A special case of stratified samples is considered where each stratum has the same number of units and from each stratum, one unit is selected in the sample with simple random sampling (SRS). The usual SRS estimator for the variance of a mean is biased under this design, and the size of the bias is estimated in this paper. The problem is related to systematic sampling.

KEYWORDS: Simple random sampling; Stratified samples; Variance estimation.

In this study a special kind of stratified samples and variance estimators is analysed. The paper is organized as follows: section 1 summarises the notations used through the whole paper. Section 2 outlines the problem, while section 3 presents different forms of the variance of the population. In sections 4 and 5 the expectations of the sample variance and the variance estimator are developed in the frame of the investigated sampling design.

1. NOTATIONS

The notations in this paper are borrowed basically from Cochran (1977). In particular,

- $U = \{1, 2, 3, \ldots, N\}$ is a finite universe,
- $s = \{1, 2, 3, \ldots, n\}$ is a sample from $U$,
- $\bar{y}$ is the population mean of a study variable $y$,
- $S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})^2$ is the variance of $y$ in $U$,
- $\bar{y} = \frac{y_1 + y_2 + \ldots + y_n}{n}$ is the estimate of $\bar{y}$ under simple random sampling (SRS),
- $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$ is the sample estimate of $S^2$ under SRS (care will be taken to prevent misinterpreting the square root of this for the notation of a sample),

1 Head of Section of the Hungarian Central Statistical Office.
\[ V(\overline{y}) = \frac{N-n}{N} S^2 = \frac{(1-f) S^2}{n} \]

is the variance of \( \overline{y} \), where \( f = n/N \),

\[ \nu(\overline{y}) = \frac{(1-f) S^2}{n} \]

is the sample estimate for \( V(\overline{y}) \),

\[ \overline{y}_{st} = \sum_{h=1}^{L} W_h \overline{y}_h \]

is the estimate of \( \overline{y} \) from a stratified sample where the \( W_h \)s and the \( \overline{y}_h \)s are stratum weights and estimated stratum means, respectively,

\[ V(\overline{y}_{st}) = \sum_{h=1}^{L} (1-f_h) W_h^2 \frac{S_h^2}{n_h} \]

is the variance of \( \overline{y}_{st} \) where \( n_h, f_h = n_h/N_h \) and \( S_h^2 \) are sample size, sampling fraction and population variance for stratum \( h \), respectively,

\[ \overline{y}_{sy} = \frac{y_1 + y_1 + k + y_1 + 2k + \ldots + y_1 + (n-1)k}{n} \]

is the estimate of \( \overline{y} \) from the \( i^{th} \) systematic sample provided \( N = nk, 1 \leq i \leq k \),

\[ V(\overline{y}_{sy}) = \frac{N-1}{N} S^2 - \frac{1}{N} \sum_{i=1}^{n} (y_i - \overline{y})^2 \]

is the variance of \( \overline{y}_{sy} \) where \( y_1 = y_i \), \( y_{i+1} = y_{1+k} \), \ldots, \( y_{in} = y_{i+(n-1)k} \), and \( \overline{y}_i \) is the mean of \( i^{th} \) sample.

The relation \( N = nk \) is supposed to hold for some integer \( k \) throughout the paper.

2. THE PROBLEM

The research was motivated by the following modification of systematic sampling: for a fixed order of the units in \( U \), in place of the customary systematic sample

\[ s_{sy} = \{ i, i+k, i+2k, \ldots, i+(n-1)k \} \]

use

\[ s_{st} = \{ i_1, i_2+k, i_3+2k, \ldots, i_{n}+(n-1)k \} \]

where \( i_1, i_2, i_3, \ldots, i_n \) are different random integers between 1 and \( k \). \( s_{st} \) is obviously a stratified random sample with one unit per stratum. If \( n \) and \( N \) are fixed, the numbers of different systematic samples and stratified samples with one unit per stratum are \( k^n \) and \( nk^n \), respectively, thus one might expect that the latter are superior to the former. However, the comparison of the two designs in terms of variance turns to be quite hard.

Considering samples of equal size, Cochran (1977) specifies situations where systematic sampling is superior to simple random sampling. Among the drawbacks of the method, he mentions that hidden periodicity in the order of units may result in poor precision of \( \overline{y}_{sy} \) and no reliable procedure is available to estimate \( V(\overline{y}_{sy}) \) from the sample.

Nevertheless, the sample estimate

\[ \nu(\overline{y}) = \frac{1-f}{n(n-1)} \sum_{i=1}^{n} (y_i - \overline{y})^2 \]

where \( f = n/N \), the stratum variance is

\[ S_h^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (y_{ih} - \overline{y}_h)^2 \]

and the sample size is

\[ n_h = \frac{n}{N_h} \]

are stratum weights and estimated stratum means, respectively.
of the variance of an estimated mean in simple random sampling is widely used to estimate \( V(\bar{y}_k) \). This practice may be supported by a result of W.G. Madow–L.H. Madow (1944), quoted also in Cochran (1977); this asserts that

\[
E(V(\bar{y}_k)) = V(\bar{y}) = (1 - f) \frac{S^2}{n},
\]

where the expectation is taken over all permutations of the units 1, 2, ..., \( N \) of \( U \).

As for the stratified sample with one unit per stratum, the theoretical variance in this case is

\[
V(\bar{y}_{st}) = (1 - \frac{1}{k}) \frac{1}{n^2} \sum_{h=1}^{c} \sum_{j=1}^{n_h} S^2_h ,
\]

since each \( n_h = 1 \), the sampling fraction in each stratum is \( 1/k \), and each stratum weight \( W_h \) equals \( k/N = 1/n \). Taking the expectation of \( V(\bar{y}_{st}) \) over all permutations of the units of \( U \) leads to the same result as in the case of systematic sampling, i.e. the theoretical variance of the mean under SRS. In other words, comparing systematic sampling and stratified sampling with one unit per stratum on the basis of the expectations of \( s_y^2 \) and \( s^2 \) results in a draw.

Since \( S^2_h \) in /3/ cannot be estimated on the basis of a single observation, the usual sample-based estimator of \( V(\bar{y}_{st}) \) breaks down. Cochran (1977) enlists a number of different approaches to estimate \( V(\bar{y}_{st}) \) from the sample in the case of one unit per stratum; some of these work with collapsing adjacent strata, while other methods use auxiliary variables or specific hypotheses on the properties of the units of \( U \). W.A. Fuller's unbiased estimator (1970) for \( V(\bar{y}_k) \) does not use collapsed strata or auxiliary variables, but randomizes the strata boundaries. The bulk of this paper is the relation between the expectation of \( \sigma^2(\bar{y}) \) in /2/ under the stratified design with one unit per stratum and the variance \( V(\bar{y}_{st}) \).

3. ALTERNATIVE VARIANCE EXPRESSIONS

Lemma 1. With the notations of section 1, we have

\[
S^2 = \frac{1}{N-1} \sum_{i=1}^{N} y_i^2 - \frac{1}{N(N-1)} \sum_{i,j}^{N} y_i y_j \quad /4/
\]

and

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{i,j}^{n} y_i y_j . \quad /5/
\]

The statement is proved by routine computation.
Equations /4/ and /5/ can be rewritten in matrix-vector form. Denote e.g. \( \mathbf{I} \) the unit matrix of order \( N \), \( \mathbf{E} \) an \( N \times N \) matrix whose entries are all equal to 1, and \( \mathbf{y} \) the \( N \)-vector with the components \( y_1, y_2, \ldots, y_N \). /4/ becomes then

\[
S^2 = \frac{1}{N-1} \mathbf{y} \cdot (\mathbf{I} - \frac{1}{N} \mathbf{E}) \mathbf{y},
\]

where the prime denotes transpose. Note that \( \mathbf{C} = \frac{1}{N-1}(\mathbf{I} - \frac{1}{N} \mathbf{E}) \) is symmetric, positive semidefinite, having \( N-1 \) eigenvalues equal to \( 1/(N-1) \) and one eigenvalue equal to 0.

**Lemma 2. (Decomposition of the Variance).** Let \( n, N \) and \( k \) be integers and \( N=kn \). Decompose the universe \( \mathcal{U} \) in two parts \( \mathcal{U}_1 = \{1, 2, 3, \ldots, N-k\} \) and \( \mathcal{U}_2 = \{N-k+1, N-k+2, \ldots, N\} \), and denote \( S(\mathcal{U}_1)^2 \) and \( S(\mathcal{U}_2)^2 \) the corresponding variances of the study variable \( y \).

We have

\[
S^2 = S(\mathcal{U})^2 = \frac{(N-k-1)(N-k)}{N(N-1)} S(\mathcal{U}_1)^2 + \frac{k}{N(N-1)} \sum_{i=1}^{N-k} y_i^2 - \frac{2}{N(N-1)} \sum_{i=1}^{N-k} \sum_{j=N-k+1}^{N} y_i y_j +
\]

\[
+ \frac{k(k-1)}{N(N-1)} S(\mathcal{U}_2)^2 + \frac{N-k}{N(N-1)} \sum_{j=N-k+1}^{N} y_j^2
\]

The proof is done by routine computation.

**Corollary 1.** Under the conditions of Lemma 2, the variance for the subpopulation \( \mathcal{U}_2 \) is given as:

\[
S(\mathcal{U}_2)^2 = \frac{n(N-1)}{k-1} S(\mathcal{U})^2 = \frac{(n-1)(N-k-1)}{k-1} S(\mathcal{U}_1)^2 + \frac{2}{k(k-1)} \sum_{j=N-k+1}^{N} \sum_{i=1}^{N-k} y_i y_j -
\]

\[
- \frac{1}{k-1} \sum_{i=1}^{N-k} y_i^2 - \frac{n-1}{k-1} \sum_{j=N-k+1}^{N} y_j^2.
\]

This identity can be used to estimate \( S(\mathcal{U}_2)^2 \) if none of the units \( N-k+1, N-k+2, \ldots, N \) is observed.

4. The expectation of \( s^2 \) under stratified sampling

With one unit per stratum

For sample /1/ replace the indices \( i_1, i_2+k, i_3+2k, \ldots, i_n+(n-1)k \) by 1, 2, 3, \ldots, \( n \), respectively, and compute

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{i,j} y_i y_j.
\]
Lemma 3. Under stratified sampling with one unit per stratum, the expectation of $s^2$ in (8) is

$$E(s^2) = \frac{1}{N} \sum_{i=1}^{N} y_i^2 - \frac{2}{N(N-k)} \sum_{i<j, j<k} y_i y_j. \quad /9/$$

Proof: There are $n$ strata, each consisting of $k$ units. Therefore:

- $k^n$ is the number of all different samples, thus the probability of each sample $s$ is $\pi(s) = k^{-n}$,
- $k^{n-1}$ is the number of samples containing a fixed unit $i$,
- $k^{n-2}$ is the number of samples containing a fixed unit $i$ from stratum $h$ and a fixed unit $j$ from stratum $h'$, $h' \neq h$,
- there is no sample containing two different units $i$ and $j$ from the same stratum $h$.

Rewrite (8) as follows:

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{i<j} y_i y_j, \quad /8a/$$

multiply both sides by $\pi(s)$, and take the sum over all samples $s$. By symmetry, the result will be the following:

$$E(s^2) = \lambda \sum_{i=1}^{N} y_i^2 - 2\mu \sum_{i<j, j<k} y_i y_j. \quad /9/$$

The first term on the right-hand side contains a factor $(1/n) \times k^{-n}$, and since each $y_i^2$ occurs $k^{n-1}$ times, and $nk = N$, it follows that $\lambda = 1/N$. On the other hand, $s^2$ in (8a) vanishes for $y_1 = y_2 = \ldots = y_n$, which implies the similar relation for $E(s^2)$. This results in $\mu = 1/(N(N-k))$. The proof is thereby complete.

Denote $S^2$ the left-hand side of (9). The matrix-vector form of that relation is the following:

$$S^2 = \frac{1}{N} y(I - \frac{1}{N-k} D)y, \quad /10/$$

where $D$ is the direct (or Kronecker) product of the $n \times n$ matrix

$$a = \begin{pmatrix} 0 & 1 & 1 & \ldots & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \ldots & 0 \end{pmatrix}$$
and the $k \times k$ matrix $b$, whose entries are all equal to 1:

$$D = a \otimes b.$$ 

$D$ is of the following form:

$$D = \begin{pmatrix}
0 & b & b & \ldots & b \\
b & 0 & b & b & b \\
b & b & 0 & b & b \\
& & & \ddots & \\
b & b & b & b & 0
\end{pmatrix}.$$

**Lemma 4.** $D$ has

- $n-1$ eigenvalues equal to $-k$,
- one eigenvalue equal to $N-k$,
- $N-n$ eigenvalues equal to 0.

**Proof:** $D$ is a symmetric matrix, hence there is an $N \times N$ orthonormal matrix $U$ such that $U'DU$ is a diagonal matrix (here and in what follows, the prime denotes transpose). Such a matrix $U$ can be defined as follows. Let $u$ and $c$ be the following $k \times k$ and $n \times n$ matrices, respectively:

$$u = \begin{pmatrix}
2^{-1/2} & 6^{-1/2} & 12^{-1/2} & \ldots & (k(k-1))^{-1/2} & k^{-1/2} \\
-2^{-1/2} & 6^{-1/2} & 12^{-1/2} & \ldots & (k(k-1))^{-1/2} & k^{-1/2} \\
0 & -2 \times 6^{-1/2} & 12^{-1/2} & \ldots & (k(k-1))^{-1/2} & k^{-1/2} \\
0 & 0 & -3 \times 12^{-1/2} & \ldots & (k(k-1))^{-1/2} & k^{-1/2} \\
& & & \ddots & \\
0 & 0 & 0 & \ldots & -(k-1)(k(k-1))^{-1/2} & k^{-1/2}
\end{pmatrix},$$

$$c = \begin{pmatrix}
2^{-1/2} & 6^{-1/2} & 12^{-1/2} & \ldots & (n(n-1))^{-1/2} & n^{-1/2} \\
-2^{-1/2} & 6^{-1/2} & 12^{-1/2} & \ldots & (n(n-1))^{-1/2} & n^{-1/2} \\
0 & -2 \times 6^{-1/2} & 12^{-1/2} & \ldots & (n(n-1))^{-1/2} & n^{-1/2} \\
0 & 0 & -3 \times 12^{-1/2} & \ldots & (n(n-1))^{-1/2} & n^{-1/2} \\
& & & \ddots & \\
0 & 0 & 0 & \ldots & -(n-1)(n(n-1))^{-1/2} & n^{-1/2}
\end{pmatrix},$$

and set $U = c \otimes u$, the *Kronecker* product of $c$ and $u$. It is easy to check that $u'u = I_k$ and $UU = I$ where $I_k$ and $I$ are unit matrices of order $k$ and $N$, respectively.
It is also easy to see that
\[ \mathbf{M} = \mathbf{U}' \mathbf{D} \mathbf{U} \]
is a diagonal matrix such that
\[ \mathbf{M} = \text{diag}(\mu_1, \mu_2, \ldots, \mu_n), \]
the \( \mu \)s are \( k \times k \) diagonal matrices, \( \mu_1 = \mu_2 = \ldots = \mu_{n-1} = \text{diag}(0, 0, \ldots, 0, -k) \), and \( \mu_n = \text{diag}(0, 0, \ldots, 0, N-k) \), so the proof is complete.

Corollary 2. Denote \( \mathbf{C}_0 \) the matrix of the quadratic form in /10/. \( \mathbf{C}_0 \) has
\(-n-1\) eigenvalues equal to \( n/(N(n-1)) \),
\(-1\) vanishing eigenvalue and
\(-N-n\) eigenvalues equal to \( 1/N \).

Theorem 1. If the conditions of Lemma 3 hold, the relative difference between the expectation of the SRS variance estimate /8/ and the population variance \( S^2 \) does not exceed \( 1/(n-1) \), i.e.
\[ \frac{|E(s^2) - S^2|}{S^2} \leq \frac{1}{n-1} \]

Proof: Let \( \mathbf{e} = (1, 1, \ldots, 1)' \) be the \( N \)-vector with all components equal to 1, and \( \mu \mathbf{e} \) the orthogonal projection of \( \mathbf{y} \) onto \( \mathbf{e} \) where \( \mathbf{y} \) represents the values of the study variable on the units of \( \mathbf{U} \). Then \( \mathbf{y} = \mathbf{z} + \mu \mathbf{e} \) and \( \mathbf{z} \) is orthogonal to \( \mathbf{e} \), \( \mathbf{z}' \mathbf{e} = 0 \). Consider the matrix-vector representations /6/ and /10/ of \( S^2 \) and \( S^2* \), respectively. Note that \( \mathbf{y}' \mathbf{C}_0 \mathbf{y} = \mathbf{z}' \mathbf{C}_0 \mathbf{z} \), and that the columns of \( \mathbf{U} \) in the proof of Lemma 4 are eigenvectors of the matrix \( \mathbf{E} \), too. Set \( \mathbf{A} = \mathbf{U}' \mathbf{E} \mathbf{U} \). We have
\[ |S^2 - S^2*| = |\mathbf{z}' \mathbf{C}_0 \mathbf{z} - \mathbf{z}' \mathbf{C}_0 \mathbf{z}| = |\mathbf{z}' \mathbf{U}' \left( \frac{1}{(N-1)N} \mathbf{I} - \frac{1}{N(N-1)} \mathbf{A} \right) \mathbf{U} \mathbf{z}| = \]
\[ = |\mathbf{z}' \mathbf{U}' \left( \frac{1}{(N-1)N} \mathbf{I} - \frac{1}{N(N-1)} \mathbf{M} \right) \mathbf{U} \mathbf{z}| \leq \rho |\mathbf{z}' \mathbf{U}' \mathbf{U} \mathbf{z}| = \rho \mathbf{z}' \mathbf{z}, \]
where \( \rho \) is the maximum of the absolute values of diagonal entries in the diagonal matrix within the brackets. The latter will be denoted by \( \mathbf{T} = \text{diag}(\tau_1, \tau_2, \ldots, \tau_N) \). Let \( \lambda_i \) and \( \mu_i \) be the \( i \)-th diagonal entry in \( \mathbf{A} \) and \( \mathbf{M} \), respectively. The following cases occur:

<table>
<thead>
<tr>
<th>( \lambda_i )</th>
<th>( \mu_i )</th>
<th>( \tau_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1/(N(N-1))</td>
</tr>
<tr>
<td>0</td>
<td>-k</td>
<td>1/(N(N-1))-1/(N(N-1))</td>
</tr>
<tr>
<td>N</td>
<td>N-k</td>
<td>0</td>
</tr>
</tbody>
</table>
Ignoring the term \(1/((N-1)N)\), we obtain \(\rho = 1/((n-1)N)\) and

\[
|S^2 - \hat{S}^2| \leq \frac{1}{(n-1)N} \mathbf{z}' \mathbf{z} < \frac{1}{n-1} \mathbf{z}' \mathbf{C} \mathbf{z}
\]

as was to be shown.

5. THE EXPECTATION OF \(v(\bar{y})\) UNDER STRATIFIED SAMPLING WITH ONE UNIT PER STRATUM

Recall that the variance of the mean \(\bar{y} = \bar{y}_h\) under stratified sampling with one unit per stratum is

\[
V(\bar{y}) = (1 - \frac{1}{k}) \frac{1}{n} \sum_{h=1}^{n} S_h^2
\]

and that direct estimation of the stratum variance \(S_h^2\) on the basis of a single observation is not possible. In this section we assume that \(n > 2\) and introduce the following estimator for \(S_h^2\):

\[
s^2_{(h)} = \frac{n(N-1)}{k-1} \hat{s}^2 - \frac{(n-1)(N-k-1)}{k-1} s^2_{(-h)} - \frac{k}{k-1} \sum_{l \neq h} y_l^2 - \frac{(n-1)k}{k-1} \hat{y}_h^2 + \frac{2k}{k-1} \sum_{l \neq h} y_l y_h, \quad /11/\]

where

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{i,j} y_i y_j,\]

\[
s^2_{(-h)} = \frac{1}{n-2} \sum_{l \neq h} y_l^2 - \frac{1}{(n-1)(n-2)} \sum_{l \neq h} y_l y_j,\]

i.e. \(s^2_{(-h)}\) is computed similarly as \(s^2\) but without the observation from stratum \(h\). The properties of the estimator /11/ will be examined in several steps.

Lemma 5. /11/ can be rewritten as

\[
s^2_{(h)} = y_h^2 + \frac{1}{(n-1)(n-2)} \sum_{l \neq h} y_l y_j - \frac{2}{n-1} \sum_{l \neq h} y_l y_h,\]

and adding up \(s^2_{(h)}\) for \(h = 1, 2, ..., n\) we obtain

\[
\sum_{h=1}^{n} s^2_{(h)} = \sum_{h=1}^{n} y_h^2 - \frac{1}{n-1} \sum_{h \neq h'} y_h y_{h'},
\]

The result follows with routine computation.
This lemma and /3/ result in the following estimator for $V(\bar{y}_{st})$

$$
V(\bar{y}_{st}) = \left(1 - \frac{1}{k}\right) \frac{1}{n} \left(\frac{1}{n-1} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{j=1}^{n} y_j^2\right)
$$

which is formally the same as the sample estimate of the variance of an estimated mean $\bar{y}$ under simple random sampling. Note that the indices in /12/ come from those reflecting both stratum and position in the same way as at the beginning of Section 4. By Lemma 3 we have

$$
E V(\bar{y}_{st}) = \left(1 - \frac{1}{k}\right) \frac{1}{n} S^2
$$

where $S^2$ denotes the left-hand side of /9/.

Lemma 6. Set $h = n$. The expectation of $s^2_{(n)}$ under the one-unit-per-stratum design is

$$
E(s^2_{(n)}) = \frac{n(N-1)}{k-1} S^*(\mathbb{U})^2 - \frac{(n-1)(N-k-1)}{k-1} S^*(\mathbb{U}_1)^2 + \frac{2}{k(k-1)} \sum_{i=1}^{N-k} \sum_{j=N-k+1}^{N} y_i y_j - \frac{1}{k-1} \sum_{i=1}^{N-k} y_i^2 - \frac{n-1}{k-1} \sum_{j=N-k+1}^{N} y_j^2.
$$

where $S^*(\mathbb{U})^2$ and $S^*(\mathbb{U}_1)^2$ denote the expectation of $s^2$ under the one-unit-per-stratum design for the universe $\mathbb{U}$ and the union of the first $n-1$ strata $\mathbb{U}_1$, respectively.

Proof: The first two terms on the right-hand side of /13/ are obviously the expectations of those on the right-hand side of /11/. The repeat of a part of arguments in the proof of Lemma 3 implies the similar statement for the last three terms on the right-hand side of /13/.

Corollary 3. The bias of $s^2_{(n)}$, or, in general, that of $s^2_{(h)}$ is

$$
E(s^2_{(h)}) - S^2_h = \frac{n(N-1)}{k-1} (S^*(\mathbb{U})^2 - S(\mathbb{U})^2) - \frac{(n-1)(N-k-1)}{k-1} (S^*(\mathbb{U}_1)^2 - S(\mathbb{U}_1)^2).
$$

where $\mathbb{U}$ is the universe and $\mathbb{U}_1$ stands for the union of the strata except for stratum $h$.

Proof: The Corollary follows immediately from /13/ and /7/.

Lemma 5 and Corollary 3 lead to the following consequence:

Lemma 7.

$$
\text{Bias } V(\bar{y}_{st}) = \left(1 - \frac{1}{k}\right) \frac{1}{n} \frac{1}{n^2} \sum_{h=1}^{n} s^2_h =
$$

$$
= \frac{1}{n^2} \sum_{i=1}^{n} y_i^2 - \frac{1}{n^2} \sum_{i=1}^{N-k} y_i^2 - \frac{1}{n^2} \sum_{i=1}^{N-k} y_i y_j + \frac{1}{n^2} \sum_{i,j}^{N-k} y_i y_j,
$$

where $S^2$ denotes the left-hand side of /9/.
or, in matrix notations,
\[
\text{Bias } v(\bar{y}_{st}) = -\frac{1}{n^2 k^2} y'(I + \frac{N-1}{N-k} D - E)y = -\frac{1}{n^2 k^2} y'Ky \quad /15a/
\]
where \(y\) is the vector of the study variable in the universe.

Proof: /15/ follows from /14/ by adding up the terms \(E\left(\nu_s^2(h)\right)-S_h^2\) for \(h = 1, 2, \ldots, n\) and using the matrix representation for \(S^*(\cdot)^2\) and \(S(\cdot)^2\). Note that the first term on the right-hand side of /14/ repeats \(n\) times. In the second term, different \((N-k)\times(N-k)\) versions of \(D\) and \(E\) occur; the coefficients in /15/ depend on the number of occurrences of the \(k \times k\) submatrices or blocks of these matrices. Off-diagonal blocks of \(D\) and \(E\) occur \(n-2\) times and the diagonal blocks of \(E\) occur \(n-1\) times.

Lemma 8. For the orthonormal matrix \(U\) introduced in Lemma 4 we have
\[
K = U\Delta U',
\]
where \(\Delta = \text{diag}(\delta_1, \delta_2, \ldots, \delta_{n-1}, \delta_n)\), the \(\delta\)s are \(k \times k\) diagonal matrices, and \(\delta_1=\delta_2=\cdots=\delta_{n-1}=\delta_n=\text{diag}(1, 1, \ldots, 1, 0)\).

Proof: The Lemma is an immediate consequence of the previous results.

Consider an arbitrary study variable represented by its \(y\) vector and decompose it in two components \(x\) and \(z\) such that within each stratum \(h\), \(z_i\) is identical to the stratum mean of \(y\):
\[
z_i = \frac{1}{k} \sum_{j \in h} y_j, \quad \text{for } i \in \text{stratum } h,
\]
and \(x = y - z\). It is easy to see that \(z = z_0 + \alpha e\), where \(e = (1, 1, \ldots, 1)'\), \(\alpha\) is the population mean of the study variable, \(\alpha = \bar{y}\), and \(Kz_0 = \beta z_0\) with \(\beta = n(k-1)/(n-1)\). \(z_0\), \(e\) and \(x\) are pairwise orthogonal to each other, and \(Ex = 0, Dx = 0\). It follows that
\[
\text{Bias } v(\bar{y}_{st}) = -\frac{1}{n^2 k^2} x'x + \frac{n(k-1)}{n(n-1)} z_0'z_0 = -\frac{1}{n^2 k^2} x'x + \frac{1}{n} \frac{1}{N-k} \frac{1}{z_0'z_0}. /16/
\]

The results of this section are summarised in the following.

Theorem 2.
\[
V(\bar{y}_{st}) = (1 - \frac{1}{k}) \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^{n} y_i^2 - \frac{1}{n(n-1)} \sum_{i,j} y_i y_j) \quad /12/
\]
is an estimator for
\[
V(\bar{y}_{st}) = (1 - \frac{1}{k}) \frac{1}{n^2} \sum_{h=1}^{H} S_h^2
\]
with expectation

\[ E \, \text{var}(\bar{y}_{st}) = (1 - \frac{1}{k}) \frac{1}{n} \left( \frac{1}{N} \sum_{i=1}^{N} y_i^2 - \frac{1}{N(N-k)} \sum_{j \neq k}^{N} y_j y_j \right), \]

and bias given by /16/.

Remark. From the aspect of stratification, the case where \( y = z \) is ideal, the variance of \( \bar{y}_{st} \) is zero. Unfortunately, this is the worst case for the estimator /12/, the upward bias being of the same order as the expectation. If \( y = x \), the variance of \( \bar{x}_{st} \) and the bias of /12/ are

\[ \frac{1}{n^2 k} x'x \quad \text{and} \quad -\frac{1}{n^2 k} x'x, \]

respectively, in this case the relative downward bias is of the order 1/k. In principle there are cases where the bias vanishes.

REFERENCES

