ISAS

Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 77(1) 2023 11–18

Robust Estimation in Finite Population Sampling under Model based Prediction Approach

B.V.S. Sisodia[1] and R.P. Kaushal[2]

[1]Acharya Narendra Deva University of Agriculture and Technology, Ayodhya
[2]SDJ PG College, Chandeshwar, Azamgarh

Received 10 April 2023; Revised 22 April 2023; Accepted 29 April 2023

SUMMARY

In the present paper, some important contributions on estimation of finite population parameters under model-based prediction approach are reviewed. Following Scott *et al.* (1978), a BLU predictor of population total in model-based prediction approach under the model $\xi(0,1{:}x_{hk}^2)$ in stratified sampling when the slope of the model is common across the strata is constructed. Its robustness and optimality is studied when some general polynomial model of degree j, i.e., $\xi(\delta_0,\delta_1,.....,\delta_J : x_{hk}^2)$ is true in real practice. It has been found that the proposed predictor is robust and optimal for stratified balanced sample and it is also more efficient than the predictor developed by Scott *et al.* (1978) when slopes are uncommon across the stratum.

Keywords: Super population models; Stratified sampling; Balanced sample; Overbalanced sample; General polynomial models.

1. INTRODUCTION

In finite population survey sampling, the aim is to estimate the finite population parameters such as population mean, population total etc. There are two fundamentally different approaches to finite population sampling theory. First one is based on sampling design in which use of probability distribution generated by random sampling plan P is the basis for inference. Second approach is referred to as model based prediction approach (Royall, 1970, 71) which depends on the assumption of a super-population model ξ (say) that plays an essential role in finite population inference. Another third approach is a hybrid of the two approaches mentioned above, which is referred to as model-assisted approach (Cassel *et al.*, 1976; 1977; 1979; Sarndal *et al.*, 1992), where model-based and design-based & principles are combined for finite population inference. In the present paper, we will deal with model based prediction approach for finite population sampling theory. So, firstly we brief review some of important contributions on model-based prediction approach in this section.

Royall (1970) was first person who advocated model-based prediction approach for finite population sampling theory. He considered the following super-population model

$$Y_k = \beta x_k + \varepsilon_k \left[v\left(x_k\right)\right]^{1/2}, k = 1,2,\ldots,N$$

$$V\left(Y_k\right) = \sigma^2 v\left(x_k\right), E_\xi\left(\varepsilon_k\right) = 0. \tag{1.1}$$

where, Y_k's (k = 1, 2, …, N) are independent random variables, $v(x_k)$ is a variance function of x_k, σ^2 and β are parameters of the model. Associated with each k^th unit are pair of number (y_k, x_k), where y_k is unknown and x_k is known quantity being auxiliary variable, y_k is considered as realised value of Y_k. Thus, the objective is to estimate $T = \sum_{k=1}^{N} y_k$. For this, a sample s is selected from the population under study. He showed that for

Corresponding author: B.V.S. Sisodia
E-mail address: bvssisodia@gmail.com

given sample s, model-based best linear unbiased estimator (BLUE) of T is

$$\hat{T}[0,1:v(x_k)] = \sum_{s} y_k + \hat{\beta}\sum_{\bar{s}} x_k \qquad (1.2)$$

Where $\hat{\beta} = \dfrac{\sum_{s} y_k x_k / v(x_k)}{\sum_{s} x^2{}_k / v(x_k)}$, which is BLUE of β

and $\bar{s}$ contains the units not sampled, i.e., complement of s. The model (1.1) is referred to as ξ-model, and it is denoted as $\xi\left[0,1:v\left(x_k\right)\right]$. For $\mathrm{v}\left(\mathrm{x}_\mathrm{k}\right)=\mathrm{x}_\mathrm{k}$, the estimator in (1.2) reduces to $\hat{T}_1\left[0,1:x_k\right]$ and it can be verified that it is the usual ratio estimator for a given sample s.

In fact, the approach is referred to as prediction approach because the second term of the estimator given in equation (1.2) is simply a prediction of $\sum_{\bar{s}} y_k$ based on fitting of the model (1.1) by least square technique using data in sample s.

The model-variance of the estimator in equation (1.2) is

$$V\left[\hat{T}\left\{0,1:v(x_k)\right\}\right] = E_{\xi}[\hat{T}\left\{0,1:v(x_k)\right\} - T]^2$$

$$= \sigma^2\left[\frac{\left(\sum_{\bar{s}} x_k\right)^2}{\sum_{s} x^2{}_k / v(x_k)} + \sum_{\bar{s}} v(x_k)\right] \qquad (1.3)$$

For $\mathrm{v}\left(\mathrm{x}_\mathrm{k}\right)=\mathrm{x}_\mathrm{k}$, the $\hat{T}_1\left[0,1:x_k\right]$ is the usual ratio estimator and its model variance is

$$V\left[\hat{T}_1(0,1:x_k)\right] = \sigma^2 \frac{\sum_{\bar{s}} x_k}{\sum_{s} x_k} \sum_{1}^{N} x_k \qquad (1.4)$$

The variance given in equation (1.4) can be further minimised by choosing a sample consisting of those n units whose x-values are largest, i.e., $\sum_{s} x_k$ attains its maximum value. Such sample is referred to as optimal sample denoted as $s(0,1:x)$. Thus, the estimator $\hat{T}_1\left[0,1:x_k\right]$ is the optimal estimator under model $\xi\left[0,1:x_k\right]$ for the optimal sample $s(0,1:x)$. Royall (1971) further demonstrated that mean-square error (MSE) of the usual ratio estimator under simple random sampling without replacement is inferior in many applications to the model-variance of the usual ratio estimator under super population model $\xi\left[0,1:x_k\right]$.

However, the correctness of the results in the prediction approach depends on the validity of the assumed super-population model. Royall & Herson (1973a) studied the robustness of the predictor/estimator if the assumed model fails to hold. Suppose that the assumed model does not hold true and if some other model, i.e., a general polynomial model of degree J

$$Y_k = \delta_0\beta_0 + \delta_1\beta_1 x_k + \delta_2\beta_2 x_k^2 + ... + \delta_J\beta_J x_k^J + \varepsilon_k\left[v\left(x_k\right)\right]^{1/2},$$

$$k = 1,2,...,N$$

$$j = 0,1,2,...,J \qquad (1.5)$$

holds true, where $\delta_j = 1$ if j$^{\text{th}}$ component (j = 0, 1, 2, ..., J) appears in the model, otherwise, zero.. The model (1.5) is denoted as $\xi[\delta_0,\delta_1,\delta_2,...,\delta_J : v(x_k)]$. They showed that the estimator $\hat{T}\left[0,1:x_k\right]$ is robust and BLUE even under model $\xi[\delta_0,\delta_1,\delta_2,...,\delta_J : v(x_k)]$, for $v\left(x_k\right) = \sum_{j=1}^{J} \delta_j\, a_j\, x_k^j$, a_0, a_1, …, a_J being some constant, provided that all the terms which appear in the variance function v(x$_k$) must also appear in the general polynomial model and sample selected is balanced one, i.e., $\bar{x}_s{}^{(j)} = \bar{x}^{(j)}$ for all j = 1, 2, ..., J, where $\bar{x}^{(j)} = \sum_{k=1}^{N} x_k^j \Big/ N$. For positive integer J, let s(J) denote any sample satisfying $\bar{x}_s{}^{(j)} = \bar{x}^{(j)}$ for j= 1, 2, …, J, then s(J) is referred to as balanced sample. However, they remarked that the cost of this insurance (balanced sample) could be quite high.

Royall and Herson (1973b) further suggested an alternative strategy, i.e., stratified sampling with separate ratio estimator within the strata, which could offer same protection at somewhat low cost. They suggested stratification of the population on the basis of size variable(x) and use of balance sampling along with separate ratio estimator of T. The strata are formed as follows: the N$_1$ units whose x-values are smallest form stratum 1, the next N$_2$ smallest units form stratum 2, and so on, such that $\sum_{h=1}^{H} N_h = N$. Thus, no unit in h$^{\text{th}}$ stratum is larger than any unit in (h+1)$^{\text{th}}$ stratum. Assume that the working model is $\xi\left[0,1:v\left(x_{hk}\right)\right]$, i.e.,

$$Y_{hk} = \beta_h x_{hk} + \varepsilon_{hk}[v(x_{hk})]^{1/2}, h = 1,2,...,H; k = 1,2,...,N_h \qquad (1.6)$$

$$V\left(Y_{hk}\right) = \sigma^2 v(x_{hk}),\ E\ (\varepsilon_{hk}) = 0$$

The model based unbiased estimator of $T=\sum_{h=1}^{H}\sum_{k=1}^{N_h} y_{hk}$ for given sample s_h of size n_h units from h^{th} stratum is given by

$$\hat{T}_{st}=\sum_{h=1}^{H}\hat{T}_h[0,1:v(x_{hk})] \tag{1.7}$$

Where

$$\hat{T}_h[0,1:v(x_{hk})]=\sum_{s_h} y_{hk}+\frac{\sum_{s_h} y_{hk}x_{hx}/v(x_{hk})}{\sum_{s_h} x_{hk}^2/v(x_{hk})}\sum_{\bar{s}_h} x_{hk} \tag{1.8}$$

For $v(x_{hk})=x_{hk}$, $\hat{T}_{st}$ become $\hat{T}'_{st}=\sum_{1}^{H}\hat{T}_h(0,1:x_{hk})$ (1.9)

and $\hat{T}_h[0_1 1\ x_{hk}]=\frac{\sum_{s_h} y_{hk}}{\sum_{s_h} x_{hk}}\sum_{k=1}^{N_h} x_{hk}$

Variance of $\hat{T}'_{st}$ is given model as

$$V\left(\hat{T}'_{st}\right)=\sigma^2\sum_{h=1}^{H}\left(\frac{\sum_{\bar{s}_h} x_{hk}}{\sum_{s_h} x_{hk}}\right)\sum_{1}^{N_h} x_{hk} \tag{1.10}$$

If the working model is $\xi(0,1:x_{hk})$ and true model is $\xi[\delta_0,\delta_1,....\delta_J:v(x_{hk})]$, then the estimator $\hat{T}'_{st}$ given in eqn. (1.9) under this true model is biased one. If $\bar{x}_{sh}^{(j)}=\bar{x}_h^{(j)}$ (stratum mean of degree j) for j = 1, 2, … J and for all h = 1,2, … H, then such sample is referred to as stratified balanced sample denoted by s* (J). For s* (J), $\hat{T}'_{st}$ is unbiased and its variance in eqn. (1.10) reduces to

$$V\left(\hat{T}^{'}_{st}\right)=\sigma^2\sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(1-\frac{n_h}{N_h}\right)\bar{x}_h \tag{1.11}$$

The variance expression in eqn. (1.11) is applicable to any polynomial regression model of degree J or less than J if $v(x_{hk})=x_{hk}$.

It has been shown by Royall & Herson (1973b) that the strategy $[\hat{T}^{'}_{st},s*(J)]$ is more efficient than the strategy $[\hat{T}_1,s(J)]$.

For a fixed cost $C=C_O+\sum_{h=1}^{H}c_h n_h$, where Co is overhead cost and c_h is a cost of remuneration of each unit of the sample s_h in stratum h, the variance in eqn. (1.11) is minimised, and we get

$$n_h=n\frac{N_h(\bar{x}_h/c_h)^{1/2}}{\sum_{h=1}^{H}N_h(\bar{x}_h/c_h)^{1/2}} \tag{1.12}$$

If c_h = c for h = 1, 2, … H, then eqn. (1.12) reduces to

$$n_h=n\frac{N_h(\bar{x}_h)^{1/2}}{\sum_{h=1}^{H}N_h(\bar{x}_h)^{1/2}} \tag{1.13}$$

The above allocation in eqn. (1.13) is referred to as an optimum allocation, and optimum variance under this optimum allocation is given by

$$V(\hat{T}'_{st})_o=\sigma^2\left[\frac{\left\{\sum_{1}^{H}N_h(\bar{x}_h)^{1/2}\right\}^2}{n}-N\bar{x}\right] \tag{1.14}$$

Royall and Herson (1973b) proved a theorem which is stated here without proof

Theorem 1.1: If $n_h=n\frac{N_h(\bar{x}_h)^{1/2}}{\sum_{1}^{H}N_h(\bar{x}_h)^{1/2}}$, then under the model $\xi(\delta_0,\delta_1,...,\delta_J:x_{nk})$ the strategy $[\hat{T}'_{st},s^*(J)]$ is efficient than the strategy [$\hat{T}_1$ (0,1:x), s(J)]

Royall (1976) extended the prediction approach to two stage sampling and developed optimal (BLU) estimator and its variance under general linear super population model.

Scott *et al.* (1978) extended the work of Royall & Herson (1973a) by considering the super population model denoted by $\xi[0,1:x^2]$ which is described by

$$Y_k=\beta x_k+\varepsilon_k x_k,\ \text{k = 1, 2,…, N} \tag{1.15}$$

$$E_\xi(Y_k)=\beta x_k,V(Y_k)=\sigma^2 x_k^{\ 2},E_\xi(\varepsilon_k^{\ 2})=\sigma^2$$

For a given sample s of size n regardless of the manner in which the sample is drawn from finite population of size N, the ξ-best linear unbiased (BLU) estimator of $T=\sum_{k=1}^{N}y_k$ under the model (1.15) is

$$\hat{T}_2(0,1:x^2)=\sum_s y_k+\left(\frac{1}{n}\sum_s\frac{y_k}{x_k}\right)\sum_{\bar{s}}x_k \tag{1.16}$$

with its model variance

$$V\left[\hat{T}_2(0,1:x^2)\right] = E_\xi\left[\hat{T}_2\left(0,1:x^2\right) - T\right]^2$$

$$= \sigma^2\left[\sum_{\bar{s}} x_k^2 + \frac{1}{n}\left(\sum_{\bar{s}} x_k\right)^2\right] \qquad (1.17)$$

Let $s = s'(J)$ be a particular sample for which

$$\frac{\sum_{\bar{s}} x_k^j}{\sum_{\bar{s}} x_k} = \frac{\sum_{s} x_k^{j+1}/v(x_k)}{\sum_{s} x_k^2/v(x_k)} \quad j = 0,\ 1\ 2,\ \ldots.\ J. \qquad (1.18)$$

The sample s' (J) was referred to as overbalanced sample by Scott, *et al.* (1978). They proved that for s = s'(J), $\hat{T}$ [0,1 : v*(x)] is model based BLU estimator under the model $\xi[\delta_0, \delta_1, ...\delta_J : v^*(x)]$ for any variance function of the form $v^*(x) = v(x)\sum_{j=0}^{J} \delta_j\, a_j\, x^{j-1}$, where a_j's are some positive constants. They also showed that for a wide class of models, $\hat{T}[0,1:v(x)]$ is in fact the BLU estimator when s = s' (J).

Note that, the condition (1.18) reduces to the case of balanced sample, i.e. $\bar{x}_s^{(j)} = \bar{x}_{\bar{s}}^{(j)}$, j= 1,2,...,J, which in denoted as $s = s(J)$ [Royall & Herson, 1973 a].

For v(x)= x^2, the condition (1.18) for overbalanced sample s'(J) reduces to

$$\frac{\sum_{\bar{s}} x_k^j}{\sum_{\bar{s}} x_k} = \frac{\sum_{\bar{s}} x_k^{j-1}}{n},\ j = 0,\ 1,\ 2,\ \ldots,\ J. \qquad (1.19)$$

Note that for overbalanced samples s' (J) satisfying the eqn. (1.19) the variance of $\hat{T}_2\left(0,1:x^2\right)$ is given by

$$V\left[\hat{T}_2\left(0,1:x^2\right)\right] = \frac{\sigma^2 N\bar{x}}{n}\sum_{\bar{s}} x_k \qquad (1.20)$$

It has been further shown by Scott *et al.* (1978) that under the model $\xi\left[\delta_0, \delta_1, \ldots, \delta_J : v(x)\right]$ with v(x) = $\sigma_1^2 x + \sigma_2^2 x^2$, both $\hat{T}_1$ with balanced sampling and $\hat{T}_2$ with overbalanced sampling are BLU for their respective samples, and $\hat{T}_2$ is more efficient than $\hat{T}_1$.

Scott *et al.* (1978*)* also extended their work in stratified sampling under the model (1.6) with $v(x_{hk}) = x_{hk}^2$. Therefore, the estimator of T under this model is given by taking $v(x_{hk}) = x_{hk}^2$ in equation (1.7). Let it be denoted as $\hat{T}''_{st}(0,1:x_{hk}^2)$ and it is given by

$$\hat{T}''_{st}(0,1:x_{hk}^2) = \sum_{1}^{H}\hat{T}_h\,(0,1:x_{hk}^2) \qquad (1.21)$$

where

$$\hat{T}_h(0,1:x_{hk}^2) = \sum_{s_h} y_{hk} + \left(\frac{1}{n}\sum_{s_h} y_{hk}/x_{hk}\right)\sum_{\bar{s}_h} x_{hk}. \qquad (1.22)$$

The model variance of $\hat{T}''_{st}(0,1:x_{hk}^2)$ is given by

$$V[\hat{T}''_{st}(0,1:x_{hk}^2)] = \sigma^2\left[\sum_{h=1}^{H}(N_h - n_h)\bar{x}^2_{\bar{s}_h} + \sum_{h=1}^{H}\left\{(N_h - n_h)\bar{x}_{\bar{s}_h}\right\}^2/n_h\right] \qquad (1.23)$$

Tam (1986) developed a linear predictor $T = \sum_{1}^{N} y_k$ by considering a model

$$E\ (y) = X\beta,\ D\ (Y) = \sigma^2 V \qquad (1.24)$$

where, V is positive definite diagonal matrix. Tam (1987) further considered a Gaussian super population model.

$$Y = X\beta + \varepsilon, E(\varepsilon) = 0, D(\varepsilon) = \sigma^2 V \qquad (1.25)$$

where, V is positive definite symmetric matrix. He showed that under this model, model based predictor of Royall (1976) is the unbiased minimum mean square error predictor of the total T. Royall (1992) showed that under the model in eqn. (1.24), $\hat{T}$ (X : V) remained optimal and robust under some different models. Bouza (1994) studied the robustness of shrunken predictors in stratified sampling under the model ξ (0,1:1) in the case of misspecification of the model in terms of whether the slope of model is $\beta_h = \beta$ or $\beta_h \neq \beta$.

Using the results of Royall (1992), Valliant *et al.* (2000) showed that unstratified weighted balanced sample yielded the same variance as stratification by size with optimum allocation of stratified weighted balanced sample when BLU predictor of T is used. However, there results are based on the assumptions of equal variance σ^2 in each stratum and equal cost of the survey for each unit in the sample. Moreover, the composition of the weighted balanced sample and stratified weighted balanced sample leading to the overall weighted balanced sample cannot be exactly the same in practical situations.

Kaushal *et al.* (2011) have proposed new shrunken estimators in stratified sampling under the model ξ (0,1:1) and studied their properties and robustness on the similar line of Bouza (1994). They found that the proposed new shrunken estimators are more efficient than that due to Bouza (1994). They also conducted a simulation study to show the superiority of the proposed estimators. Sisodia *et al.* (2015) have extended the work of Kaushal *et al.* (2011) in stratified sampling under the model ξ (0,1:x_{hk}). They showed that the properties of the proposed shrunken estimators under model ξ (0,1:x_{hk}) still hold true even under the model ξ [(0,1:v(x_{hk})] for any variance function v(x_{hk}) if sample selected in each stratum is stratified balanced sample.

However, considering the piecewise super-population model (1.6) for each stratum separately in stratified sampling poses the following problems:

It requires to making linear approximation of the true function in piecewise manner.

Although for stratified balanced sample (balanced sample in each stratum) the BLU estimator $\hat{T}_h\left[\delta_0,\delta_{1,},...,\delta_J{:}v(x)\right]$ reduces to expansion estimator $\frac{N_h\sum_{s_h} y_{hk}}{n_h}$ for $v(x)=\sum_{j=0}^{J}\delta_j\, a_j\, x_{hk}^{j}$, where a_j 's are some constants, it still requires to estimate parameters of regression function to estimate the error-variance in each stratum separately for the evaluation of the estimator. It, therefore, results to fitting of H models.

If n_h is small and if it is less than J, estimation of β_j 's are not possible in the hth stratum (h=1, 2, …, H).

To overcome the above problems, it is better to make linear approximation of the true function across the strata. That is a single super-population model across the strata in which regression coefficients are common across the strata. In fact, the way the stratification of the population is carried out on the basis of size-variable x, the specification of single super population model across the strata is meaningful and justifiable. Kaushal and Sisodia (2021) has developed robust and BLU estimator of T under the model ξ (0,1:x_{hk}) in stratified sampling where the slopes are common across the strata. For stratified balanced samples with $n_h \propto N_h\bar{x}_h$, it has been shown that the estimator under ξ (0,1:x_{hk}) with common slopes across the strata is more efficient than that due to Royall & Herson (1973b) where the slopes are different from stratum to stratum under certain conditions which has been shown empirically to exist.

Models of type $\hat{\imath}\left(0,1{:}x^{2\gamma}\right)$ have been used by various investigators including Smith (1938), Jessen (1942), Raj (1958), Rao & Bayless (1969) and Bayless & Rao (1970) for empirical results in finite population sampling. Empirical results obtained by various investigators for wide variety of socio-economic variables showed that the γ value varies between one and half and one.

Therefore, an attempt has been made in the present paper to develop the BLU estimator under the model $\xi\left(0,1:x_{hk}^{2}\right)$ in stratified sampling when the slopes are common across the strata. Robustness and other properties of the estimator are studied. The relative efficiency of the proposed estimator is also examined as compared to the estimator developed by Scott *et al.* (1978) under the model $\xi\left(0,1:x_{hk}^{2}\right)$ when slopes are uncommon across the strata.

2. BEST LINEAR UNBIASED (BLU) ESTIMATOR UNDER SINGLE SUPER-POPULATION MODEL

Consider the super-population model of type $\xi\left(0,1:x_{hk}^{2}\right)$ when the slopes are common across the strata, i.e.

$$Y_{hk}=\beta x_{hk}+\varepsilon_{hk}x_{hk};\ h = 1, 2 \ldots, H,\ k = 1, 2, \ldots, N_h \tag{2.1}$$

V (Y_{hk}) = $\sigma^2 x_{hk}^2$ and $E_\xi(Y_{hk})=\beta\, x_{hk}$

The strata are formed in similar way as mentioned in section-1. Let a sample s_h of size n_h such that $\sum_{h=1}^{H} n_h = n$ be selected from N_h units in hth stratum. The model –based BLU estimator of T under the model (2.1) is constructed as

$$\hat{T}(0,1:x_{hk}^{2})=\sum_{h=1}^{H}\sum_{s_h} y_{hk}+\frac{\sum_{h=1}^{H}\sum_{s_h}(y_{hk}/x_{hk})}{n}\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk} \tag{2.2}$$

$$\text{where, } \frac{\sum_{h=1}^{H}\sum_{s_h}\left(y_{hk}/x_{hk}\right)}{n}=\hat{\beta} \tag{2.3}$$

The model variance of $\hat{T}$ is derived as

$$V\left[\hat{T}\left(0,1:x_{hk}^{2}\right)\right]=E_{\xi}\left[\hat{T}(0,1:x_{hk}^{2})-T\right]^{2}$$

$$= E_{\xi}\left[\sum_{h=l}^{H}\sum_{s_h} y_{hk} + \frac{\sum_{h=l}^{H}\sum_{s_h}(y_{hk}/x_{hk})}{n}\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk} - \sum_{h=1}^{H}\sum_{k=1}^{N_h} y_{hk}\right]^2$$

$$= \sigma^2\left[\frac{1}{n}\left\{\sum_{h=1}^{H}(N_h - n_h)\bar{x}_{\bar{s}_h}\right\}^2 + \sum_{h=1}^{H}(N_h - n_h)\bar{x}^{(2)}_{\bar{s}_h}\right] \quad (2.4)$$

Note that the estimator in equation (2.2) is similar to the design –based combined regression estimator in stratified sampling (See, Sukhatme *et al.*, 1984)

Further, let the samples from each stratum be selected in such a way that

$$\frac{\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk}^{j}}{\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk}} = \frac{\sum_{h=1}^{H}\sum_{s_h} x_{hk}^{j+1}/v(x_{hk})}{\sum_{h=1}^{H}\sum_{s_h} x_{hk}^{2}/v(x_{hk})} \quad (2.5)$$

is satisfied for all j = 0, 1, 2, …, J. This may be referred to as overall stratified overbalanced samples. If $v(x_{hk}) = x_{hk}^2$, then (2.5) reduces to

$$\frac{\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk}^{j}}{\sum_{h=1}^{H}\sum_{\bar{s}_h} x_{hk}} = \frac{\sum_{h=1}^{H}\sum_{s_h} x_{hk}^{j-l}}{n} \quad (2.6)$$

For j = 2, this reduces to

$$\sum_{h=1}^{H}(N_h - n_h)\bar{x}^{(2)}_{\bar{s}_h} = \frac{1}{n}\left(\sum_{h=1}^{H} n_h\bar{x}_{s_h}\right)\sum_{h=1}^{H}(N_h - n_h)\bar{x}_{\bar{s}_h}\hat{T}_2 \quad (2.7)$$

Substituting for $\sum_{h=1}^{H}(N_h - n_h)\bar{x}^{(2)}_{\bar{s}_h}$ from (2.7) into (2.4), and after simplification, we get

$$V\left[\hat{T}\left(0,1:x_{hk}^2\right)\right] = \frac{\sigma^2 N\bar{x}}{n}\sum_{h=1}^{H}(N_h - n_h)\bar{x}_{\bar{s}_h} \quad (2.8)$$

Comparing the variance expression (1.23) and (2.4) for given samples (h = 1, 2, ….H), we get

$$V[\hat{T}''_{st}(0,1:x_{hk}{}^2)] = \sigma^2\left[\sum_{h=1}^{H}\left(\frac{1}{n_h} - \frac{1}{n}\right)V_h^2 - \frac{1}{n}\sum_{h\neq h'=1}^{H}\sum^{H} V_h V_{h'}\right] \quad (2.9)$$

where $V_h = (N_h - n_h)\bar{x}_{\bar{s}_h}$

It implies that the estimator under the model $\xi\left(0,1:x_{hk}^2\right)$ with common slope across the strata is more precise than the estimator under the model $\xi\left(0,1:x_{hk}^2\right)$ with slope varying from stratum to stratum if RHS of eqn.(2.9) is greater than zero, i. e.

$$\sigma^2\left[\sum_{h=1}^{H}\left(\frac{1}{n_h} - \frac{1}{n}\right)V_h^2 - \frac{1}{n}\sum_{h\neq h'=1}^{H}\sum^{H} V_h V_{h'}\right] > 0 \quad (2.10)$$

which may hold true in general. The above result can be summarized in the following theorem:

Hence, we have the following theorem:

Theorem 2.1. The BLU estimator $\hat{T}\left(0,1:x_{hk}^2\right)$ of T under the model $\xi\left(0,1:x_{hk}^2\right)$ with common slope across the strata is more efficient than the BLU estimator $\hat{T}''_{st}$ of T under the model $\xi\left(0,1:x_{hk}^2\right)$ with uncommon slopes across the strata if the inequality (2.10) holds true.

The direct comparison of estimator $\hat{T}_2$ in (1.16) developed without stratification on size variable and $\hat{T}\left(0,1:x_{hk}^2\right)$ is difficult since the two samples are not the same in general. However, when sampling fraction is small in each stratum, $\bar{x}_{\bar{s}_h}$ is expected to be very close to $\bar{x}_h$ but $\bar{x}_{\bar{s}_h} \leq \bar{x}_h$, then variance expression of $\hat{T}\left(0,1:x_{hk}^2\right)$ in (2.4) can be at most

$$V\left[\hat{T}\left(0,1:x_{hk}^2\right)\right] \leq \sigma^2\left[\frac{1}{n}\left\{\sum_{h=1}^{H}(N_h - n_h)\bar{x}_h\right\}^2 + \sum_{h=1}^{H}(N_h - n_h)\bar{x}_h^{(2)}\right]$$

$$= \sigma^2\left[\frac{1}{n}\left\{(N-n)\bar{x}_{\bar{s}}\right\}^2 + (N-n)\bar{x}^{(2)}_{\bar{s}}\right]$$

$$= \sigma^2(N-n)\left[\frac{N-n}{n}\bar{x}_{\bar{s}}^2 + \bar{x}^{(2)}_{\bar{s}}\right] \quad (2.11)$$

Comparing variance expression of in (1.17) and upper bound of variance of $\hat{T}\left(0,1:x_{hk}^2\right)$ in (2.11), it is obvious that $\hat{T}\left(0,1:x_{hk}^2\right)$ will be more efficient than $\hat{T}_2$ unless the equality holds in (2.11). It may be noted that equality would hold true when two samples are exactly the same.

3. ROBUSTNESS OF $\hat{T}\left(0,1:x_{hk}^2\right)$

Now, if the model $\xi\left(0,1:x_{hk}^2\right)$ is true, then $\hat{T}\left(0,1:x_{hk}^2\right)$ is BLUE of T with variance given in expression (2.4). If the model of general polynomial form of degree J at most, i.e., $\xi\left(\delta_0,\delta_1,....,\delta_J:x_{hk}^2\right)$ is true and the estimator $\hat{T}\left(0,1:x_{hk}^2\right)$ is used, then how this estimator performs in terms of precision is a matter of investigation. We first derive the bias of $\hat{T}\left(0,1:x_{hk}^2\right)$ under the model $\xi\left(\delta_0,\delta_1,....,\delta_J:x_{hk}^2\right)$.

$$E_{\xi}\left[\hat{T}\left(0,1:x_{hk}^{2}\right)-T\right]=E_{\xi}\left[\sum_{h=1}^{H}\sum_{s_h}y_{hk}+\frac{1}{n}\sum_{h=1}^{H}\sum_{s_h}\left(y_{hk}/x_{hk}\right)\sum_{h=1}^{H}\sum_{\bar{s}_h}x_{hk}-\sum_{h=1}^{H}\sum_{k=1}^{N_h}y_{hk}\right]$$

$$=\sum_{j=0}^{J}\delta_j\beta_j\sum_{h=1}^{H}\sum_{\bar{s}_h}x_{hk}\left[\frac{\sum_{h=1}^{H}\sum_{s_h}x_{hk}^{j-1}}{n}-\frac{\sum_{h=1}^{H}\sum_{\bar{s}_h}x_{hk}^{j}}{\sum_{h=1}^{H}\sum_{\bar{s}_h}x_{hk}}\right] \quad (3.1)$$

For overall stratified overbalanced samples (see equation 2.6), it can easily be verified that the bias in equation (3.1) becomes zero and hence, the estimator $\hat{T}\left(0,1:x_{hk}^{2}\right)$ is unbiased even under the polynomial model $\xi\left(\delta_0,\delta_1,...,\delta_J:x_{hk}^{2}\right)$ but it may not be robust and optimal (BLUE) in general. However, some special cases are demonstrated below where the estimators are robust and optimal (BLUE).

Special cases:

Case I: Let the true model be $\xi\left(1,1:x_{hk}^{2}\right)$. The expression for bias in (3.1) for j = 0, 1 reduces to

$$E_{\xi}\left[\hat{T}\left(0,1:x_{hk}^{2}\right)-T\right]=\beta_0(N-n)\bar{x}_{\bar{s}}\left[\bar{x}_{s}^{(-1)}-\frac{1}{\bar{x}_{\bar{s}}}\right] \quad (3.2)$$

Where, $\bar{x}_{\bar{s}}=\frac{\sum_{h=1}^{H}\left(N_h-n_h\right)\bar{x}_{\bar{s}_h}}{N-n}$ and $\bar{\mathrm{x}}_{\mathrm{s}}^{(-1)}=\frac{\sum_{\mathrm{h}=1}^{\mathrm{H}}\mathrm{n_h}\bar{\mathrm{x}}_{\mathrm{s_h}}^{(-1)}}{\mathrm{n}}$

For j = 1, the condition (2.6) is clearly satisfied and this term contribute nothing to the bias. For j = 0, we have the term as shown in (3.2). If the samples $\mathrm{s_h}$ (h=1, 2 …, H) are so selected that the strategy $\bar{x}_{\bar{s}}\bar{x}_{s}^{(-1)}$ = 1, as per condition of (2.6), then this term will also become zero. Thus, under such situations $\hat{T}\left(0,1:x_{hk}^{2}\right)$ would be robust and optimal (BLUE) even under the model $\xi\left(1,1:x_{hk}^{2}\right)$ as variance of $\hat{T}\left(0,1:x_{hk}^{2}\right)$ in the model $\xi\left(1,1:x_{hk}^{2}\right)$ under condition (2.6) will be exactly similar to the variance of $\hat{T}\left(0,1:x_{hk}^{2}\right)$ in model $\xi\left(0,1:x_{hk}^{2}\right)$.

Case II: Let the true model be $\xi\left(1,1,1:x_{hk}^{2}\right)$. The expression for bias in (3.1) reduces to

$$E_{\xi}[\hat{T}(0,1:x_{hk}^{2})-T]=\beta_0(N-n)\bar{x}_{\bar{s}}\left[\bar{x}_{s}^{(-1)}-\frac{1}{\bar{x}_{\bar{s}}}\right]+\beta_2(N-n)\bar{x}_{\bar{s}}\left[\bar{x}_{s}-\bar{x}_{\bar{s}}^{(2)}/\bar{x}_{\bar{s}}\right] \quad (3.3)$$

Where, $\bar{x}_{\bar{s}}^{(2)}=\frac{\sum_{h=1}^{H}\left(N_h-n_h\right)\bar{x}_{\bar{s}_h}^{(2)}}{N-n}$.

Obviously, if the samples $\mathrm{s_h}$ are so selected that satisfy $\bar{x}_{\bar{s}}\bar{x}_{s}^{(-1)}=1$ and $\bar{x}_{s}\bar{x}_{\bar{s}}=\bar{x}_{\bar{s}}^{(2)}$ as per condition of (2.6), for j = 0, 1, 2, then bias in (3.3) becomes zero and hence, $\hat{T}\left(0,1:x_{hk}^{2}\right)$ is robust and optimal (BLUE) even under the model $\xi\left(1,1,1:x_{hk}^{2}\right)$ with variance given in (2.8).

Remarks:

(i) It can be verified that $\hat{T}\left(0,1:x_{hk}^{2}\right)$ will remain unbiased for stratified overbalanced samples satisfying the condition (2.6) even under the general polynomial model $\xi\left[\delta_0,\delta_1,...,\delta_J:v\left(x_{hk}\right)\right]$ for any variance function v($\mathrm{x_{hk}}$) but in general it will not be robust and optimal(BLUE).

(ii) For unstratified population with overbalanced sample $\mathrm{s}=\mathrm{s}'(\mathrm{J})$, Scott *et al.* (1978) have proved that $\hat{T}_0$ [0,1:v(x)] is BLU estimator under the model $\xi\left[\delta_0,\delta_1,...,\delta_J:v^{*}\left(x_{hk}\right)\right]$ for any variance function of the form $v^{*}(x)=v(x)\sum_{j=0}^{J}\delta_j a_j x^{j-1}$, where $\mathrm{a_j}$'s are some constant. Analogous result for $\hat{T}\left(0,1:x_{hk}^{2}\right)$ can easily be achieved for any variance function of the form $v^{*}(x_{hk})=x_{hk}^{2}\sum_{h=1}^{H}\sum_{j=0}^{J}\delta_j a_{hj}x_{hk}^{j-1}$ in stratified population with stratified overbalanced sample $\mathrm{s}=\mathrm{s}''(\mathrm{J})$, where $\mathrm{a_{hj}}$'s are some constants. In case of departure from $v^{*}(x_{hk})$, $\hat{T}\left(0,1:x_{hk}^{2}\right)$ remains ξ –unbiased but not optimal. It is, however, a matter of an investigation to find the extent of departure from the optimality if the variance function is different from $v^{*}(x_{hk})$.

ACKNOWLEDGEMENTS

The authors are highly thankful to the reviewer of the paper for his valuable suggestions which have improved the paper.

REFERENCES

Bayless, D.L. and Rao, J.N.K. (1970). An empirical study of stabilities of estimators and variance estimators in unequal probabilities sampling (n = 3 or 4). *Journal of the American Statistical Association,* **65**, 1645-1667.

Bouza, C.N. (1994). Robustness of shrunken predictors in stratified populations. *Biometrical Journal*, **36**, 95-102.

Cassel, C.M., Sarndal, C.E. and Wretman, J.H. (1976). Some results on generalised difference estimation and generalised regression estimation for finite population. *Biometrika,* **63**(3), 615-620.

Cassel, C.M., Sarndal, C.E. and Wretman, J.H. (1977). Foundations of inference in survey sampling. *New York, Wiley.*

Cassel, C.M., Sarndal, C.E. and Wretman, J.H.(1979). Prediction theory for finite population when model –based and design-based principles are combined. *Scandinavian journal of statistics*, **6**, 97-106.

Dorfman, A.H. and Valliant, R. (2000). Stratification by size revised, *Journal of official Statistics*, **16**(2), 139-154.

Jessen, R. (1942). Statistical investigation of a sample survey for obtaining farm facts. *Iowa Agricultural Experiment Station Research Bulletin,* 104.

Kaushal, R.P., Sisodia, B.V.S. and Sud, U.C.(2011).Shrunken predictors in stratified sampling under super population model. *Statistics*, **45**(3), 281-291

Kaushal, R.P. and Sisodia, B.V.S. (2021). Robust estimation in stratified sampling under super-population model. *Communication in Statistics-Theory & Methods* (Submitted)

Raj, Des (1958). On the relative accuracy of some sampling techniques. *Journal of the American Statistical Association,* **53**, 98-101.

Rao, J.N.K. and Bayless, D.L. (1969). An empirical study of stabilities of estimators and variance estimators in unequal probability sampling of two units per stratum. *Journal of the American Statistical Association,* **64**, 540-559.

Royall, R.M. (1970). On finite population sampling theory under certain linear regression models. *Biometrika*, **57**(2),377-387.

Royall, R.M. (1971). Linear regression models in finite population sampling theory, in Godambe, and V.P, Sprott, D.A. Eds. Foundations of Statistical Inference, Toronto Holt, Rinehart and Winston of Canada Ltd.

Royall, R.M. and Herson, J. (1973a). Robust estimation in finite populations. I. *Journal of the American Statistical Association,* **68**, 880-889.

Royall, R.M. and Herson, J. (1973b). Robust estimation in finite populations II. Stratification on a size variable. *Journal of the American Statistical Association,* **68**, 890-893.

Royall, R.M. (1976). The linear least squares prediction approach to two –stage sampling. *Journal of the American Statistical Association*, **71**, 657-664.

Royall, R.M. (1992). Robustness and optimal design under prediction models for finite populations. *Survey Methodology*, **2**, 179-195.

Scott, A.J., Brewer, K.R.W. and Ho, E.W.H. (1978). Finite population sampling and robust estimation. *Journal of the American Statistical Association,* **73**, 359-361.

Sarndal, C.E., Swensson, B. and Wretman, J. (1992). Model assisted survey sampling. *Springer-Verlag, New-York Inc*.

Sisodia, B.V.S., Kaushal, R.P. and Chandra, Hukum (2015). Shrunken predictors in stratified sampling*. International Journal of Agriculture and Statistical Sciences*, **11**(01), 123-129.

Smith, H. Fairfield (1938). An experimental law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Sciences*, **28**, 1-23.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). Sampling theory of surveys with applications, Third Edition, *Iowa State University Press, USA, and Indian Society of Agricultural Statistics*, New Delhi, India.

Tam, S.M. (1986). Characterisations of best model-based predictors in survey sampling. *Biometrika*, **73**, 232-235.

Tam, S.M. (1987). Optimality of Royall's predictor under a Gaussian super population model. *Biometrika*, **74**, 659-660.

Valliant, R., Dorfman, A.H. and Royall, R.M. (2000). Finite population sampling and inferences. Wiley, New York.