

Estimation of Variance for Domain Total: A High Level Calibration Approach

Ashutosh^{1*}

¹Department of Statistics, Institute of Science, Banaras Hindu University, India-221005

*Corresponding Author: kumarashubhustat@gmail.com

Received: 28th February 2020 / Revised: 15th October 2020 / Published: 30th November 2020

©IApstat-SL2020

ABSTRACT

Sample surveys are used to estimate interest about the population parameters. But, the objective of this study is to estimate about the sub-populations (termed domain). The proposed method is based on direct generalized calibration estimator for domain total using auxiliary character. The properties of the estimators also studied using simple random sampling process. It is also evaluate a class of estimators for domain total. A comparative study of the proposed estimator is given theoretically in the support of the problem. It was found that that the high level calibration approach better than the low level calibration approach with respect to (Sarndal et al. 1992, appendix B). A real data set is used to illustrate our methodology.

Keywords: Calibration Approach, Domain, Direct Estimator, Auxiliary Character, Ratio and Regression Estimator.

1. Introduction to Research Problem

Sample surveys are basically used to estimates the information for population parameters. But, we are seldom estimates the information for the sub-population (domain) rather than the population. The imminence of estimation of domain parameter has become very elegant regarding the formulation of the program and policy. The main idea of the domain estimation is that, the number of sufficient sampling units available in the study domains, and to gain in the desire level of the accuracy for the estimator. In this relation,

imminence books (Sarndal et al., 1992 and Rao, 2003) have been discussed about the small area estimation. I consider generalized linear regression (GREG) estimator for domain total which depends on the single auxiliary character x and use calibration approach under different sampling procedure. The main idea of the use of calibration estimator is to improve the efficiency of the estimator with help of the available information of the auxiliary character. (Horvitz and Thompson, 1952) have been contributed to estimate the population total using usual sampling designs with weights $d_i = 1/\pi_i$ is given by

$$\hat{Y}_{HT} = \sum_{i=1}^n d_i y_i \quad (1)$$

Later by using new weights w_i and proposed the calibration estimator for population total has been discussed (Deville and Sarndal, 1992)

$$\hat{Y}_{DS} = \sum_{i=1}^n w_i y_i \quad (2)$$

where, w_i is nearer to d_i under the calibration equation with available auxiliary character

$$\sum_{i=1}^n w_i x_i = X \quad (3)$$

With the assumption that the minimum chi-square type distance function is given by

$$\sum_{i=1}^n \frac{(w_i - d_i)^2}{w_i q_i} \quad (4)$$

where, q_i is suitable chosen constant, and using the calibration equations (3) and (4) with help of the Lagrange's multiplier λ , the new weight is obtained using the calibration functions is given by

$$\sum_{i=1}^n \frac{(w_i - d_i)^2}{w_i q_i} - 2\lambda \left(X - \sum_{i=1}^n d_i x_i \right) = 0 \quad (5)$$

Partially differentiate w.r.to λ and equate to zero, then substitute the value of λ in the Equation (3), then we have

$$w_i = d_i + \frac{d_i q_i x_i}{\sum_{i=1}^n d_i q_i x_i^2} \left(X - \sum_{i=1}^n d_i x_i \right) \quad (6)$$

Substitute in the Equation (2), then the new estimator is given as:

$$\hat{Y}_{DS} = \sum_{i=1}^n d_i y_i + \frac{\sum_{i=1}^n d_i q_i x_i y_i}{\sum_{i=1}^n d_i q_i x_i^2} \left(X - \sum_{i=1}^n d_i x_i \right) \quad (7)$$

The estimation of variance of the Equation (7) has been considered at two levels using calibration approach which are ratio and regression. Various estimates using the low level calibration approach have been considered (Sarndal, 1996) in which they utilize only available single auxiliary character x for population total. However, in the high level calibration approach, use of population total as well as available variance for the auxiliary character.

2. Methodology

In the present context, I am studying the sub-population characters. In this connection, the estimation of domain total has been considered by (Tikkiwal et al., 2013). Let us consider pre-defined a^{th} domains $\Omega_a : \{a = 1, 2, \dots, A\}$ where,

$$\Omega_1 + \Omega_2 + \Omega_3 + \dots + \Omega_A = \sum_{i=1}^A \Omega_a \quad (8)$$

Probability sample of a^{th} domains $s_a \{s_a \subset \Omega_a\}$ is drawn with known sampling design, $p(\cdot)$. The inclusion probability of i^{th} unit for a^{th} domains, $\pi_{i,a} = P_r \{i_a \in s_a\} \geq 0$, the inclusion probability of j^{th} unit for a^{th} domains, $\pi_{j,a} = P_r \{j_a \in s_a\} \geq 0$ and the inclusion probability of i^{th} and j^{th} units for a^{th} domains, $\pi_{ij,a} = P_r \{i_a \text{ and } j_a \in s_a\} \geq 0$ (9)

The propose direct generalized regression estimator of domain total is given by

$$\hat{Y}_{PD,a} = \sum_{i=1}^{n_a} w_{i,a} y_{i,a} \quad (10)$$

It is obtaining weights of the a^{th} domains $w_{i,a}, (i = 1, 2, \dots, n_a)$ which is close to $d_{i,a}$, under the assumption, and property of unbiased of the Horvitz and Thompson estimator. It got the value of the new weight $w_{i,a}$ using the calibration estimator for Equation (10). (Deville and Sarndal, 1992) estimator has been discussed for population character. But here, I apply in the case of the domain. The calibration equation of the domain total using the auxiliary character x is given by

$$\sum_{i=1}^{n_a} w_{i,a} x_{i,a} = X_a \quad (11)$$

Under the assumption of minimizing the chi-square type distance function is given by

$$\sum_{i=1}^{n_a} \frac{(w_{i,a} - d_{i,a})^2}{w_{i,a} q_{i,a}} \quad (12)$$

where, $q_{i,a}$ is suitably chosen constants, now proceeding using the calibration equations (11) and (12), the new weights obtained by using the calibration function:

$$f(w_{i,a}, \lambda_a) = \sum_{i=1}^{n_a} \frac{(w_{i,a} - d_{i,a})^2}{w_{i,a} q_{i,a}} - 2\lambda_a \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right) = 0 \quad (13)$$

Partially differentiate w.r.to λ_a and equate to zero, then substitute value of λ_a in the Equation (11), then we have

$$w_{i,a} = d_{i,a} + \frac{d_{i,a} q_{i,a} x_{i,a}}{\sum_{i=1}^{n_a} w_{i,a} q_{i,a} x_{i,a}^2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right) \quad (14)$$

Now substitute in the Equation (10), then the new form of the proposed calibration estimator is given by

$$\hat{Y}_{PD,a} = \sum_{i=1}^{n_a} d_{i,a} y_{i,a} + \frac{\sum_{i=1}^{n_a} d_{i,a} q_{i,a} x_{i,a} y_{a,i}}{\sum_{i=1}^{n_a} w_{i,a} q_{i,a} x_{i,a}^2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right) \quad (15)$$

3. Estimation of Calibration Estimator for Domain

3.1 Estimation of Variance of the GREG Using Low Level Calibration Approach

Using variance of the population total using the calibration approach for population total has been discussed (Yates and Grundy, 1953)

$$\hat{V}_{YG}(\hat{Y}_{DS}) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n D_{ij} (d_i y_i - d_j y_j)^2 \quad (16)$$

Estimation of Variance for Domain Total:

A High Level Calibration Approach

$$\text{where, } D_{ij} = \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}}, \quad i \neq j, \quad e_i = y_i - \hat{\beta} x_i \quad \text{and} \quad \hat{\beta} = \frac{\sum_{i=1}^n d_i y_i}{\sum_{i=1}^n d_i x_i}$$

It may also write for estimation of the domain total:

$$\hat{V}_{YG}(\hat{Y}_{PD,DS,a}) = \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} \left(d_{i,a} y_{i,a} - d_{j,a} y_{j,a} \right)^2 \quad (17)$$

$$\text{Where, } D_{ij,a} = \frac{\pi_{i,a} \pi_{j,a} - \pi_{ij,a}}{\pi_{ij,a}}, \quad i \neq j,$$

$$e_{i,a} = y_{i,a} - \hat{\beta}_a x_{i,a} \quad \text{and} \quad \hat{\beta}_a = \frac{\sum_{i=1}^{n_a} d_{i,a} y_{i,a}}{\sum_{i=1}^{n_a} d_{i,a} x_{i,a}}.$$

Special cases of the variance of the propose estimator is given as follows:

$$\text{Case (i) } \hat{Y}_{YG,1,a} = \sum_{i=1}^{n_a} d_{i,a} y_{i,a} + \frac{\sum_{i=1}^{n_a} d_{i,a} x_{i,a} y_{i,a}}{\sum_{i=1}^{n_a} d_{i,a} x_{i,a}^2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right),$$

if $q_{i,a} = 1$ (Regression estimator) (18)

$$\text{Case (ii) } \hat{Y}_{YG,2,a} = \frac{\sum_{i=1}^{n_a} d_{i,a} y_{i,a}}{\sum_{i=1}^{n_a} d_{i,a} x_{i,a}} X_a, \quad \text{if } q_{i,a} = \frac{1}{x_{i,a}} \quad (\text{Ratio estimator})$$

(19)

The variance of the members of the proposed estimator using Equation (16) is given as:

$$\hat{V}(\hat{Y}_{PD,DS,a}) = \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} \left(w_{i,a} e_{i,a} - w_{j,a} e_{j,a} \right)^2 \quad (20)$$

$$\text{where, } D_{ij,a} = \frac{\pi_{ij,a} - \pi_{i,a} \pi_{j,a}}{\pi_{ij,a}} \quad i \neq j, \quad e_{i,a} = y_{i,a} - \hat{\beta}_a x_{i,a} \quad (21)$$

The Equation (18) may be written in the other form which is given as:

$$\hat{V}(\hat{Y}_{PD,a,DS}) = \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} \left(\frac{\pi_{ij,a} - \pi_{i,a}\pi_{j,a}}{\pi_{ij,a}} \right) (d_{i,a}e_{i,a} - d_{j,a}e_{j,a})^2 + \hat{\Psi}_{a1} \left(X_a - \sum_{i=1}^{n_a} d_{i,a}x_{i,a} \right) + \hat{\Psi}_{a2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a}x_{i,a} \right)^2 \quad (22)$$

where,

$$\hat{\Psi}_{a1} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a}e_{i,a} - d_{j,a}e_{j,a}) (d_{i,a}q_{i,a}x_{i,a}e_{i,a} - d_{j,a}q_{j,a}x_{j,a}e_{j,a})}{\sum_{i=1}^{n_a} d_{i,a}q_{i,a}x_{i,a}^2} \quad (23)$$

$$\hat{\Psi}_{a2} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a}e_{i,a} - d_{j,a}e_{j,a}) (d_{i,a}q_{i,a}x_{i,a}e_{i,a} - d_{j,a}q_{j,a}x_{j,a}e_{j,a})}{2 \left(\sum_{i=1}^{n_a} d_{i,a}q_{i,a}x_{i,a}^2 \right)^2} \quad (24)$$

Special cases 1: Value of the Equation (18) under the SRSWOR

$$\pi_{i,a} = \pi_{j,a} = \frac{n_a}{N_a} \text{ and } \pi_{ij,a} = \frac{n_a(n_a - 1)}{N_a(N_a - 1)} \quad (25)$$

$$\hat{V}_L(\hat{Y}_{YG,1,a}) = \frac{N_a^2(1 - f_a)}{n_a(n_a - 1)} \sum_{i=1}^{n_a} e_{i,a}^2 \quad (26)$$

where, $e_{i,a} = y_{i,a} - \hat{\beta} x_{i,a}$ and $f_a = \frac{n_a}{N_a}$

$$\hat{V}_L(\hat{Y}_{YG,2,a}) = \frac{N_a^2(1 - f_a)}{n_a(n_a - 1)} \sum_{i=1}^{n_a} e_{i,a}^2 \left(\frac{X_a}{\hat{X}_a} \right)^2, \text{ is a special case of (22)} \quad (27)$$

where, $e_{i,a} = y_{i,a} - \frac{\bar{y}_a}{\bar{x}_a} x_{i,a}$ and $\hat{X}_a = \frac{N_a}{n_a} \sum_{i=1}^{n_a} x_{i,a}$.

The estimator $\hat{V}_L(\hat{Y}_{YG,2,a})$ is a special case of the estimator of variance of the ratio estimator of domain total (Deng and Wu, 1987)

$$\hat{V}_L(\hat{Y}_{YG,3,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 \left(\frac{X_a}{\hat{X}_a} \right)^g, \text{ for } g=2 \quad (28)$$

where, $e_{i,a} = y_{i,a} - \hat{\beta} x_{i,a}$ and expand $\left(\frac{X_a}{\hat{X}_a} \right)^g$ and neglecting third and

higher, we have

$$\left(\frac{X_a}{\hat{X}_a} \right)^g = \left[1 + g \left(\frac{X_a}{\hat{X}_a} - 1 \right) + g(g-1) \left(\frac{X_a}{\hat{X}_a} - 1 \right)^2 + \dots \right] \quad (29)$$

If $\left(\frac{X_a}{\hat{X}_a} \right) = 1$, it reduces into linear form of the class of estimators (28)

$$\hat{V}_L(\hat{Y}_{YG,4,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 \quad (30)$$

Which is similar to the estimator $\hat{V}_L(\hat{Y}_{YG,1,a})$ in this situation there is no choice of an improved efficiency of the estimation of variance of the ratio and regression estimator and there is no any other option of $q_{i,a}$ which reduces to (7). Estimation of variance of GREG of domain total is obtained which is near to a class of estimators have been proposed (Deng and Wu, 1987). This idea is also applied for estimation of domain character. The next section devoted to the high level calibration approach which is consist of the low level calibration approach as a special case. The efficiency of the high level calibration estimation is more preferable than the low level calibration approach.

3.2 Estimation of Variance of GREG Using High Level Calibration Approach

Calibration approach is used to estimate the variance of the GREG estimator for (9). The weight $D_{ij,a}$ using (Yates and Grundy, 1953) for estimation of variance for domain total is given in the Equation (12) are calibrated such as

the estimator of variance for the auxiliary character is exact variance. The proposed estimator for domain total is given by

$$\hat{V}(\hat{Y}_{PD,DS,a}) = \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} \Omega_{ij,a} (w_{i,a} e_{i,a} - w_{j,a} e_{j,a})^2 \quad (31)$$

where, $\Omega_{ij,a}$ is the modified weights which is related to (Yates and Grundy, 1953) form of the estimator and very close to $D_{ij,a}$ with respect to the calibration equation is given by

$$\frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} \Omega_{ij,a} (d_{i,a} e_{i,a} - d_{j,a} e_{j,a})^2 = \hat{V}_{YG}(\hat{X}_{HT,a}) \quad (32)$$

where

$$\hat{V}_{YG}(\hat{X}_{HT,a}) = \frac{1}{2} \sum_{i=1}^{N_a} \sum_{j=1}^{N_a} (\pi_{i,a} \pi_{j,a} - \pi_{ij,a}) (d_{i,a} x_{i,a} - d_{j,a} x_{j,a})^2$$

Represent the known variance of the estimator of the auxiliary total

$X_a : \left(\sum_{i=1}^{N_a} x_{i,a} \right)$ is given by $\hat{X}_{HT,a} = \sum_{i=1}^{N_a} d_{i,a} x_{i,a}$ on right hand side. To

compute the value of (17), we need to either information on every unit of the auxiliary character for the interest domain or it may be obtained by the past survey or specific pilot survey, census register or administrative records. The known variance of the auxiliary character has been discussed (Das and Tripathi, 1978), (Singh and Srivastava, 1980), (Srivastava and Jhajj, 1980, 1981) and (Isaki, 1983). (Singh, et al., 1995) have reviewed class of estimators for unknown population parameters, making use of the known variance of auxiliary character. The idea of an adjustment of weight $D_{ij,a}$ has discussed (Fullar, 1970) via regression type estimation tools. I restrict ourselves to the two dimensional chi-square type distance D, of $n_a \times n_a$ grids formed by the new domain weight $\Omega_{ij,a}$ and $D_{ij,a}$ for $i, j = 1, 2, \dots, n_a$, is given by

$$D_a = \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} \frac{(\Omega_{ij,a} - D_{ij,a})^2}{D_{ij,a} Q_{ij,a}} \quad (33)$$

For a^{th} domain $Q_{ij,a} = 1$, it reduced to linear regression estimator and

$Q_{ij,a} = \frac{1}{x_1}$, reduce to ratio estimator, also some other types of choice can be

possible. My aim is to minimize (33) under the restriction that (32), the then modified weights for the a^{th} domain is given by

Estimation of Variance for Domain Total:

A High Level Calibration Approach

$$\Omega_{ij,a} = D_{ij,a} + P_a \left[V_{YG}(\hat{X}_{HT,a}) - \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^2 \right] \quad (34)$$

$$\text{where, } P_a = \frac{(d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^2}{\frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} Q_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^4}$$

For optimum choice of the Lagrange Multiplier λ_a , given as

$$\lambda_a = \frac{V_{YG}(\hat{X}_{HT,a}) - \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^2}{\frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} Q_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^4} \quad (35)$$

Now, substitute the value of $\Omega_{ij,a}$ from (34) in (31) then regression estimator is given by

$$\hat{V}_{SS}(\hat{X}_{GREG,a}) = \hat{V}_{YG}(\hat{X}_{DS,a}) + \phi_{1,a} [V_{YG}(\hat{X}_{HT,a}) - \hat{V}_{YG}(\hat{X}_{HT,a})] \quad (36)$$

Where,

$$\hat{\phi}_{1,a} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} Q_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^2 (w_{i,a}e_{i,a} - w_{j,a}e_{j,a})^2}{\frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} Q_{ij,a} (d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^4} \quad (37)$$

The special cases of the higher level calibration are given as follows:

Case 3.2.1: If $q_{i,a} = \frac{1}{x_{i,a}}$, low level calibration weights and

$Q_{ij,a} = \frac{1}{(d_{i,a}x_{i,a} - d_{j,a}x_{j,a})^2}$, high level calibration weights for a^{th} domain

and the proposed estimator reduced to

$$\hat{V}_H(\hat{Y}_{YG,1,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 \left(\frac{X_a}{\hat{X}_a} \right)^2 \left(\frac{S_{x,a}^2}{s_{x,a}^2} \right) \quad (38)$$

where, $s_{x,a}^2 = \frac{1}{(n_a-1)} \sum_{i=1}^{n_a} (x_{i,a} - \bar{x}_a)^2$ and $S_{x,a}^2 = \frac{1}{(N_a-1)} \sum_{i=1}^{N_a} (x_{i,a} - \bar{X}_a)^2$, I

know sample variance is an unbiased estimator for domain variance.

Case 3.2.2: If $q_{i,a} = 1$, low level calibration weights and $Q_{ij,a} = 1$, for all i and j for each domain and I got

$$\hat{V}_H(\hat{Y}_{YG,2,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 + \hat{\psi}_{1,a}(X_a - \hat{X}_a) + \hat{\psi}_{2,a}(X_a - \hat{X}_a)^2 + \hat{\psi}_{3,a}(S_{x,a}^2 - s_{x,a}^2) \quad (39)$$

where,

$$\hat{\psi}_{1,a} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a} e_{i,a} - d_{j,a} e_{j,a}) (q_{i,a} q_{i,a} x_{i,a} e_{i,a} - d_{j,a} q_{i,a} x_{i,a} e_{j,a})}{\sum_{i=1}^{n_a} d_{i,a} q_{i,a} x_{i,a}^2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right) \quad (40)$$

$$\hat{\psi}_{2,a} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (q_{i,a} q_{i,a} x_{i,a} e_{i,a} - d_{j,a} q_{i,a} x_{i,a} e_{j,a})^2}{2 \left(\sum_{i=1}^{n_a} d_{i,a} q_{i,a} x_{i,a}^2 \right)^2} \left(X_a - \sum_{i=1}^{n_a} d_{i,a} x_{i,a} \right)^2 \quad (41)$$

$$\hat{\psi}_{3,a} = \frac{N_a^2(1-f_a)}{n_a \sum_{i=1}^{n_a} (x_{i,a} - x_{j,a})^4} \left[\left\{ (x_{i,a} - x_{j,a}) (e_{i,a} - e_{j,a}) + \frac{(X_a - \hat{X}_a) (x_{i,a} - x_{j,a})^2}{\sum_{i=1}^{n_a} x_{i,a}^2} \right\}^2 \right] \quad (42)$$

From the equations (38) and (39), it is clear that the high level calibration estimator is not same as given by Deng and Wu (1987). These estimators are member of a class of estimators for estimating variance of the GREG has been discussed (Srivastava and Jhajj, 1981)

$$\hat{V}_{SJ}(\hat{Y}_{YG,2,a}) = \left(\frac{N_a^2(1-f_a)}{n_a(1-n_a)} \sum_{i=1}^{n_a} e_{i,a}^2 \right) L \left(\frac{X_a}{\hat{X}_a}, \frac{S_{x,a}^2}{s_{x,a}^2} \right) \quad (43)$$

where, $L(\cdot, \cdot)$ is a parametric function such that $L(1,1)=1$ and satisfied certain regularity conditions. The Equation (43) is superior to the Equation (17), (Srivastava and Jhajj, 1981)), the Equation (21) and (Deng and Wu, 1987).

4. Class of Estimators of Domain Total

I consider $u = X_a / \sum_{i=1}^{n_a} d_{i,a} x_{i,a}$ and $v = V(\hat{X}_{HT,a}) / \hat{V}(\hat{X}_{HT,a})$, a class of estimator is given as:

$$\hat{V}_{SS}(\hat{Y}_{YG,2,a}) = \left\{ \frac{1}{2} \sum_{i=1}^{n_a} \sum_{j=1}^{n_a} D_{ij,a} (d_{i,a} e_{i,a} - d_{j,a} e_{j,a})^2 \right\} L(u, v). \quad (44)$$

where $L(u, v)$ is a parametric function of u and v for example $L(1,1) = 1$ and which pose the following regularity condition:

1: The function $L(u, v)$ exists for all the values of (u, v) and it contains the

points $\left(X_a / \sum_{i=1}^{n_a} d_{i,a} x_{i,a}, V(\hat{X}_{HT,a}) / \hat{V}(\hat{X}_{HT,a}) \right)$ in a bounded subset D of two

dimensional real spaces.

2: The first and second order partial derivatives of $L(u, v)$ exists and are continuous and bounded.

There are so many members of the class of estimator which pose the regularity conditions. But here some member of the class of estimators for domain total which exist are given as

$$L(u, v) = u^\alpha v^\beta, \quad L(u, v) = \frac{1 + \alpha(u-1)}{1 + \beta(v-1)}, \quad L(u, v) = 1 + \alpha(u-1) + \beta(v-1) \text{ and}$$

$L(u, v) = \{1 + \alpha(u-1) + \beta(v-1)\}^{-1}$ are special cases of the higher level calibration approach where, α and β are unknown parameters involved in the function $L(u, v)$. In each of the estimators value of the constants α and β are different in the Equation (44) leads to the asymptotic same variance has shown by (Srivastava and Jhajj, 1983) and (Singh and Singh, 1984). The extension of the present estimators using two phase sampling has been discussed by (Hidiroglou and Sarndal, 1995). But, here I consider for estimation of the domain total. The next section is devoted to the studied the performance of the proposed estimator using both calibration approaches through comparative study.

5. Comparative Study

Comparative study via theoretically is also a view to show the performance of the proposed calibration estimator of domain total, the variance of the high level calibration approach and the low level calibration approach of the ratio estimator from equations (38) and (42) for a^{th} domain are given as:

$$\hat{V}_H(\hat{Y}_{YG,1,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 \left(\frac{X_a}{\hat{X}_a} \right)^2 \left(\frac{S_{x,a}^2}{s_{x,a}^2} \right) \quad (45)$$

and

$$\hat{V}_L(\hat{Y}_{YG,1,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 \left(\frac{X_a}{\hat{X}_a} \right)^2 \quad (46)$$

$$\text{If, } \left(\frac{S_{x,a}^2}{s_{x,a}^2} \right) < 1 \quad (47)$$

$$\text{i.e. } \hat{V}_H(\hat{Y}_{YG,1,a}) < \hat{V}_L(\hat{Y}_{YG,1,a}) \quad (48)$$

The comparison between low level calibration approach and high level calibration approach of the regression estimator from the Equation (39) and (41) for a^{th} domain is given as:

$$\hat{V}_H(\hat{Y}_{YG,2,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 + \hat{\psi}_{1,a}(X_a - \hat{X}_a) + \hat{\psi}_{2,a}(X_a - \hat{X}_a)^2 + \hat{\psi}_{3,a}(S_{x,a}^2 - s_{x,a}^2) \quad (49)$$

And

$$\hat{V}_L(\hat{Y}_{YG,2,a}) = \frac{N_a^2(1-f_a)}{n_a(n_a-1)} \sum_{i=1}^{n_a} e_{i,a}^2 + \hat{\psi}_{1,a}(X_a - \hat{X}_a) + \hat{\psi}_{2,a}(X_a - \hat{X}_a)^2 \quad (50)$$

From Equation (47), $(S_{x,a}^2 - s_{x,a}^2) = -ve$

$$\text{i.e. } \hat{V}_H(\hat{Y}_{YG,2,a}) < \hat{V}_L(\hat{Y}_{YG,2,a}) \quad (51)$$

Hence, High level calibration is a better choice than the low level calibration approach for both ratio and regression estimator of a^{th} domain total.

6. Simulation Study

For the purpose of simulation study, I have taken a real data set (Sarndal et al. (1992), in appendix B). In which I take only four geographical areas (domains) (1, 2, 7 and 6) out of the eight geographical areas (domains) 1, 2, 3, 4, 5, 6, 7 and 8. The study domains have sizes 25, 48, 15 and 29. I select a sample of approximately 20% units from each domain by study and auxiliary character y and x. The variance of the estimator for domain total of low level calibration approach and high level calibration approach of the ratio estimator and generalized estimator for a^{th} domain total are obtained which are given in the equations (45), (46), (49) and (50) section of the proposed estimator.

Study character and auxiliary character are given as:

y= Population 1985 (in thousands), x= Population 1975 (in thousands)

Table 1: The value of the variance of the ratio $\hat{V}_L(\hat{Y}_{YG,1,a})$, $\hat{V}_H(\hat{Y}_{YG,1,a})$, and regression estimators $\hat{V}_L(\hat{Y}_{YG,2,a})$, $\hat{V}_H(\hat{Y}_{YG,2,a})$ for domains (1, 2, 3 and 4):

Estimator	Domain Values			
	1	2	3	4
$\hat{V}_L(\hat{Y}_{YG,1,a})$	47150.3	568342.3	997231.3	127026.5
$\hat{V}_H(\hat{Y}_{YG,1,a})$	42210.6	589590.1	970261.2	110055.2
$\hat{V}_L(\hat{Y}_{YG,2,a})$	49521.4	591323.4	1004598.7	137026.5
$\hat{V}_H(\hat{Y}_{YG,2,a})$	45356.7	582127.6	984123.4	121037.9

It is obtained from the table 1 that the variance of the ratio estimator using high level calibration is less than low level calibration approximately (3% to 15%) for the domains. It is also seen, that the variance of the regression estimator using high level calibration is less than the low level calibration approximately (2% to 13%) for the domains (1, 2, 3 and 4). Hence, high level calibration approach is better than the low level calibration approach for both the estimators.

7. Results and Discussion

The important points are obtained which are given as follows:

1. It is obtained that from the equations (26) and (27), low level ratio calibration estimator is lower than the low level regression estimator of a^{th} domain and from the equations (48) and (51) high level ratio calibration estimator for domain total is lower to high level regression estimator for domain total.

2. It is seen that high level calibration approach is better than low level calibration for both the estimator ratio and regression estimator for domain total.

3. It is obtained that from real data using simulation study, theoretical point is justified through simulation study. The variance of the high level calibration approach is less than the low level calibration approach for all domains.

4. It is also seen that proposed estimator is a better choice than the (Deng and Wu, 1987) estimator for domain mean and the proposed estimator for domain mean is a special case of the GREG by (Srivastava and Jhajj, 1981) given in the Equation (43).

8. Conclusion

The High level calibration approach is used when the variance of each domain of the auxiliary character is known. The proposed direct generalized calibration estimator for domain total gives good option when sufficient number of the units present in the interested domain. It is option that high level calibration approach is better option to low level calibration via theoretically and empirically.

The proposed calibration approach may apply for domains problem such as health related problem (COVID-19), environmental problem and welfare program in the deprived area of the country.

Acknowledgements

I am very thankful to Prof. K. K. Singh provided a project, SAP-DRS II at Department of Statistics, Banaras Hindu University, India which help me to imminence contribution in the domain parameter estimation.

References

1. Das, A. K., and Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya*, 40(C): 139-148.
2. Deng, L. Y. and Wu, C. F. J. (1987). Estimation of variance of the regression estimator. *Journal of the American Statistical Association*, 82: 568-576.
3. Deville, J. C. and Sarndal, C. E. (1992). Estimation of variance of the regression estimator. *Journal of the American Statistical Association*, 82: 568-576.
4. Deville, J. C. and Sarndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87: 376-382.
5. Fuller, W. A. (1970). Sampling with random stratum boundaries. *Journal of the Royal Statistical Society*, 32: 209-226.
6. Hidiroglou, M. A. and Sarndal, C. E. (1995). Use of auxiliary information for two-phase sampling. *Proceeding of the Section on Survey Research Methods, American Statistical Association*, Volume II: 873-878.
7. Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*.
8. Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78(381): 117-123.
9. Rao, J. N. K. (2003): *Small Area Estimation*. New York: Wiley and Sons.
10. Sarndal, C. E. Swensson, B. and Wretman, J. H. (1992). *Model assisted survey sampling*, Springer-Verlog: New York.
11. Singh, S, Horn, S and Yu, F. (1998). Estimation of variance of general regression estimator, high level calibration approach. *Survey Methodology*, 24(1): 41-50.
12. Singh, R. K. and Singh, S. (1984). A class of estimators with estimated optimum values in sample surveys. *Statistics & Probability Letters*, 2: 319-321.
13. Singh, P. and Srivastava, S. K. (1980). Sampling scheme providing unbiased regression estimators. *Biometrika*, 67: 205-209.
14. Srivastava, S. K. and Jhaji, H. S. (1981). A class of estimators of the population mean survey sampling using auxiliary information. *Biometrika*, 68: 341-343.
15. Srivastava, S. K. and Jhaji, H. S. (1983). A class of estimators of estimators of the population mean using multi-auxiliary information. *Calcutta Statistical Association Bulletin*, 32: 47-56.

16. Tikkiwal et al. (2013). On the performance of generalized regression estimator for small domains. *Communication in Statistics-Simulation and Computation*, 42: 891-909.
17. Wu, C. F. J. (1985). Variance estimation for combined ratio and combined regression estimators. *Journal of the Royal Statistical Society*, 47(B): 147-154.
18. Yates, F. and Grundy, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society*, 15(B): 253-261.