh} - \gamma_{N^v}) \mid (F_{N^v}, h = 1, 2, \ldots, H) \xrightarrow{d} N(0, \sigma^2), \]
where
\[ \sigma^2 = \sum_{h=1}^{H} \frac{f_{h}^2}{f_{h}^*} \sigma_h^2. \]

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

References

Asymptotic properties of the sample mean under MAS design


Afshin Parvardeh
Department of Statistics, University of Isfahan, Isfahan, Iran
Email: a.parvardeh@sci.ui.ac.ir

Bardia Panahbejahgh
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