A CLASS-CUM-CLASS ESTIMATORS FOR POPULATION RATIO IN POSTSTRATIFIED SAMPLING OVER TWO OCCASIONS

A. C. ONYEKA

(Received 14 February 2003; Revision accepted 19 June, 2003).

ABSTRACT

Onyeka (2002) proposed two estimators of the second occasion population ratio, $R_2$, in poststratified sampling over two occasions along the line of Tripathi and Sinha (1976) and Okafor (1985). The estimators are a ratio-cum-product type ($d_1$) estimator and a product-cum-ratio type ($d_2$) estimator. The present study provides a generalization of the estimators, $d_1$ and $d_2$, proposed by Onyeka (2002). This generalization is in the form of a class-cum-class estimators, $d_i$, for the second occasion population ratio, $R_2$, in poststratified sampling over two occasions. The proposed class-cum-class estimators, $d_i$, do not assume knowledge of the first occasion population ratio, $R_1$. Conditions under which some of the estimators in the proposed class perform better than the others (in terms of having smaller mean square errors) are obtained for repeated samples of fixed sizes.

KEYWORDS: Successive sampling, poststratified sampling, ratio estimation, matching (or replacement) fraction, repeated samples.

INTRODUCTION

The present study is an extension of the study by Onyeka (2002) who proposed two estimators, $d_1$ and $d_2$, for the estimation of the second occasion population ratio, $R_2$, in poststratified sampling over two occasions. The said-estimators, $d_1$ and $d_2$, are a ratio-cum-product type estimator and a product-cum-ratio type estimator respectively. The estimators were proposed along the line of Tripathi and Sinha (1976) and Okafor (1985). Tripathi and Sinha (1976) studied the use of a unistage cluster sampling scheme on two occasions, while Okafor (1985) worked on the use of a multistage cluster sampling scheme on two occasions.

The estimators proposed by Onyeka (2002), Tripathi and Sinha (1976) and Okafor (1985) for the estimation of $R_2$ were ratios (or quotients) of two estimators where the numerators and denominators were some ratio-type, product-type and difference-type estimators. In the present study, we propose an estimator, $d$, with both the numerator and denominator as classes of some estimators including the usual ratio-type, product-type and difference-type estimators. This approach gives rise to a class-cum-class estimators of the second occasion population ratio, $R_2$, in poststratified sampling over two occasions.

Properties of the proposed class of estimators are obtained for repeated samples of fixed sizes. Also obtained are the conditions under which some of the estimators in the proposed class-cum-class estimators perform better than the others in terms of having smaller mean square errors. The proposed estimators do not require knowledge of the first occasion population ratio, $R_1$, unlike the estimators proposed by authors like Rao (1957) and Rao and Pereira (1968).

A. C. ONYEKA, Department of Mathematics and Computer Science, Federal University of Technology, Owerri, Nigeria
THE PROPOSED ESTIMATORS

The sampling design used by Onyeka (2002) is the same sampling design for poststratified sampling over two occasions proposed by Onyeka (2001). The sampling design is as follows:

A random sample of size \( n \) is drawn from a population of \( N \) units using simple random sampling without replacement (SRSWOR) method on the first occasion. The sampled units are allocated to their respective strata where \( n_h \) is the number of units that fall into the \( h^{th} \) stratum such that \( \sum_h n_h = n \). ( \( h = 1, 2, \ldots, L \)).

It is assumed that \( n \) is large enough such that \( \text{Prob} (n_h = 0) = 0 \) for all \( h \). On the second occasion, \( m_h = \lambda n_h \) units of the first occasion sample are retained in the \( h^{th} \) stratum, \( \sum_h m_h = m = \lambda n \). ( \( 0 < \lambda < 1 \)). The remaining \( u_{1h} = n_h - m_h = n_h - \lambda n_h = \mu n_h \) units are discarded, \( \sum_h u_{1h} = u = \mu n \), and \( \mu + \lambda = 1 \). Then, the matched sample of size \( m \) is supplemented with a fresh (unmatched) sample of \( u \) units drawn independently from the entire population, again using SRSWOR method. The \( u \) sampled units are allocated to their respective strata where \( u_{2h} \) is the number of units that fall into the \( h^{th} \) stratum such that \( \sum_h u_{2h} = u \) \( = \sum_h u_{1h} \). Again, it is assumed that \( u \) is large enough such that \( \text{Prob} (u_{2h} = 0) = 0 \) for all \( h \).

Let \( y_{i1h} \) and \( x_{i1h} \) denote observations on the \( i^{th} \) unit of the two characters of study in the \( h^{th} \) stratum on the \( j^{th} \) occasion, \( i = 1, 2, \ldots, N; h = 1, 2, \ldots, L \) and \( j = 1, 2 \). The variate, \( X_i \), in some cases can be the stratification variate. But generally, the variates, \( Y \) and \( X \), are to be taken as any two characters of study.

Let \( \bar{Y}_i (\bar{Y}_2) \) and \( \bar{X}_i (\bar{X}_2) \) respectively denote the population means of \( Y \) and \( X \) on the first (second) occasion.

Let \( R_1 = \bar{Y}_2 / \bar{X}_2 \) and \( R_1 = \bar{Y}_1 / \bar{X}_1 \) respectively denote the second and first occasion population ratios of the two characters of study. We now propose a class-cum-class estimators, \( d \), for the second occasion population ratio, \( R_2 \), in poststratified sampling over two occasions as:

\[
d = \theta \left\{ \frac{\bar{y}_{2m} - b(\bar{y}_{1m} - \bar{y}_{1n})}{\left( t \bar{y}_{1m} + (1 - t) \bar{y}_{1n} \right)^\alpha (\bar{y}_{1n})^\alpha} \right\} \left( \frac{\bar{x}_{2m} - b_1(\bar{x}_{1m} - \bar{x}_{1n})}{\left( t_1 \bar{x}_{1m} + (1 - t_1) \bar{x}_{1n} \right)^{\alpha_1} (\bar{x}_{1n})^{\alpha_1}} \right) + (1 - \theta)(\bar{y}_{2u} / \bar{x}_{2u}) \quad (2.1)
\]

or:

\[
d = \theta d_m + (1 - \theta) \gamma_{2u} \quad (2.2)
\]

where \( d_m = \bar{y}_{2p} / \bar{x}_{2p} \), \( \gamma_{2u} = \bar{y}_{2u} / \bar{x}_{2u} \) \quad (2.3)

\[
\bar{y}_{2p} = \frac{\bar{y}_{2m} - b(\bar{y}_{1m} - \bar{y}_{1n})}{\left( t \bar{y}_{1m} + (1 - t) \bar{y}_{1n} \right)^\alpha (\bar{y}_{1n})^\alpha} \quad (2.4)
\]

\[
\bar{x}_{2p} = \frac{\bar{x}_{2m} - b_1(\bar{x}_{1m} - \bar{x}_{1n})}{\left( t_1 \bar{x}_{1m} + (1 - t_1) \bar{x}_{1n} \right)^{\alpha_1} (\bar{x}_{1n})^{\alpha_1}} \quad (2.5)
\]
\[ y_{mn} = \sum_{h} W_h y_{h}^m, \quad y_{1m} = \sum_{h} W_h y_{h}^1, \quad y_{1n} = \sum_{h} W_h y_{h}^1, \quad y_{1n} = \sum_{h} W_h y_{h}^1 \]

\[ \bar{y}_{1h}, \bar{x}_{1h} \text{ are sample means based on the entire first occasion sample of size } n_{1h} \]

\[ y_{2h}^m, x_{2h}^m, y_{2h}^n, x_{2h}^n \text{ are sample means based on the matched sample of size } m_n \]

\[ y_{2h}^u, x_{2h}^u \text{ are sample means based on the second occasion unmatched sample of size } u_{2h} \]

\[ (b, t, t, 1) \text{ are suitable constants used in defining the class of estimators, } \bar{y}_{2p}, \text{ of } \bar{y}_2. \]

\[ (b_1, t_1, 1, 1) \text{ are suitable constants used in defining the class of estimators, } \bar{x}_{2p}, \text{ of } \bar{x}_2. \]

and

\[ \theta \text{ is a constant (weighting) fraction of the matched and unmatched parts of the proposed class-cum-class estimators, } d. \]

**PROPERTIES OF THE PROPOSED ESTIMATORS**

Let \( S_{yh}^2 \) and \( S_{xh}^2 \) respectively denote the population variances of the study characters \( Y \) and \( X \), in the \( h \)-th stratum on both occasions. Also, let \( S_{yxh} \) denote the covariance of \( Y \) and \( X \) in the \( h \)-th stratum on both occasions. Following Onyeka (2002), the properties of the proposed class-cum-class estimators, \( d \), can be obtained as stated in the following theorems.

**Theorem 1**

For repeated samples of fixed sizes \( n, m \) and \( u \), the optimum value of the weighting fraction, \( \theta \), and the associated mean square error of the proposed class-cum-class estimators, \( d \), are respectively given by:

\[
\theta_0 = \frac{\lambda \sum_{h} W_h \sigma_{2h}}{\lambda \sum_{h} W_h \sigma_{2h} + \mu^2 \sum_{h} W_h \sigma_{21h} - 2\mu \sum_{h} W_h \sigma_{22h}} 
\]

\[
\text{MSE}(d) = \frac{\sum_{h} W_h \sigma_{2h} + \mu \sum_{h} W_h \sigma_{21h} - 2\mu \sum_{h} W_h \sigma_{22h}}{\sum_{h} W_h \sigma_{2h} + \mu^2 \sum_{h} W_h \sigma_{21h} - 2\mu \sum_{h} W_h \sigma_{22h}} \frac{\sum_{h} W_h \sigma_{2h}}{nX_2^2} 
\]

where

\[
\sigma_{2h} = S_{yh}^2 + R_{2}^2 S_{xh}^2 - 2R_{2} S_{yxh} ; \quad \sigma_{21h} = Q_y S_{yh}^2 + Q_x R_{2}^2 S_{xh}^2 - 2Q_y Q_x R_{2} S_{yxh} 
\]

\[
\sigma_{22h} = Q_y S_{yh}^2 + Q_x R_{2}^2 S_{xh}^2 - (Q_y + Q_x) R_{2} S_{yxh} ; \quad Q_y = b + \alpha (1 - t) R_y ; 
\]

\[ Q_x = b_1 + \alpha_1 (1 - t_1) R_x ; \quad R_2 = \frac{\bar{Y}_2}{\bar{X}_2}, \; R_1 = \frac{\bar{Y}_1}{\bar{X}_1}, \; R_y = \frac{\bar{Y}_2}{\bar{Y}_1}, \; R_x = \frac{\bar{X}_2}{\bar{X}_1} \]
Theorem 1 can be proved by following similar procedure as in the proof of Theorem 1 of Onyeka (2002).

**Theorem 2**

The optimum replacement or matching fraction, \( \lambda \), of the proposed class-cum-class estimators, \( d \), and the associated optimum mean square error of \( d \) are respectively given by:

\[
\lambda_0 = \frac{1 + \Lambda_0}{\sqrt{1 + \Lambda_0}}
\]

and \( \text{MSE}_0(d) = (2n\bar{X}_2)^{-1} \left( 1 + \sqrt{1 + \Lambda_0} \right) \sum_h W_h \sigma_{2h} \)  \hspace{1cm} (3.4)

where \( \Lambda_0 = \frac{\sum_h W_h \sigma_{21h} - 2 \sum_h W_h \sigma_{22h}}{\sum_h W_h \sigma_{2h}} \) \hspace{1cm} (3.6)

and \( \sigma_{21h} \), \( \sigma_{22h} \) and \( \sigma_{22h} \) are as previously defined in equation (3.3).

Theorem 2 can be proved by following similar procedure as in the proof of Theorem 3 of Onyeka (2002).

**COMPARISON OF PROPOSED ESTIMATORS**

The proposed class-cum-class estimators, \( d \), can generate a wide range of estimators of \( R_2 \) by suitable choices of the constants \( (b, t, C) \) and \( (b_1, t_1, C_1) \) as used in equation (2.1). Consider any two particular estimators in the proposed class, \( d \), say \( d(1) \) and \( d(2) \), where

\[
d(1) = 0 \left( \begin{array}{c}
\tilde{y}_{2m} - b'(\tilde{y}_{lm} - \tilde{y}_{ln}) \\
t'\tilde{y}_{ln} + (1-t')\tilde{y}_{lm} \\
^t_{\mu}I_{\mu}X_{\mu}\end{array} \right) + (1-0)(\tilde{y}_{2u}/\tilde{x}_{2u}) \hspace{1cm} (4.1)
\]

and \( d(2) = 0 \left( \begin{array}{c}
\tilde{y}_{2m} - b''(\tilde{y}_{lm} - \tilde{y}_{ln}) \\
t''\tilde{y}_{ln} + (1-t'')\tilde{y}_{lm} \\
^t_{\mu}I_{\mu}X_{\mu}\end{array} \right) + (1-0)(\tilde{y}_{2u}/\tilde{x}_{2u}) \hspace{1cm} (4.2)

The following theorem compares the efficiencies of the estimators, \( d(1) \) and \( d(2) \), in terms of the estimator with the smaller mean square error.

**Theorem 3**

For any two estimators, \( d(1) \) and \( d(2) \), in the proposed class-cum-class estimators, \( d \), the estimator \( d(1) \)
would be more efficient than the estimator d (2) in terms of having a smaller mean square error if

\[
(Q_{2x} - Q_{1x}^2 + 2Q_{1x} - 2Q_{2x}) \sum_h W_h S_{xh}' \\
+ (Q_{2x}^2 - Q_{1x}^2 + 2Q_{1x} - 2Q_{2x}) R_2^2 \sum_h W_h S_{xh}' \\
- 2(Q_{1x} + Q_{1x} - Q_{2x} - Q_{2x} - Q_{1x} Q_{1x} + Q_{2x} Q_{2x}) R_2 \sum_h W_h S_{xh} < 0
\]  

(4.3)

where

\[
Q_{iy} = b' + \alpha' (1 - t') R_y \\
Q_{ix} = b_1' + \alpha_1' (1 - t_1') R_x
\]  

(4.4)

\[
Q_{2y} = b'' + \alpha'' (1 - t'' ) R_y \\
Q_{2x} = b_1'' + \alpha_1'' (1 - t_1'') R_x
\]  

(4.5)

\[b', \alpha', t', b_1', \alpha_1', t_1' \text{ are suitable constants used in the estimator } d (1)\]

\[b'', \alpha'', t'', b_1'', \alpha_1'', t_1'' \text{ are suitable constants used in the estimator } d (2)\]

and \[R_y, R_x \text{ are as given in equation (3.3)}\]

PROOF:

Using equation (3.2), we obtain the respective mean square errors of the estimators d (1) and d (2) as:

\[
\text{MSE}(d(1)) = \frac{\sum_h W_h \sigma_{2h} + \mu \sum_h W_h \sigma_{21h} - 2\mu \sum_h W_h \sigma_{22h} - \sum_h W_h \sigma_{2h}^2}{n \bar{X}_2^2}
\]

(4.6)

and

\[
\text{MSE}(d(2)) = \frac{\sum_h W_h \sigma_{2h} + \mu \sum_h W_h \sigma_{21h} - 2\mu \sum_h W_h \sigma_{22h} - \sum_h W_h \sigma_{2h}^2}{n \bar{X}_2^2}
\]

(4.7)

where

\[
\sigma_{2h} = S_{xh}^2 + R_2^2 S_{xh}^2 - 2R_2 S_{xyh}
\]

\[
\sigma_{21h} = Q_{1y}^2 S_{xy}^2 + Q_{1x}^2 R_2^2 S_{xh}^2 - 2Q_{1y} Q_{1x} R_2 S_{xyh}
\]

\[
\sigma_{22h} = Q_{1y} S_{xy}^2 + Q_{1x} R_2^2 S_{xh}^2 - (Q_{1y} + Q_{1x}) R_2 S_{xyh}
\]

\[
\sigma_{21h}^* = Q_{2y}^2 S_{xy}^2 + Q_{2x}^2 R_2^2 S_{xh}^2 - 2Q_{2y} Q_{2x} R_2 S_{xyh}
\]

\[
\sigma_{22h}^* = Q_{2y} S_{xy}^2 + Q_{2x} R_2^2 S_{xh}^2 - (Q_{2y} + Q_{2x}) R_2 S_{xyh}
\]

(4.8)
Equation (4.3) is obtained by considering the difference of the mean square errors of $d(1)$ and $d(2)$ as respectively given in equations (4.6) and (4.7). This completes the proof.

PARTICULAR CASES OF THE PROPOSED ESTIMATORS

The following are some particular estimators in the proposed class-cum-class estimators, $d$, for the estimation of the second occasion population ratio, $R_2$, in poststratified sampling over two occasions.

(a) **Ratio-cum-Product Estimator**
Substituting $b = b_t = 0$; $t = t_i = 0$; $\alpha = 1$; $\alpha_i = -1$ in equation (2.1) gives the ratio-cum-product estimator as

$$d_i = \theta_1 \left( \frac{\bar{y}_{2m} - \bar{y}_{1m}}{\bar{y}_{1m}} / \frac{\bar{x}_{2m} - \bar{x}_{1m}}{\bar{x}_{1m}} \right) + (1 - \theta_1)(\bar{y}_{2u} / \bar{x}_{2u})$$  \hspace{1cm} (5.1)

Here, the matched estimator is a ratio of a ratio-type estimator of $\bar{Y}_2$ to a product-type estimator of $\bar{X}_2$. It is easy to verify that $Q_{1y} = R_y$ and $Q_{1x} = -R_x$.

(b) **Product-cum-Ratio Estimator**
Substituting $b = b_t = 0$; $t = t_i = 0$; $\alpha = -1$; $\alpha_i = 1$ in equation (2.1) gives the product-cum-ratio estimator as:

$$d_2 = \theta_2 \left( \frac{\bar{y}_{2m} - \bar{y}_{1m}}{\bar{y}_{1m}} / \frac{\bar{x}_{2m} - \bar{x}_{1m}}{\bar{x}_{1m}} \right) + (1 - \theta_2)(\bar{y}_{2u} / \bar{x}_{2u})$$  \hspace{1cm} (5.2)

Here, the matched estimator is a ratio of a product-type estimator of $\bar{Y}_2$ to a ratio-type estimator of $\bar{X}_2$. It is easy to verify that $Q_{2y} = -R_y$ and $Q_{2x} = R_x$.

(c) **Difference-cum-Ratio Estimator**
With $\alpha = 0$ or $t = 1$; $b_t = 0$; $t_i = 0$; $\alpha_i = 1$; we have the difference-cum-ratio estimator as:

$$d_3 = \theta_3 \left( \frac{\bar{y}_{2m} - b(\bar{y}_{1m} - \bar{y}_{1n})}{\bar{y}_{1m}} / \frac{\bar{x}_{2m} - \bar{x}_{1m}}{\bar{x}_{1m}} \right) + (1 - \theta_3)(\bar{y}_{2u} / \bar{x}_{2u})$$  \hspace{1cm} (5.3)

This is a ratio of a difference-type estimator of $\bar{Y}_2$ to a ratio-type estimator of $\bar{X}_2$ with $Q_{3y} = b$ and $Q_{3x} = R_x$.

(d) **Ratio-cum-Ratio Estimator or Ratio-Type Estimator**
With $b = b_t = 0$; $t = t_i = 0$; $\alpha = \alpha_i = 1$; we have the ratio-cum-ratio estimator as:

$$d_4 = \theta_4 \left( \frac{\bar{y}_{2m}}{\bar{y}_{1m}} / \frac{\bar{x}_{2m} - \bar{x}_{1m}}{\bar{x}_{1m}} \right) + (1 - \theta_4)(\bar{y}_{2u} / \bar{x}_{2u})$$  \hspace{1cm} (5.4)
This is a ratio of a ratio-type estimator of $\bar{Y}_2$ to another ratio-type estimator of $\bar{X}_2$ with $Q_{4y} = R_y$ and $Q_{4x} = R_x$.

(e) **Product-cum-Product Estimator or Product-Type Estimator**

With $b = b_1 = 0$; $t = t_1 = 0$; $\alpha = \alpha_1 = -1$; we have the product-cum-product estimator as:

$$d_x = \delta_3 \left( \frac{\bar{Y}_{2m} \cdot \bar{Y}_1m}{\bar{Y}_{1n}} / \frac{\bar{X}_{2m} \cdot \bar{X}_1m}{\bar{X}_{1n}} \right) + (1 - \delta_3)(\bar{Y}_{2n} / \bar{X}_{2n}) \quad (5.5)$$

This is a ratio of a product-type estimator of $\bar{Y}_2$ to another product-type estimator of $\bar{X}_2$ with $Q_{5y} = -R_y$ and $Q_{5x} = -R_x$.

Theoretically, an infinite number of estimators can be generated from the proposed class-cum-class estimators, $d$. The results in Theorem 3 can always be used to compare the efficiencies of the generated (or particular) cases of the proposed estimators. For instance, it is easy to verify from Theorem 3 that the ratio-cum-product estimator, $d_1$, would perform better than the estimators $d_2, d_3, d_4$ and $d_5$ if conditions (1), (2), (3) and (4) are respectively satisfied, where:

1. $$\sum_h W_h S_{y,h}^2 > R_2 R_1 \sum_h W_h S_{x,h}^2 + (R_2 - R_1) \sum_h W_h S_{xy,h} \quad (5.6)$$

2. $$b^2 - R_y - 2b + 2R_y \sum_h W_h S_{y,h}^2 < 4R_x R_2^2 \sum_h W_h S_{x,h}^2$$
$$+ 2(R_y - 2R_x - b + R_y R_x + bR_x) R_2 \sum_h W_h S_{xy,h} \quad (5.7)$$

3. $$R_2 \sum_h W_h S_{x,h}^2 > (2 - R_y) \sum_h W_h S_{xy,h} \quad (5.8)$$

4. $$\sum_h W_h S_{y,h}^2 < (1 - R_x) R_2 \sum_h W_h S_{xy,h} \quad (5.9)$$

In summary, the proposed class-cum-class estimators, $d$, is a generalization of many estimators of the second occasion population ratio, $R_2$, in poststratified sampling over two occasions. With the class of estimators, $d$, proposed in this study, we can generate a wide range of estimators of $R_2$ in poststratified sampling over two occasions. Theorem 3 offers a very useful tool in terms of comparing the efficiencies of generated or particular cases of the proposed class of estimators. Interestingly also is the fact that the proposed estimators do not require knowledge of the first occasion population ratio, $R_1$. This, obviously, is an added advantage over many other ratio estimators that require information on $R_1$.

**REFERENCES**

Okafor, F. C., 1985. Use of multistage sampling over two occasions. A Ph.D. thesis submitted to the Faculty of Science, University of Ibadan, Nigeria.


