Thailand Statistician
July 2009; 7(2) : 101-129
www.statassoc.or.th
Contributed paper

Investigation of Variance Estimators for Adaptive Cluster Sampling with a Single Primary Unit

Urairat Netharn

School of Applied Statistics, National Institute of Development Administration,
Bangkapi, Bangkok 10240, Thailand.

E-mail: ffisurnh@ku.ac.th, kokayja@yahoo.com

Received: 15 October2008

Accepted: 8 June 2009.

Abstract

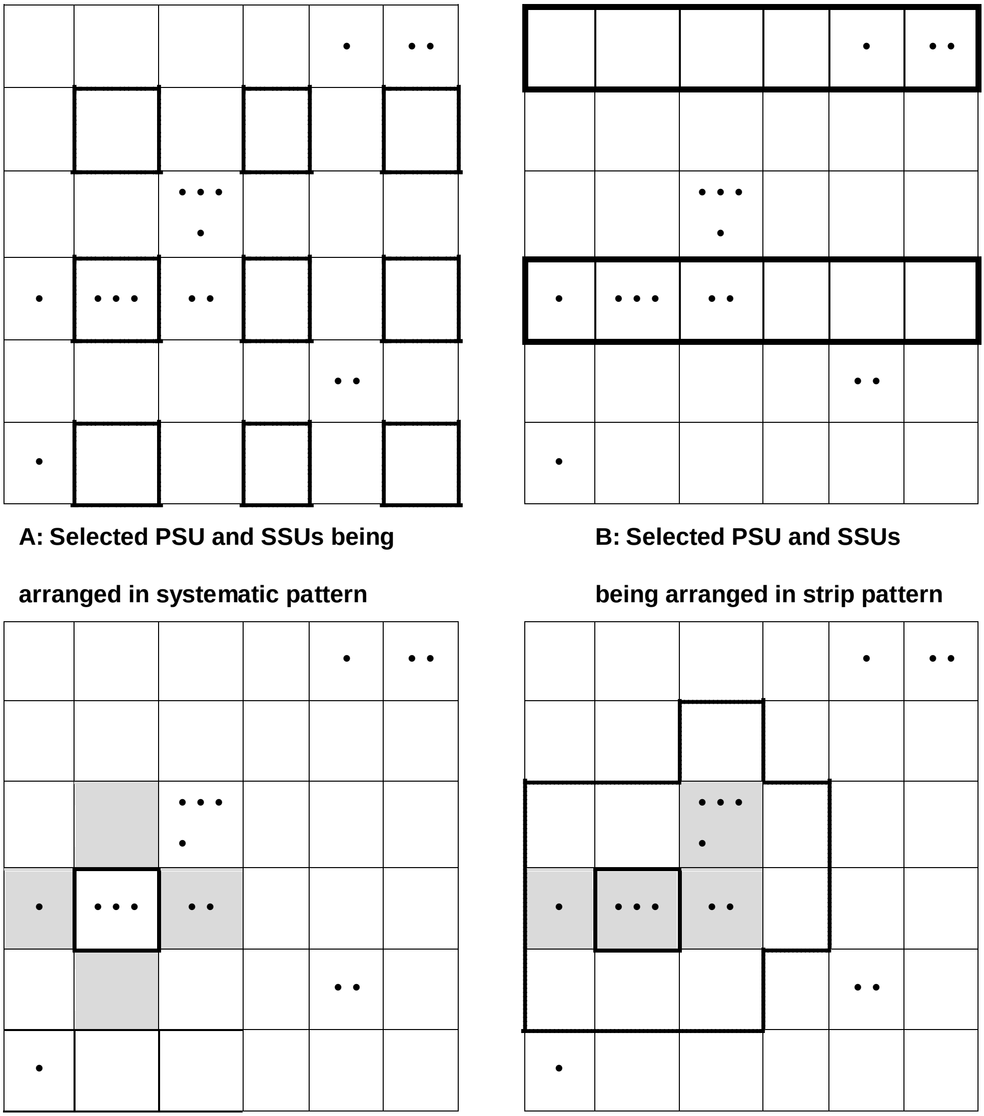
In adaptive cluster sampling design, the initial sample is considered in terms of primary and secondary sampling units. The secondary sampling units being arranged in systematic or strip pattern. The problem of unbiased variance estimation occurs when a single primary sampling unit is selected; consequently, the objective of this study is to introduce three alternative estimators of the variance of unbiased estimator of the population mean proposed by Thompson [1] for the design of adaptive cluster sampling. All three new variance estimators are performed based on the method of splitting the sample and regarding the sample as a stratified sample of equal size and a simple random sample. The numerical study is carried out with real blue-wing teal data in order to investigate the properties of the variance estimators. These three variance estimators are compared to each other empirically in terms of relative bias, MSE and relative frequency of coverage the population mean of 95% confidence interval that is formed by each variance estimator. In addition, all three estimators are also compared to other variance estimators in systematic sampling. The confidence interval containing the population mean is performed based on the assumption of asymptotic normality.

Keywords: adaptive cluster sampling, systematic sampling, variance estimation.

1. Introduction

Adaptive cluster sampling, a design with primary and secondary units, was proposed by Thompson [1], and is a suitable design for rare and cluster populations, especially for a biological population. In this design, a primary sampling unit (PSU) is composed of units that are called secondary sampling units (SSUs). Secondary sampling units within any primary sampling unit may arrange in strip or systematic pattern. As we see in the figure 1, the study area of size 36 meter squares is divided into 36 plots of equal size. The y-value within a plot is the count of black dots. Suppose that a case of a single primary sampling unit chosen is considered. By using the systematic sampling method there are 4 possible primary sampling units. Figure 1A is one of them: the selected primary sampling unit is a set of 9 units with dash lines. Notice that all of these units being arranged in systematic order. The selected primary sampling unit consists of 12 secondary sampling units being arranged in strip pattern is shown in figure 1B. The selected primary sampling unit is called an initial sample. The condition for adding units is that y-values must be equal to or greater than *c,* a pre-specified value*.* Whenever a secondary sampling unit within an initial sample satisfies the condition, its neighborhood will be added to the sample. Note that if a u_{ij} is in the neighborhood of unit $u_{i'j'}$, then unit $u_{i'j'}$ is also in the neighborhood of unit u_{ij}. A secondary sampling unit that is not in the initial sample is called an adaptively added unit. An adaptively added unit may or may not satisfy the condition, but if it does not, it will be called an *edge unit*. A set of secondary sampling units that satisfies the condition is called a *network*. A network and its associated edge units make up a *cluster*.

Although unbiased estimators of the population mean and their variances for adaptive cluster sampling design with primary and secondary sampling units [1] are available, all of the variance estimators require an initial sample of at least two primary sampling units (PSUs). This requirement sometimes is too difficult to apply in a real situation, for example, when a study area is large or the study area is composed of a large number of secondary sampling units in a primary sampling unit. In addition, an unbiased estimator of the variance of unbiased estimator of the population mean proposed by Thompson [1] cannot be obtained in this case. As stated in Thompson [1], “For an initial systematic sample with only one starting point (i.e., only one primary sampling unit is selected), some of the joint inclusion probabilities are zero, underscoring the fact that an unbiased estimator of variance is not available for such a design.” This problem is likely to occur with conventional sampling when a conventional systematic sample is used. In such a case, eight biased estimators of the variance of the sample mean are reviewed and compared, as for example in Cochran [2] and Wolter [3].



A: Selected PSU and SSUs being arranged in systematic pattern

B: Selected PSU and SSUs being arranged in strip pattern

C: Network

D: Cluster

Figure 1. An example of a selected initial sample from adaptive clustered sampling with a single primary sampling unit.

According to the study of Wolter a variance estimator obtained by regarding the sample as a simple random sample is recommended to be used in real practice because its bias and MSE is reasonably small for all population types. The second variance estimator obtained by regarding the sample as a stratified sample might be preferable in a real population where the characteristic type is linear trend or stratification effect. Though it is biased its bias often smaller than the other estimators. Another interesting variance estimator is the estimator obtained from spiting the sample into sub-samples. Though its bias is very large when the number of sub-samples is two [3,4], its form is simple for application and its bias be reduced when the number of sub-samples is increased. These three methods of estimation are to be used as a guide form variance estimation when a single primary sampling unit is selected.

The variance estimator proposed based on the above three methods is given in section 3. Consequently of a properties comparison of three new variance estimators using mathematics is more complicated so that their properties have been investigated using simulation study with carried out on the real blue-wing teal data. The efficiency comparison in terms of relative bias, mean squared error and the relative frequency of coverage population mean of 95% confidence interval that is formed by each variance estimator are presented in section 4. In simulation study section, a primary unit is defined as a set of secondary units being arranged in strip pattern.

2. Design and Methodology

Consider a finite rare population of size NM units, the population consists of N primary sampling units (PSUs) with each PSU contains M secondary sampling units (SSUs). Let (i, j) denote j^{th} SSU in the i^{th} PSU with the associated y-value Y_{ij}, where $i = 1, 2, ..., N$ and $j = 1, 2, ..., M$. Note that the i^{th} PSU in ACS corresponds to i^{th} systematic sample in conventional systematic sampling.

Suppose s^{th} PSU is selected, let $y_j^* = Y_{sj}$, where $j = 1, 2, ..., M$. Let $y_{j.}$ be the total y-value of units within the network associated with y_j^* and x_j be the number of PSUs in the population that are in the network associated with y_j. The draw-by-draw probability of the selection of primary sampling unit associated with network of y_j^* is

x_j/N. A set of number of network associated with the s^{th} PSU is defined as κ_s. For the conventional systematic sampling with the population size is $NM = qnM$; n denote the number of systematic sampled. When $n=1$ systematic sample is chosen, the sampling interval is $q = N$ and the probability of a systematic sample chosen is $1/q = 1/N$. The sample mean $\overline{y}$ that is an estimator ignoring all units adaptively added to the sample is used to estimate the population mean but an unbiased estimator of $Var(\overline{y})$ cannot be obtained. In this case, the variance estimation can be done based on three methods of estimation. First, each systematic sample of size $M = mp$, say, is treated as if it were divided into p sub-samples at random, each of size $m = M/p$ units. Second, the sample is treated as if $m = 2$ units were selected at random from within each of $p = M/2$ equal-sized strata (assume that M is an even number). Finally, the sample is treated as if it were a simple random sample. All of methods give the biased variance estimators [2,3] as follows:

Method I: $M = mp, p \geq 2$ and $m \geq 2$.

$$\hat{v}_{sys(1)}(\overline{y}) = \hat{v}_{sys(1)} = \left(\frac{N-1}{N}\right)\frac{1}{p}\frac{\sum_{t=1}^{p}(\overline{y}_t - \overline{y})^2}{p-1}, \qquad (1)$$

where $\overline{y}_t$ represents the sample mean of the t^{th} sub-sample of size $m = \dfrac{M}{p}$ and $\overline{y}$ represents the systematic sample mean.

Method II: $M = 2p, p \geq 2$ and $m = 2$.

$$\hat{v}_{sys(2)}(\overline{y}) = \hat{v}_{sys(2)} = \left(\frac{N-1}{N}\right)\frac{1}{M}\frac{\sum_{j=1}^{M/2}(y_{2j} - y_{2j-1})^2}{M}, \qquad (2)$$

Method III

$$\hat{v}_{sys(3)}(\overline{y}) = \hat{v}_{sys(3)} = \left(\frac{N-1}{N}\right)\frac{1}{M}\frac{\sum_{j=1}^{M}(y_j - \overline{y})^2}{M-1} \qquad (3)$$

where $\overline{y}$ represents the systematic sample mean.

For ACS, an unbiased estimator of the population mean of interest in this paper is a modified Hansen-Hurwitz (HH) type estimator proposed by Thompson in 1991. The terms of network is considered instead of terms of individual units. When a single PSU is selected, an unbiased estimator of the population mean is of the form [1]

$$\hat{\mu}_{acs} = \frac{1}{M}\sum_{k=1}^{\kappa_i}\frac{y_k}{x_k} , \tag{4}$$

with the variance

$$V\left(\hat{\mu}_{acs}\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\mu}_i - \mu\right)^2 , \tag{5}$$

where i represent the i^{th} PSU that is chosen and only one PSU is chosen.

The problem of variance estimation for ACS will be considered here according to the above three methods. The details have been shown below.

3. The proposed variance estimators

Method I: $p \geq 2$ *and* $m \geq 2$.

Based on method I, the selected PSU or the initial sample is divided into p sub-samples, where $p \geq 2$. Let $\hat{\mu}_t = \frac{1}{m}\sum_{k=1}^{\kappa_t}\frac{y_k}{h_k x_k}$ be the mean of the t^{th} sub-sample, t represent the sub-samples, where $t = 1, 2, .., p$. Let m SSUs and $m \geq 2$ be assigned to the sub-samples at random, h_k be the number of sub-samples in which the k^{th} network appears, κ_t be the number of network associated with the t^{th} sub-sample. An estimator $\hat{\mu}_{acs}$ in equation (4) can be rewritten as

$$\hat{\mu}_{acs} = \frac{1}{p}\sum_{t=1}^{p}\hat{\mu}_t , \tag{6}$$

which is also an unbiased estimator of the population mean.

Since $\hat{\mu}_{acs}$ is the mean of $\hat{\mu}_t$ for simple random sample of size p so the variance estimator based on equation (1) is suggested to be

$$\hat{v}_{acs(1)} = \frac{(1-f)}{p(p-1)} \sum_{t=1}^{p} \left(\hat{\mu}_t - \hat{\mu}_{acs}\right)^2 \ , \tag{7}$$

where $f = 1/N$. This estimator is biased (see appendix A).

Method II: $p \geq 2$ *and* $m = 2$ *.*

Based on method II, the selected PSU of size M is regarding as stratified random sample with equal size. Assume that M is an even integer and no prior information about population. A reasonable choice for allocated units into stratum would be to assign equal sample size to the stratum [1]. So that for stratum h the sample size would be $m_h = \frac{M}{(M/2)} = 2$ SSUs. Let $\hat{\mu}_h = \frac{1}{m_h} \sum_{j=1}^{m_h} u_{hj}$ be the mean of stratum h, $w_h = \frac{m_h}{M}$ be the weight of stratum h and $u_{hj} = \sum_{k \in \kappa_{hj}} \frac{y_k}{v_k x_k}$, where k represent network label, v_k is the number of times that the k^{th} network appears and κ_{hj} is the number of networks associated with the j^{th} SSU in stratum h . An estimator in equation (4) can be rewritten as

$$\hat{\mu}_{acs} = \sum_{h=1}^{p=M/2} w_h \hat{\mu}_h = \frac{1}{M/2} \sum_{h=1}^{p=M/2} \hat{\mu}_h \ , \tag{8}$$

which is also an unbiased estimator of the population mean.

Since $\hat{\mu}_{acs}$ in equation 8 is the mean of $\hat{\mu}_h$ for stratified random sample of size p strata so the variance estimator based on equation (2) is suggested to be

$$\hat{v}_{acs(2)} = \left(\left(\frac{1-f}{M} \right) \sum_{h=1}^{p=M/2} \frac{\left(u_{h1} - u_{h2}\right)^2}{M} \right) \tag{9}$$

where $f = 1/N$. This estimator is biased (see appendix B).

Method III:

For method III, the selected PSU is regarding as simple random sample of size M SSUs. The terms of individual value is considered instead of terms of network. Define m_j^* is the number of unit in the initial sample associated with networks that unit j belongs. The estimator in equation (4) can be rewritten as

$$\hat{\mu}_{acs} = \frac{1}{M}\sum_{j=1}^{M}\frac{y_{j.}}{x_j m_j^*}\ , \tag{10}$$

which is also an unbiased estimator of the population mean.

The variance estimator based on equation (3) is suggested to be

$$\hat{v}_{acs(3)} = \frac{(1-f)}{M^2}\sum_{j=1}^{M}\left(\frac{y_{j.}}{x_j m_j^*} - \hat{\mu}_{acs}\right)^2 \tag{11}$$

where $f = 1/N$. This estimator is biased (see appendix C).

4. Numerical Study

In order to observe the efficiency of the three new variance estimators in the previous section, two numerical studies were carried out. The first study concerns with the artificial population area of 12 meter-square. This population was divided into $NM = 12$ SSUs or plots. Each plot is of size 1.0 meter square. The remaining study, which is more extensive than the first, was based on real blue-winged teal data which was given by Smith, Conroy and Brakhage in 1995 (cited in Salehi and Smith, [5]). This data was study in area of 5000 kilometer-squares of central Florida and was partitioned into $NM = 200$ SSUs with each of size 25 kilometer-squares.

4.1 Illustrative Example

This artificial population was considered here in the case of $N = 3$ PSUs. Details of individual values Y_{ij} and associated networks are shown in table 1. The population mean is μ = (4+3+0+0+2+0+1+5+1+2+6+3)/12 = 2.25 with the population variance σ^2 = 3.69. A single PSU was defined as a set of SSUs being arranged in strip pattern so a single random start is selected from 3x1 squares in the left hand side of table 1 with probability $1/N = 1/3$. All units with the same position are formed the selected PSU or the initial sample. The condition for adding units is defined by $C = \{Y_{ij} \geq 2\}$. For the first-two method of estimation, the case of $p = 2$ sub-samples/ strata was performed and there are $\frac{M!}{(m!)^p \, p!} = \frac{4!}{(2!)^2 \, 2!} = 3$ possible ways to assign units into sub-samples/ stratum at random. The list of all possible samples corresponding to each method is shown in table 2.

Table 1. A population consists of $N = 3$ PSUs and its associated network with the number of PSU corresponded to that network and possible sample mean corresponded to each PSU.

Population; Y_{ij}				**Corresponding network value;** y_k				**Number of PSUs associated with the network;** x_k			
4	3	0	0	9 = 4+3+2	9	0	0	2	2	1	1
2	0	1	5	9	0	1	16 = 5+3+6+2	2	1	1	2
1	2	6	3	1	16	16	16	1	2	2	2

Based on information in table 1, it will be shown how to calculate estimators of the population mean and variance estimators. Suppose that the 2nd PSU with associated y-values {2, 0, 1, 5; 3, 6, 2, 0} is selected as the initial sample, where additional units are presenting after the semicolon. These 4 SSUs are split up at random into 2 sub-samples/ strata. Suppose the first-two SSUs are assigned to the 1st sub-sample/ strata and the remaining SSUs in the initial sample are assigned to the 2nd sub-sample/ stratum. Values of y_k are {9, 0} and {16, 16} for units in the 1st sub-sample/ stratum and the 2nd sub-samples/ stratum, respectively. Notice that the information of edge units is not used in this study. Values of x_k are {2, 1} and {1, 2} for units in the 1st sub-sample/ stratum and the 2nd sub-samples/ stratum, respectively.

Based on method I, values of h_k are {1, 1} and {1, 1} for units in the 1st sub-sample and the 2nd sub-sample.

Thus we get $\hat{\mu}_1 = \frac{1}{2}\left(\frac{9}{1(2)} + \frac{0}{1(1)}\right) = \frac{9}{4}$ and $\hat{\mu}_2 = \frac{1}{2}\left(\frac{1}{1(1)} + \frac{16}{1(2)}\right) = \frac{9}{2}$.

Estimators $\hat{\mu}_{acs}$ in equation 6 and $\hat{v}_{acs(1)}$ are calculated by

$$\hat{\mu}_{acs} = \frac{1}{2}\left(\frac{9}{4} + \frac{9}{2}\right) = \frac{20}{4} = 3.375, \tag{12}$$

and

$$\hat{v}_{acs(1)} = \left(\frac{3-1}{3}\right)\frac{1}{2(2-1)}\left(\left(\frac{9}{4} - 3.375\right)^2 + \left(\frac{9}{2} - 3.375\right)^2\right) = 0.844 \tag{13}$$

For non-adaptive estimators, we get $\bar{y}_1 = \left(\frac{2+0}{2}\right) = 1$ and $\bar{y}_2 = \left(\frac{1+5}{2}\right) = 3$.

Estimators $\bar{y}_{sys}$ and $\hat{v}_{sys(1)}$ are calculated by

$$\bar{y}_{sys} = \left(\frac{1+3}{2}\right) = 2, \tag{14}$$

and

$$\hat{v}_{sys(1)} = \left(\frac{3-1}{3}\right)\frac{1}{2(2-1)}\left((1-2)^2 + (3-2)^2\right) = 0.667 \tag{15}$$

Based on method II, values of v_k are {1, 1} and {1, 1} for units in the 1st stratum and the 2nd stratum. Thus we get $U_{11} = \frac{9}{1(2)} = \frac{9}{2}$, $U_{21} = \frac{0}{1(1)} = 0$ and $U_{12} = \frac{1}{1(1)} = 1$, $U_{22} = \frac{16}{1(2)} = 8$, $\hat{\mu}_1 = \frac{(9+0)}{2} = \frac{9}{2}$ and $\hat{\mu}_2 = \frac{(1+8)}{2} = \frac{9}{2}$.

Estimators $\hat{\mu}_{acs}$ in equation 8 and $\hat{v}_{acs(2)}$ are calculated by

$$\hat{\mu}_{acs} = \frac{1}{2}\left(\frac{9}{4} + \frac{9}{2}\right) = \frac{20}{4} = 3.375, \tag{16}$$

and

$$\hat{v}_{acs(2)} = \left(\frac{3-1}{3}\right)\frac{1}{4^2}\left(\left(\frac{9}{2} - 0\right)^2 + (1-8)^2\right) = 2.885 \tag{17}$$

For non-adaptive estimators, we get $y_1 = 2$, $y_2 = 0$ for the 1st stratum and $y_1 = 1$, $y_2 = 5$ for the 2nd stratum. Thus an estimator $\hat{v}_{sys(2)}$ is calculated from

$$\hat{v}_{sys(2)} = \left(\frac{3-1}{3}\right)\frac{1}{4^2}\left((2-0)^2 + (1-5)^2\right) = 0.833 \tag{18}$$

Based on method III, values of m_j^* are {1, 1, 1, 1}. Thus an estimator $\hat{\mu}_{acs}$ in equation 10 is calculated by

$$\hat{\mu}_{acs} = \frac{1}{4}\left(\frac{9}{2} + \frac{0}{1} + \frac{1}{1} + \frac{16}{2}\right) = \frac{25}{8} = 3.375, \tag{19}$$

and $\hat{v}_{acs(3)}$ is calculated by

$$\hat{v}_{acs(3)} = \left(\frac{3-1}{3}\right)\frac{1}{4(4-1)}\left(\begin{array}{l}\left(\frac{9}{2} - 3.375\right)^2 + (0-3.375)^2 \\ +(1-3.375)^2 + (8-3.375)^2\end{array}\right) \tag{20}$$
$$= 2.205$$

For non-adaptive estimators, we get $y_1 = 2$, $y_2 = 0$ for the 1st stratum and $y_1 = 1$, $y_2 = 5$ for the 2nd stratum. Thus an estimator $\hat{v}_{sys(2)}$ is calculated by

$$\hat{v}_{sys(3)} = \left(\frac{3-1}{3}\right)\frac{1}{4(4-1)}\left((2-2)^2 + (0-2)^2 + (1-2)^2 + (5-2)^2\right) \tag{21}$$
$$= 0.778$$

It can be seen in table 2, possible sample means obtained from each method of estimation as $\hat{\mu}_{acs}$ = 1.125, 3.375 and 2.250 and hence its expectation is 2.25 and the variance is 0.844. For non-adaptive sampling, three possible sample means obtained from each method of estimation as $\overline{y}_{sys}$ is shown in table 2 and its expectation is 2.25. This means that both estimators $\hat{\mu}_{acs}$ and $\overline{y}_{sys}$ are unbiased estimator of μ . Since the initial sample consists of $M = 4$ SSUs, the sub-sample/ strata can take on only one value; that is, $p = 2$ sub-samples/ strata. All of estimators $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$ and $\hat{v}_{acs(3)}$ are biased estimator of $V\left(\hat{\mu}_{acs}\right)$ (see table 2)

Table 2. All possible samples when $N = 3$ PSUs, $p = 2$ and the corresponding estimators $\hat{\mu}_{acs}$, $\bar{y}_{sys}$, $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$, $\hat{v}_{acs(3)}$, $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$ and $\hat{v}_{sys(3)}$.

Possible PSU chosen	p_i	$\hat{\mu}_{acs}$	$\bar{y}_{sys}$	Method I & II						Method III	
				$P(s)$	Sub-sample	$\hat{v}_{acs(1)}$	$\hat{v}_{sys(1)}$	$\hat{v}_{acs(2)}$	$\hat{v}_{sys(2)}$	$\hat{v}_{acs(3)}$	$\hat{v}_{sys(3)}$
(4, 3, 0, 0; 2, 0)	1/3	1.13	1.75	1/9	(4, 3) (0, 0)	0.844	2.042	0.000	0.042	0.281	0.708
				1/9	(4, 0) (3, 0)	0.000	0.042	0.422	1.042		
				1/9	(4, 0) (3, 0)	0.000	0.042	0.422	1.042		
(2, 0, 1, 5; 3, 6, 2, 0)	1/3	3.38	2.00	1/9	(2, 0) (1, 5)	0.844	0.667	2.885	0.833	2.205	0.778
				1/9	(2, 1) (0, 5)	0.260	0.167	3.177	1.083		
				1/9	(2, 5) (0, 1)	5.510	1.500	0.552	0.417		
(1, 2, 6, 3; 5, 1, 0, 0)	1/3	2.25	3.00	1/9	(1, 2) (6, 3)	0.042	1.500	0.116	0.417	0.116	0.778
				1/9	(1, 6) (2, 3)	0.042	0.167	0.116	1.083		
				1/9	(1, 3) (2, 6)	0.042	0.667	0.116	0.833		
Expectation		**2.25**	**2.25**			**0.843**	**0.755**	**0.867**	**0.755**	**0.867**	**0.755**
Relative bias		**0.000**	**0.000**			**-0.001**	**1.587**	**0.024**	**1.587**	**0.024**	**1.587**
MSE		**0.844**	**0.292**			**2.830**	**0.215**	**1.373**	**0.215**	**0.900**	**0.215**

Note: Additional units are presenting after the semicolon in the parenthesis.

4.2 Real Blue-Winged Teal Data

The real blue-winged teal data which was given by Smith, Conroy and Brakhage in 1995 (cited in Salehi and Smith, [5]) was used in this study. Its detail is represented in table 3. This data was divided into 200 secondary sampling units (SSUs). The population mean is μ = 70.605 and the population variance is σ^2 = 451,440.97. These data were considered in two cases of $N = 10$ and $N = 5$ PSUs in the population. For simulation, a PSU was defined as a set of SSUs that being arranged in strip pattern. So that single random start is chosen at random in 10-by-1 squares and in 5-by-1 squares as seen in table 2 when the population consists of $N = 10$ and $N = 5$ PSUs, respectively. The positions repeated throughout the study area and these SSUs form the selected PSU. The pre-condition for adding an unit is defined by $C = \{Y_{ij} \geq 1\}$.

Table3. Blue-winged teal data of size $MN = 200$ SSUs with μ = 70.605.

0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	20	4	2	12	0	0	0	0	0	10	103	0	0	0
0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	150	7144	1	0
0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	6	6339	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	122	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	114	60
0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	3	0

Secondary sampling units (SSUs) within the selected PSU will be divided into sub-samples at random when the first method of estimation is used. For the second method of estimation, the selected PSU will be treated as if it were a stratified sample of size p strata with equal size. Within each stratum, $m = 2$ SSUs are assigned to the stratum at random. The selected PSU will be treated as if it were a simple random sample for the third method of estimation. In this study, there are possible N PSUs can be chosen when the initial sample of size 1 PSU. All possible initial samples were performed. Within the initial sample, each method for the variance estimation was performed r =2,000 times, where j denote the iteration label and $j = 1, 2, ..., r$. Let subscript $(\bullet)$ be replaced with *"acs(1)", "acs(2)", "acs(3)'*, *"sys(1)"*, *"sys(2)"* and *"sys(3)"* for the variance estimators in ACS, and SYS, respectively. Let subscript i denote the selected PSU, $p_i = 1/N$ be the probability of choosing the i^{th} PSU.

The simulation mean and variance of $\hat{\mu}_{(\bullet)}$ was calculated based on the formula

$$E\left(\hat{\mu}_{(\bullet)}\right) = \bar{\mu}_{(\bullet)} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{r}\left(\hat{\mu}_{(\bullet)}\right)_{ij}}{rN}, \tag{22}$$

and

$$V\left(\hat{\mu}_{(\bullet)}\right) = \frac{1}{rN}\sum_{i=1}^{N}\sum_{j=1}^{r}\left(\hat{\mu}_{(\bullet)ij} - \bar{\mu}_{(\bullet)}\right)^2, \tag{23}$$

where $\bar{\mu}_{(\bullet)}$ is as in equation 22.

The simulation mean of $\hat{v}_{(\bullet)}$ was done based on the formula

$$E\left(\hat{v}_{(\bullet)}\right)=\bar{v}_{(\bullet)}=\frac{\sum_{i=1}^{N}\sum_{j=1}^{r}\hat{v}_{(\bullet)ij}}{rN}\ , \tag{24}$$

The relative bias of $\hat{v}_{(\bullet)}$ was calculated based on the formula

$$RB\left(\hat{v}_{(\bullet)}\right)=\frac{E\left(\hat{v}_{(\bullet)}\right)}{V\left(\hat{\mu}_{(\bullet)}\right)}-1\ , \tag{25}$$

where $V\left(\hat{\mu}_{(\bullet)}\right)$ and $E\left(\hat{v}_{(\bullet)}\right)$ are as in equation 23 and 24, respectively.

The mean squared error (MSE) of $\hat{v}_{(\bullet)}$ was approximated based on the formula

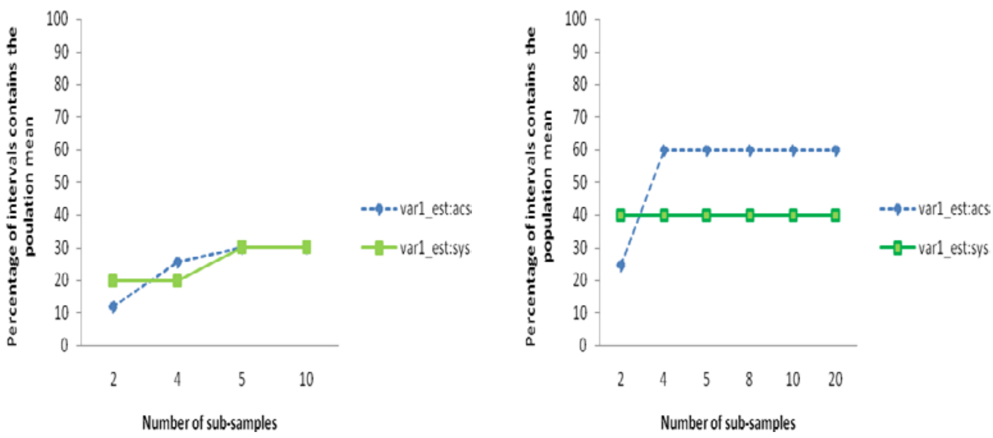
$$MSE\left(\hat{v}_{(\bullet)}\right)=\frac{\sum_{i=1}^{N}\sum_{j=1}^{r}\left(\hat{v}_{(\bullet)ij}-V\left(\hat{\mu}_{(\bullet)}\right)\right)^{2}}{rN} \tag{26}$$

Finally, the relative frequency of coverage μ of the 95% confidence interval that is formed by using each estimator $\hat{v}_{(\bullet)}$ was considered. In this study, it is called as the ability to construct the confidence interval of μ. Define $u_{ij}\left(\hat{v}_{(\bullet)}\right)$ to be 1 if the 95% confidence interval that is produced by $\hat{\mu}_{(\bullet)ij}\pm t_{\alpha/2,df}\sqrt{\hat{v}_{(\bullet)ij}}$ contains μ, and 0 if the 95% confidence interval does not contain μ, where $df=p-1$ for the first-two methods and $df=M-1$ for the 3rd method of estimation. The confidence interval is performed based on a normality assumption. So that the ability to construct the confidence interval of μ was calculated base on the formula

$$CI\left(\hat{v}_{(\bullet)}\right)=100\times\frac{1}{rN}\sum_{i=1}^{N}\sum_{j=1}^{r}u_{ij}\left(\hat{v}_{(\bullet)}\right) \tag{27}$$

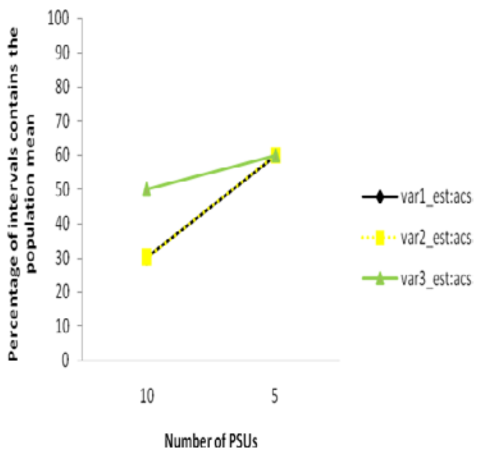
The four possible values of p are performed here when N =10 PSUs, these are, p =2, 4, 5 and 10. When p =2, 4, 5, the estimator $\hat{v}_{acs(2)}$ and $\hat{v}_{sys(2)}$ cannot be obtained because the number of units in each stratum is not 2. This limitation is likely to occur with a population composed of N =5 PSUs and when the values of p are 2, 4, 5, 8 and 10. In these cases, only estimators $\hat{v}_{acs(1)}$ and $\hat{v}_{sys(1)}$ are compared to each other in terms of RB and the ability to construct 95% CI for μ while estimators $\hat{v}_{acs(3)}$ and $\hat{v}_{sys(3)}$ are compared to each other in two cases of N =10 and N =5 PSUs. For N =10 and N =5 PSUs, variance estimator $\hat{v}_{acs(3)}$ is compared to $\hat{v}_{acs(1)}$ and $\hat{v}_{acs(2)}$ that can be obtained from cases of p = 10 and p = 20. In cases of N =10 and N =5 PSUs in the population, the simulation mean of estimators $\hat{\mu}_{acs}$, $\bar{y}_{sys}$ for each value of p are shown in table 4. The results indicate that both estimators $\hat{\mu}_{acs}$ and $\bar{y}_{sys}$ are unbiased estimators of μ for all values of p. When N =10 and p = 2, 4, 5, 8 and 10, as seen in table 5, estimator $\hat{v}_{acs(1)}$ is underestimated while estimator $\hat{v}_{sys(1)}$ is overestimated. In these cases, an estimator $\hat{v}_{sys(1)}$ gives a smaller RB than estimator $\hat{v}_{acs(1)}$. For estimators $\hat{v}_{acs(3)}$ and $\hat{v}_{sys(3)}$, the result show that the RB of estimator $\hat{v}_{acs(3)}$ is higher than the RB of estimator $\hat{v}_{sys(3)}$. Though the RB of estimator $\hat{v}_{acs(1)}$ is too large to 50 percent when p = 2 sub-sample the RB of the estimator $\hat{v}_{acs(1)}$ is slightly decreased when the number of sub-samples increased. In addition, the MSE of both estimators $\hat{v}_{acs(1)}$ and $\hat{v}_{sys(1)}$ is decreased when number of sub-samples increased. When p = 2, the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{sys(1)}$ is 8.2 percent higher than the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{acs(1)}$ (figure 2A). When a value of p is changed from 2 to 4, the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{acs(1)}$ will increase about 13.9 percent (figure 2A). The maximum relative frequency of coverage μ of the 95% confidence interval that of both variance estimators $\hat{v}_{acs(1)}$ and $\hat{v}_{sys(1)}$ is about 30 percent. When N = 5 and p = 4, 5, 8, 10 and 20, the results in table 6 show that estimator $\hat{v}_{acs(1)}$ is better than estimator $\hat{v}_{sys(1)}$ in terms of minimum RB. Its RB will decrease when the number of sub-samples increases. The estimator $\hat{v}_{acs(1)}$ is 15.74 percent underestimated when p =20 sub-samples. In addition, the MSE of both estimators $\hat{v}_{acs(1)}$ and $\hat{v}_{sys(1)}$ will decrease when the sub-samples increase. When p = 2, the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{sys(1)}$ is 15.5 percent higher than the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{acs(1)}$ (figure 2B). However, the relative frequency of coverage μ of the 95% confidence interval that is formed by using variance estimator $\hat{v}_{acs(1)}$ will increase about 35.5 percent when a value of p is changed from 2 to 4, (figure 2B). Though the maximum relative frequency of coverage μ of the 95% confidence interval that of variance estimator $\hat{v}_{acs(1)}$ is higher than of variance estimator $\hat{v}_{sys(1)}$ its relative frequency of coverage μ of the 95% confidence interval is less than 0.95 or 95 percent.

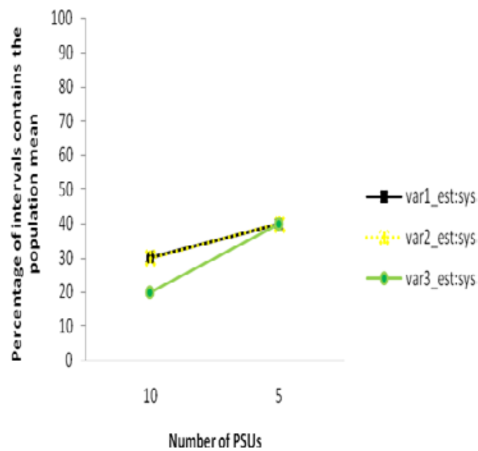
When all of three estimators $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$ and $\hat{v}_{acs(3)}$ are compared to each other, the result in table 5 show that an estimator $\hat{v}_{acs(3)}$ seems to be a good choice in terms of minimum RB and MSE in case of N =10 PSUs. For N = 5 PSUs and p = 20, as seen in table 6, an estimator $\hat{v}_{acs(2)}$ gives the smallest RB than the others. In cases of N =10 PSUs, p = 20 and N =5 PSUs, p = 20, estimators $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$ and $\hat{v}_{acs(3)}$ are better than estimators $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$ and $\hat{v}_{sys(3)}$ because they have the smaller RB and MSE than non-adaptive estimators $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$ and $\hat{v}_{sys(3)}$. For a comparison in terms of relative frequency of coverage μ of the 95% confidence interval that is formed by using estimators $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$ and $\hat{v}_{acs(3)}$, the result in figure 3A show that they have the same efficiency in terms of relative frequency of coverage μ. The maximum relative frequency of coverage μ of the 95% confidence interval of these three variance estimators is 0.6 or 60 percent while the maximum relative frequency of coverage μ of the 95% confidence interval of non-adaptive estimators is 0.4 or 40 percent (figure 3A and figure 3B). However, their relative frequencies of coverage μ are less than 0.95 or 95 percent.



A: when $N = 10$ **B:** when $N = 5$

Figure 2. Percent of relative frequency coverage μ of the 95% CI that is formed with $\hat{v}_{acs(1)}$ and $\hat{v}_{sys(1)}$ when $N = 10$, $N = 5$ PSUs*.





A **B**

Figure 3. Percent of relative frequency coverage μ of the 95% CI that is formed with $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$, $\hat{v}_{acs(3)}$ and $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$, $\hat{v}_{sys(3)}$ when $N=10$, $N=5$ PSUs*.

Note: * var1_est:acs= $\hat{v}_{acs(1)}$, var1_est:sys= $\hat{v}_{sys(1)}$, var2_est:acs= $\hat{v}_{acs(2)}$, var2_est:sys= $\hat{v}_{sys(2)}$, var3_est:acs = $\hat{v}_{acs(3)}$ and var3_est:sys = $\hat{v}_{sys(3)}$.

Table 4. The simulation mean and variance of $\hat{\mu}_{acs}$ and $\overline{y}_{sys}$ when the real blue-winged teal population consists of $N=5$ and $N=10$ PSUs.

$N=5$ & $M=40$				$N=10$ & $M=20$			
p	Possible PSU chosen	$\hat{\mu}_{acs}$	$\overline{y}_{sys}$	p	Possible PSU chosen	$\hat{\mu}_{acs}$	$\overline{y}_{sys}$
	1	114.833	158.849		1	0.25	0.25
	2	0.075	0.075		2	0.15	0.15
2	3	2.608	3.399		3	0.00	0.00
	4	118.167	8.125		4	231.12	7.55
	5	117.342	182.575	2	5	229.37	364.90
Mean		**70.605**	**70.605**		6	229.42	317.45
Variance		**3201.187**	**6746.090**		7	0.00	0.00
	1	114.833	158.849		8	5.22	6.80
	2	0.075	0.075		9	5.22	8.70
4	3	2.608	3.399		10	5.32	0.25
	4	118.167	8.125	**Mean**		**70.605**	**70.605**
	5	117.342	182.575	**Variance**		**10894.258**	**20514.016**
Mean		**70.605**	**70.605**	4	1	0.25	0.25
Variance		**3201.187**	**6746.090**		2	0.15	0.15

Table 4. (continued)

$N=5$ & $M=40$			
p	Possible PSU chosen	$\hat{\mu}_{acs}$	$\overline{y}_{sys}$
5	1	114.833	158.849
	2	0.075	0.075
	3	2.608	3.399
	4	118.167	8.125
	5	117.342	182.575
Mean		**70.605**	**70.605**
Variance		**3201.187**	**6746.090**
8	1	114.833	158.849
	2	0.075	0.075
	3	2.608	3.399
	4	118.167	8.125
	5	117.342	182.575
Mean		**70.605**	**70.605**
Variance		**3201.187**	**6746.090**
10	1	114.833	158.849
	2	0.075	0.075
	3	2.608	3.399
	4	118.167	8.125
	5	117.342	182.575
Mean		**70.605**	**70.605**
Variance		**3201.187**	**6746.090**
20	1	114.833	158.849
	2	0.075	0.075
	3	2.608	3.399
	4	118.167	8.125
	5	117.342	182.575
Mean		**70.605**	**70.605**
Variance		**3201.187**	**6746.090**

$N=10$ & $M=20$			
p	Possible PSU chosen	$\hat{\mu}_{acs}$	$\overline{y}_{sys}$
	3	0.00	0.00
	4	231.12	7.55
	5	229.37	364.90
	6	229.42	317.45
	7	0.00	0.00
	8	5.22	6.80
	9	5.22	8.70
	10	5.32	0.25
Mean		**70.605**	**70.605**
Variance		**10894.258**	**20514.016**
5	1	0.25	0.25
	2	0.15	0.15
	3	0.00	0.00
	4	231.12	7.55
	5	229.37	364.90
	6	229.42	317.45
	7	0.00	0.00
	8	5.22	6.80
	9	5.22	8.70
	10	5.32	0.25
Mean		**70.605**	**70.605**
Variance		**10894.258**	**20514.016**
10	1	0.25	0.25
	2	0.15	0.15
	3	0.00	0.00
	4	231.12	7.55
	5	229.37	364.90
	6	229.42	317.45
	7	0.00	0.00
	8	5.22	6.80
	9	5.22	8.70
	10	5.32	0.25
Mean		**70.605**	**70.605**
Variance		**10894.258**	**20514.016**

Note: mean refers to the simulation mean of estimators $\hat{\mu}_{acs}$ and $\overline{y}_{sys}$.

Table 5. The simulation mean, MSE and RB of $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$, $\hat{v}_{acs(3)}$, $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$ and $\hat{v}_{sys(3)}$ when the real blue-winged teal population consists of $N = 10$ PSUs.

p	Possible PSU chosen	$\hat{v}_{acs(1)}$	$\hat{v}_{acs(2)}$	$\hat{v}_{sys(1)}$	$\hat{v}_{sys(2)}$
2	1	0.056	na	0.056	na
	2	0.020	na	0.020	na
	3	0.000	na	0.000	na
	4	22249.799	na	24.130	na
	5	10966.841	na	114683.326	na
	6	21369.729	na	90388.902	na
	7	0.000	na	0.000	na
	8	12.087	na	33.830	na
	9	11.769	na	36.141	na
	10	24.406	na	0.027	na
Mean		**5463.47**	**-**	**20516.64**	**-**
MSE		**189161255.79**	**-**	**104384950.99**	**-**
RB		**-0.4895**	**-**	**0.1133**	**-**
4	1	0.056	na	0.056	na
	2	0.020	na	0.020	na
	3	0.000	na	0.000	na
	4	22018.885	na	24.079	na
	5	12548.469	na	114648.258	na
	6	22298.869	na	90395.348	na
	7	0.000	na	0.000	na
	8	11.960	na	33.799	na
	9	11.552	na	31.251	na
	10	24.454	na	0.028	na
Mean		**5691.43**	**-**	**20513.28**	**-**
MSE		**72789251.26**	**-**	**182680129.60**	**-**
RB		**-0.4776**	**-**	**0.1128**	**-**
5	1	0.056	na	0.056	na
	2	0.0004	na	0.020	na
	3	0.000	na	0.000	na
	4	22904.927	na	24.249	na
	5	12789.197	na	114625.218	na
	6	22529.469	na	90396.829	na
	7	0.000	na	0.000	na
	8	11.726	na	33.110	na
	9	11.756	na	36.110	na
	10	24.455	na	0.028	na
Mean		**5827.159**	**-**	**20515.279**	**-**
MSE		**61085140.86**	**-**	**182346863.60**	**-**
RB		**-0.4651**	**-**	**0.1126**	**-**

Possible PSU chosen	$\hat{v}_{acs(3)}$	$\hat{v}_{sys(3)}$
1	0.056	0.056
2	0.020	0.020
3	0.000	0.000
4	22358.176	23.864
5	14099.680	114307.233
6	24882.090	90157.975
7	0.001	0.000
8	12.891	33.416

Table 5. (continued)

p	Possible PSU chosen	$\hat{v}_{acs(1)}$	$\hat{v}_{acs(2)}$	$\hat{v}_{sys(1)}$	$\hat{v}_{sys(2)}$	Possible PSU chosen	$\hat{v}_{acs(3)}$	$\hat{v}_{sys(3)}$
	1	0.056	0.056	0.056	0.056			
	2	0.0004	0.020	0.020	0.020			
	3	0.000	0.000	0.000	0.000			
	4	22529.043	22904.507	24.058	24.276	9	2428.850	29.488
10	5	13661.817	14915.501	114660.485	114987.214			
	6	22561.815	22941.246	90397.814	90395.605			
	7	0.000	0.000	0.000	0.000			
	8	11.599	11.925	33.524	33.526	10	2402.271	0.028
	9	11.634	11.909	35.802	35.648			
	10	24.459	24.445	0.028	0.028			
Mean		**5880.04**	**6080.961**	**20515.18**	**20547.637**	**Mean**	**6618.404**	**20455.208**
MSE		**39048720.75**	**22962500.02**	**180760072.03**	**203445416.20**	**MSE**	**107427444.47**	**1701048566.61**
RB		**-0.4603**	**-0.442**	**0.1129**	**0.1146**	**RB**	**-0.392**	**0.1128**

Note: "na" refers to "is not available and **"**mean" refers to the simulation mean of estimators.

Table 6. The simulation mean, MSE and RB of $\hat{v}_{acs(1)}$, $\hat{v}_{acs(2)}$, $\hat{v}_{acs(3)}$, $\hat{v}_{sys(1)}$, $\hat{v}_{sys(2)}$ and $\hat{v}_{sys(3)}$ when the real blue-winged teal population consists of $N = 5$ PSUs.

p	**Possible PSU chosen**	$\hat{v}_{acs(1)}$	$\hat{v}_{acs(2)}$	$\hat{v}_{sys(1)}$	$\hat{v}_{sys(2)}$	**Possible PSU chosen**	$\hat{v}_{acs(3)}$	$\hat{v}_{sys(3)}$
	1	5179.359	na	20089.996	na			
	2	0.0045	na	0.0045	na			
2	3	2.656	na	7.499	na			
	4	4880.370	na	13.093	na	1	5118.267	20063.175
	5	2499.187	na	25516.48	na			
Mean		**2512.315**	-	**9125.47**	-			
MSE		**27700824.09**	-	**76195014.41**	-			
RB		**-0.2080**	-	**0.3525**	-			
	1	5168.986	na	20088.592	na			
	2	0.0045	na	0.0045	na			
4	3	2.685	na	7.517	na			
	4	5047.413	na	12.789	na	2	0.0045	0.0045
	5	2950.213	na	25517.887	na			
Mean		**2633.86**	-	**9125.358**	-			
MSE		**13049298.50**	-	**76047866.73**	-			
RB		**-0.1873**	-	**0.3525**	-			
	1	5137.788	na	20088.993	na			
	2	0.0045	na	0.0045	na			
5	3	2.653	na	7.497	na			
	4	5046.251	na	12.543	na			
	5	3012.559	na	25526.513	na	3	2.652	7.484
Mean		**2639.85**	-	**9127.11**	-			
MSE		**16006235.98**	-	**76071427.63**	-			
RB		**-0.1838**	-	**0.3527**	-			

Table 6. (continued)

p	Possible PSU chosen	$\hat{v}_{acs(1)}$	$\hat{v}_{acs(2)}$	$\hat{v}_{sys(1)}$	$\hat{v}_{sys(2)}$	Possible PSU chosen	$\hat{v}_{acs(3)}$	$\hat{v}_{sys(3)}$
	1	5108.632	na	20088.537	na			
	2	0.0045	na	0.0045	na			
8	3	2.647	na	7.493	na			
	4	5129.201	na	13.158	na	4	5105.295	11.524
	5	3082.797	na	25510.203	na			
Mean		**2664.66**	**-**	**9123.88**	-			
MSE		**14556932.43**	**-**	**72293301.54**	-			
RB		**-0.1656**	**-**	**0.3522**	-			
10	1	5136.152	na	20089.014	na			
	2	0.0045	na	0.0045	na			
	3	2.673	na	7.509	na			
	4	5082.953	na	13.005	na			
	5	3141.166	na	25495.055	na			
Mean		**2672.59**	**-**	**9120.92**	-			
MSE		**14136331.78**	**-**	**75870180.76**	**-**	5	3315.644	25466.280
RB		**-0.1649**	**-**	**0.3518**	**-**			
20	1	5117.414	5186.067	20088.811	20089.167			
	2	0.0045	0.045	0.0045	0.0045			
	3	2.647	2.689	7.491	7.499			
	4	5139.842	5139.842	12.994	12.899			
	5	3267.939	3405.324	25498.518	25554.029			
Mean		**2705.57**	**2746.793**	**9121.563**	**9132.719**	**Mean**	**2708.373**	**9109.694**
MSE		**13503941.49**	**14658764.36**	**72205928.98**	**124205590.00**	**MSE**	**5557374.33**	**207947126.60**
RB		**-0.1574**	**-0.1419**	**0.3519**	**0.3536**	**RB**	**-0.1537**	**0.3504**

Note: “na” refers to “is not available and **“**mean” refers to the simulation mean of estimators.

5. Conclusion and Discussion

The numerical study shows that both estimators $\hat{\mu}_{acs}$ and $\bar{y}_{sys}$ were unbiased estimator of the population mean when a single primary sampling unit is chosen. Three new bias variance estimators, based on splitting the initial sample into sub-samples, $\hat{v}_{acs(1)}$, regarding the initial sample as a stratified sample, $\hat{v}_{acs(2)}$**,** and simple random sample, $\hat{v}_{acs(3)}$**,** were proposed. The numerical study indicated that the first estimator was not preferable for all cases of p when $N = 10$ primary sampling units, especially p =2, because its relative bias was too large to be useful. In addition, it was not a good choice in terms of relative bias when it was compared with non-adaptive variance estimator for systematic sampling. Increasing the number of sub-samples made its relative bias and MSE decrease. Similar to the stratified random sampling estimator,

$\hat{v}_{acs(2)}$, because its relative bias was too large when $p = 10$ and $N = 10$. The relative bias and MSE of the simple random sampling estimator, $\hat{v}_{acs(3)}$, may be reduced when a size of primary sampling unit increased. When all three new variance estimators were compared to each other, estimator $\hat{v}_{acs(3)}$ seems to be a good choice in terms of relative bias in case of $N = 10$ primary sampling units with each of size $M = 20$ secondary sampling units. The estimators $\hat{v}_{acs(2)}$ which obtained in case of $p = 20$, $N = 5$ gave the smallest relative bias and MSE while the relative bias of estimators $\hat{v}_{acs(3)}$ and $\hat{v}_{acs(1)}$ were same. The ability to construct confidence intervals containing the population mean of all three new estimators though was less than 95 percent their ability to construct intervals were higher than the ability to construct intervals of non-adaptive estimators. However, all three new variance estimators were not used any information of edge units. Similar to the correlation terms between units within PSU or between sub-samples was not considered in this study. These points may reduce the bias of variance estimator and needed to study in the future. Other variance estimators under various conditions for examples the population with random effects etc, are of interest for further study. In addition, the development of the variance estimator based on other issues, such as the “modified Hajek-variance estimator” [6] and the “partially systematic sampling” [7,8], might be considered.

Acknowledgements

I would like to thank and offer my appreciation to Assistant Prof. Dr. Arthur L. Dryver for his guidance and encouragement. In addition, I would like to thank Dr. Nipaporn and Mr. Paveen Chomanon who advise me on issue of computer programming.

References

[1] Thompson, S. K., Adaptive cluster sampling: Designs with primary and secondary units, *Biometrics*, 1991; 47: 1103-1105.

[2] Cochran, W. G. *Sampling Techniques*, 3rd edition, Singapore: Wiley, 1977.

[3] Wolter, K. M., An Investigation of Some Estimators of Variance for Systematic Sampling, *J. Am. Stat. Assoc*. 1984; 79: 781-790.

[4] Koop, J. C., Short Communications on Splitting a Systematic Sample for Variance Estimation, *Ann. Math Stat*, 1971; 42: 1084-1087.

[5] Salehi, M. M. and Smith, R. D., Two-Stage Sequential Sampling: A Neighborhood-Free Adaptive Sampling Procedure, *J. Agri., Bio., and Envi. Stat*., 2005; 10: 84-103.

[6] Berger, G. Y., A Modified Hajek Variance Estimator for Systematic Sampling*, Stat Tran*, 2003; 6: 5-21.

[7] Zinger, A., Variance Estimation in Partially Systematic Sampling, *J. Am. Stat. Assoc*., 1980; 75: 206-211.

[8] Leu and Tsui., New Partially Systematic Sampling, *Stat. Sinica*, 1996; 6: 617-630.

Appendix A: The expected value of variance estimator $\hat{v}_{acs(1)}$

Suppose that an estimator $\hat{v}_{acs(1)} = (1-f)\frac{1}{p}\frac{\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)} - \hat{\mu}_{acs}\right)^2}{p-1}$,

where $f = 1/N$ and i is fixed for the i[th] PSU chosen, is an unbiased estimator of $V(\hat{\mu}_{acs})$ in equation 4. Thus

$$E\left(\hat{v}_{acs(1)}\right) = E\left\{(1-f)\frac{1}{p}\frac{\sum_{t=1}^{p}(\hat{\mu}_t - \hat{\mu}_{acs})^2}{p-1}\right\} = V(\hat{\mu}_{acs})$$

Consider in terms of $\hat{\mu}_{t(i)} - \mu = \left(\hat{\mu}_{t(i)} - \hat{\mu}\right) - (\mu - \hat{\mu})$

$$\therefore \sum_{t=1}^{p}\left(\hat{\mu}_{t(i)} - \mu\right)^2 = \sum_{t=1}^{p}\left(\left(\hat{\mu}_{t(i)} - \hat{\mu}\right) - (\mu - \hat{\mu})\right)^2$$

Since $\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)} - \hat{\mu}\right) = 0$ and $(\mu - \hat{\mu})^2 = (\hat{\mu} - \mu)^2$ so we get

$$\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)} - \mu\right)^2 = \sum_{t=1}^{p}\left(\hat{\mu}_{t(i)} - \hat{\mu}\right)^2 + p(\hat{\mu} - \mu)^2$$

$$=\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\hat{\mu}\right)^2+p\left(\left(\sum_{t=1}^{p}\hat{\mu}_{t(i)}\Big/p\right)-\mu\right)^2$$

$$\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\hat{\mu}\right)^2=\left(1-\frac{1}{p}\right)\left(\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\right)^2$$
$$-\sum_{t=1}^{p}\sum_{t'\neq t}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\left(\hat{\mu}_{t'(i)}-\mu\right)$$

Multiply by $\dfrac{(1-f)}{p(p-1)}$, then we get

$$\frac{(1-f)}{p(p-1)}\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\hat{\mu}\right)^2=\frac{(1-f)}{p^2}\left(\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\right)^2$$
$$-\frac{(1-f)}{p(p-1)}\sum_{t=1}^{p}\sum_{t'\neq t}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\left(\hat{\mu}_{t'(i)}-\mu\right)$$

Since $\hat{v}_{acs(1)}(\hat{\mu})=\dfrac{(1-f)}{p}\dfrac{\sum_{t=1}^{p}(\hat{\mu}_t-\hat{\mu})^2}{p-1}$ so the expectation of $\hat{v}_{acs(1)}(\hat{\mu})$ is

$$E\left(\hat{v}_{acs(1)}(\hat{\mu})\right)=E\left\{\begin{array}{l}\dfrac{(1-f)}{p^2}\left(\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\right)^2\\ -\dfrac{(1-f)}{p(p-1)}\left(\sum_{t=1}^{p}\sum_{t'\neq t}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\left(\hat{\mu}_{t'(i)}-\mu\right)\right)\end{array}\right\}$$

Based on method I, there are possible $r=\dfrac{M!}{(m!)^p\,p!}$ sub-samples. Let j denote the number of possible sub-samples, where $j=1,2,\ldots,r$. Thus

$$E\left(\hat{v}_{acs(1)}(\hat{\mu})\right)=\frac{(1-f)}{Np^2}\sum_{i=1}^{N}\left(\sum_{t=1}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\right)^2$$
$$-\frac{(1-f)}{Nrp(p-1)}\sum_{i=1}^{N}\left(\sum_{j=1}^{r}\sum_{t=1}^{p}\sum_{t'\neq t}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\left(\hat{\mu}_{t'(i)}-\mu\right)\right)$$

$$E\left(\hat{v}_{acs(1)}\left(\hat{\mu}\right)\right)=\left(1-f\right)\left\{V\left(\hat{\mu}\right)-\frac{\sum_{i=1}^{N}\left(\sum_{j=1}^{r}\sum_{t=1}^{p}\sum_{t'\neq t}^{p}\left(\hat{\mu}_{t(i)}-\mu\right)\left(\hat{\mu}_{t'(i)}-\mu\right)\right)}{Nrp\left(p-1\right)}\right\} \quad \text{(A1)}$$

$\neq V\left(\hat{\mu}\right)$ which is contradiction.

This means that estimator $\hat{v}_{acs(1)}$ is a biased estimator of $V\left(\hat{\mu}_{acs}\right)$.

Appendix B: The expectation the value of the variance estimator $\hat{v}_{acs(2)}$

Suppose that an estimator $\hat{v}_{acs(2)}=\left(1-f\right)\frac{\sum_{t=1}^{p}\left(u_{1t}-u_{2t}\right)^{2}}{M^{2}}$, where $f=1/N$, is an unbiased estimator of $V\left(\hat{\mu}_{acs}\right)$ in equation 4. Thus

$$E\left(\hat{v}_{acs(2)}\right)=E\left\{\left(1-f\right)\frac{\sum_{h=1}^{p}\left(u_{h1}-u_{h2}\right)^{2}}{M^{2}}\right\}=V\left(\hat{\mu}_{acs}\right)$$

Consider the terms of $\left(\hat{\mu}_{acs}-\mu\right)$ it can be rewritten as

$$\left(\hat{\mu}_{acs}-\mu\right)=\frac{1}{p}\sum_{h=1}^{p}\left(\hat{\mu}_{h(i)}-\mu\right),$$

where i is fixed for the i^th PSU chosen.

$$\left(\hat{\mu}_{acs}-\mu\right)^{2}=\left\{\frac{1}{p}\sum_{h=1}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\right\}^{2}$$

$$=\frac{1}{p^{2}}\left\{\sum_{h=1}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)^{2}+\sum_{h=1}^{p}\sum_{h'\neq t}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\left(\hat{\mu}_{h'(i)}-\mu\right)\right\}$$

Since $\hat{\mu}_{h}=\frac{\left(u_{h1}+u_{h2}\right)}{2}$ so the terms of $\sum_{h=1}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)^{2}$ can be rewritten as

$$\sum_{h=1}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)^2=\sum_{h=1}^{p}\left(\frac{\left(u_{h1}+u_{h2}\right)-2\mu}{2}\right)^2$$

$$=\sum_{h=1}^{p}\left(\frac{\left(u_{h1}+u_{h2}\right)^2-2\left(2\mu\right)\left(u_{h1}+u_{h2}\right)+4\mu^2}{4}\right)$$

$$=\sum_{h=1}^{p}\left(\frac{\left[\left(u_{h1}-u_{h2}\right)^2+4u_{h1}u_{h2}\right]-4\mu\left(u_{h1}+u_{h2}\right)+4\mu^2}{4}\right)$$

$$=\sum_{h=1}^{p}\frac{\left(u_{h1}-u_{h2}\right)^2}{4}+\sum_{h=1}^{p}\frac{\left(u_{h1}-\mu\right)\left(u_{h2}-\mu\right)}{4}$$

$$\therefore\left(\hat{\mu}_{acs}-\mu\right)^2=\frac{1}{\left(M/2\right)^2}\left\{\sum_{h=1}^{p}\frac{\left(u_{h1}-u_{h2}\right)^2}{4}+\sum_{t=1}^{p}\frac{\left(u_{h1}-\mu\right)\left(u_{h2}-\mu\right)}{4}\right\}$$
$$+\frac{1}{\left(M/2\right)^2}\left\{\sum_{h=1}^{p}\sum_{h'\neq h}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\left(\hat{\mu}_{h'(i)}-\mu\right)\right\}$$

So that we get

$$\frac{1}{M^2}\sum_{h=1}^{p}\left(u_{h1}-u_{h2}\right)^2=\left(\hat{\mu}_{acs}-\mu\right)^2-\frac{1}{M^2}\sum_{h=1}^{p}\left(u_{h1}-\mu\right)\left(u_{h2}-\mu\right)$$
$$-\frac{4}{M^2}\left\{\sum_{h=1}^{p}\sum_{h'\neq h}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\left(\hat{\mu}_{h(i)}-\mu\right)\right\}$$

Multiply by $\left(1-f\right)$. Since $\hat{v}_{acs(2)}=\left(1-f\right)\frac{\sum_{h=1}^{p}\left(u_{h1}-u_{h2}\right)^2}{M^2}$ so the expectation of $\hat{v}_{acs(2)}$ is

$$E\left(\frac{(1-f)}{M^2}\sum_{h=1}^{p}(u_{h1}-u_{h2})^2\right)=(1-f)E(\hat{\mu}_{acs}-\mu)^2$$
$$-E\left(\frac{(1-f)}{M^2}\sum_{h=1}^{p}(u_{h1}-\mu)(u_{h2}-\mu)\right)$$
$$-E\left(\frac{4(1-f)}{M^2}\left\{\sum_{h=1}^{p}\sum_{h'\neq h}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\left(\hat{\mu}_{h'(i)}-\mu\right)\right\}\right)$$

Based on method II, there are possible $r=\dfrac{M!}{(2!)^p\,p!}$ strata. Let j represent the number of possible strata, where $j=1,2,...,r$.

$$E\left(\hat{v}_{acs(2)}\right)=(1-f)\left\{\begin{array}{l}V(\hat{\mu}_{acs})-\dfrac{1}{rNM^2}\sum_{i=1}^{N}\left(\sum_{j=1}^{r}\sum_{h=1}^{p}(u_{1t}-\mu)(u_{2t}-\mu)\right)\\-\dfrac{4}{rNM^2}\sum_{i=1}^{N}\left(\sum_{j=1}^{r}\left\{\sum_{h=1}^{p}\sum_{h'\neq h}^{p}\left(\hat{\mu}_{h(i)}-\mu\right)\left(\hat{\mu}_{h'(i)}-\mu\right)\right\}\right)\end{array}\right\} \quad \text{(B1)}$$

$\neq V(\hat{\mu}_{acs})$ which is contradiction.

This means that estimator $\hat{v}_{acs(2)}$ is a biased estimator of $V(\hat{\mu}_{acs})$.

Appendix C: The expected value of variance estimator $\hat{v}_{acs(3)}$

Suppose that an estimator $\hat{v}_{acs(3)}=(1-f)\dfrac{\sum_{j=1}^{M}\left(\dfrac{y_j}{x_j m_j^*}-\hat{\mu}_{acs}\right)^2}{M(M-1)}$,

where $f=1/N$, is an unbiased estimator of $V(\hat{\mu}_{acs})$ in equation 4. Thus

$$E\left(\hat{v}_{acs(3)}\right)=E\left\{\frac{(1-f)}{M(M-1)}\sum_{j=1}^{M}\left(\frac{y_j}{x_j m_j^*}-\hat{\mu}_{acs}\right)^2\right\}=V(\hat{\mu}_{acs})$$

Let $Z_j=\dfrac{y_j}{x_j m_j^*}$ and the terms of $(\hat{\mu}_{acs}-\mu)$ can be rewritten as

$$(\hat{\mu}_{acs}-\mu)=\frac{1}{M}\sum_{j=1}^{M}\left((Z_j-\hat{\mu}_{acs})+(\hat{\mu}_{acs}-\mu)\right)$$

$$\left(\hat{\mu}_{acs}-\mu\right)^2=\frac{1}{M^2}\left\{\begin{array}{l}\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)^2+2\left(\hat{\mu}_{acs}-\mu\right)\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)\\+\left(\hat{\mu}_{acs}-\mu\right)^2+\sum_{j=1}^{M}\sum_{j'=1}^{M}\left(Z_j-\mu\right)\left(Z_{j'}-\mu\right)\end{array}\right\}$$

Since $\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)=\sum_{j=1}^{M}Z_j-M\hat{\mu}_{acs}=0$ so $2\left(\hat{\mu}_{acs}-\mu\right)\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)=0$. Thus we get

$$\left(1-\frac{1}{M}\right)\left(\hat{\mu}_{acs}-\mu\right)^2=\frac{1}{M^2}\left\{\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)^2+\sum_{j=1}^{M}\sum_{j'=1}^{M}\left(Z_j-\mu\right)\left(Z_{j'}-\mu\right)\right\}$$

$$\left(\hat{\mu}_{acs}-\mu\right)^2=\frac{1}{M\left(M-1\right)}\left\{\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)^2+\sum_{j=1}^{M}\sum_{j'=1}^{M}\left(Z_j-\mu\right)\left(Z_{j'}-\mu\right)\right\}$$

Multiply by $\left(1-f\right)$, then we get

$$\frac{\left(1-f\right)}{M\left(M-1\right)}\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)^2=\left(1-f\right)\left\{\left(\hat{\mu}_{acs}-\mu\right)^2-\frac{\sum_{j=1}^{M}\sum_{j'=1}^{M}\left(Z_j-\mu\right)\left(Z_{j'}-\mu\right)}{M\left(M-1\right)}\right\}$$

Since $\hat{v}_{acs(3)}=\frac{\left(1-f\right)}{\left(M-1\right)M}\sum_{j=1}^{M}\left(Z_j-\hat{\mu}_{acs}\right)^2$ so the expectation of $\hat{v}_{acs(3)}$ is

$$E\left(\hat{v}_{acs(3)}\right)=\left(1-f\right)\left\{V\left(\hat{\mu}_{acs}\right)-\frac{\sum_{i=1}^{N}\left(\sum_{j=1}^{M}\sum_{j'=1}^{M}\left(Z_j-\mu\right)\left(Z_{j'}-\mu\right)\right)}{NM\left(M-1\right)}\right\} \quad \text{(C1)}$$

$$\neq V\left(\hat{\mu}_{acs}\right)$$ which is contradiction.

This means that estimator $\hat{v}_{acs(3)}$ is a biased estimator of $V\left(\hat{\mu}_{acs}\right)$