International Journal of Computational Intelligence Research
ISSN 0973-1873 Volume 13, Number 1 (2017), pp. 35-44

http://www.ripublication.com

An Improved Class of Unbiased Separate Regression Type Estimator under Stratified Random Sampling

Shashi Bhushan[1] and Praveen Kumar Misra[1]

[1] *Department of Mathematics and Statistics, Dr. Shakuntala Misra National Rehabilitation University, India*

Abstract

In this paper a class of regression type estimators using the auxiliary information on population mean and population variance is proposed under stratified random sampling. In order to improve the performance of the proposed class of estimator, the Jack-knifed versions are also proposed. A comparative study of the proposed estimator is made with that of separate ratio estimator, separate product estimator, separate linear regression estimator and the usual stratified sample mean. It is shown that the estimators through proposed allocation always give more efficient estimators in the sense of having smaller mean square error than those obtained through Neyman Allocation.

Keywords: Auxiliary information,ratio type Estimator, Bias, Mean Square Error.

INTRODUCTION OF THE PROPOSED ESTIMATOR

Let a population of size 'N' be stratified in to 'L' non-overlapping strata, the h^{th} stratum size being N_h (h=1,2,……,L) and $\sum_{h=1}^{L} N_h = N$. Suppose 'y' be characteristic under study and 'x' be the auxiliary variable. We denote by

y_{hj}: The observation on the j^{th} unit of the population for the charectarstic 'y' under study $(j = 1, 2, ..., N_h)$ in the h^{th} stratum (h =1,2,…..,L).

x_{hj}: The observation on the $\hat{\hat{Y}}_{gS}$ unit of the population for the auxiliary charectarstic 'x' under study $(j = 1, 2, ..., N_h)$ in the h^{th} stratum (h =1,2,…..,L).

$$\bar{Y}_h = \frac{1}{N_h}\sum_{j=1}^{N_h} y_{hj}, \bar{X}_h = \frac{1}{N_h}\sum_{j=1}^{N_h} x_{hj}$$

$$S_{yh}^2 = \frac{1}{(N_h - 1)}\sum_{h=1}^{N_h}(y_{hj} - \bar{Y}_h)^2, S_{xh}^2 = \frac{1}{(N_h - 1)}\sum_{j=1}^{N_h}(x_{hj} - \bar{X}_h)^2$$

where $\bar{Y}_h$ and $\bar{X}_h$ are population means of character 'y' and 'x' respectively in the h^{th} stratum.

$$\sigma_{xh}^2 = \frac{1}{N_h}\sum_{j=1}^{N_h}(x_{hj} - \bar{X}_h)^2, \sigma_{yh}^2 = \frac{1}{N_h}\sum_{j=1}^{N_h}(y_{hj} - \bar{Y}_h)^2$$

$$S_{xyh} = \frac{1}{(N_h - 1)}\sum_{j=1}^{N_h}(x_{hj} - \bar{X}_h)\ (y_{hj} - \bar{Y}_h) = \rho_h S_{xh} S_{hj}$$

where ρ_h is the population correlation coefficient between 'x' and 'y' for the h^{th} stratum $(j = 1, 2, ..., N_h)$.

$$R_h = \frac{\bar{Y}_h}{\bar{X}_h}, C_{yh}^2 = \frac{S_{yh}^2}{\bar{Y}_h^2} = \frac{\mu_{02h}}{\bar{Y}_h^2}, C_{xh}^2 = \frac{S_{xh}^2}{\bar{X}_h^2} = \frac{\mu_{20h}}{\bar{X}_h^2}$$

$\mu_{pqh} = \frac{1}{N_h}\sum_{j=1}^{L}(X_{hj} - \bar{X}_h)^p (Y_{hj} - \bar{Y}_h)^q$: the $(p, q)^{th}$ population product moment about mean between 'x' and 'y' for the h^{th} stratum $(h = 1, 2, ..., L)$.

$\beta_{1h} = \frac{\mu_{30h}^2}{\mu_{20h}^2}, \beta_{2h} = \frac{\mu_{40h}^2}{\mu_{20h}^2}, \beta_h = \frac{S_{xyh}}{S_{xh}^2} = \rho_h \frac{S_{yh}}{S_{xh}}$ be the population regression coefficient of y on x for the h^{th} stratum $(h = 1, 2, ..., L)$.

Let a simple random sample of size n_h be selected from the h^{th} stratum without replacement, without any loss of generality, we assume that first n_h units have been selected in the h^{th} stratum from N_h units by SRSWOR.

Moreover we assume that N_h is so large that $1 - f_h \simeq 1$.

We define

$$\bar{y}_h = \frac{1}{n_h}\sum_{j=1}^{n_h} y_{hj}, \bar{x}_h = \frac{1}{n_h}\sum_{j=1}^{n_h} x_{hj}$$

$$s_{xh}^2 = \frac{1}{n_h - 1}\sum_{j=1}^{n_h}(x_{hj} - \bar{x}_h)^2, s_{yh}^2 = \frac{1}{n_h - 1}\sum_{j=1}^{n_h}(y_{hj} - \bar{y}_h)^2$$

$$\hat{\sigma}_{xh}^2 = \frac{1}{n_h}\sum_{j=1}^{n_h}(x_{hj} - \bar{x}_h)^2, \hat{\sigma}_{yh}^2 = \frac{1}{n_h}\sum_{j=1}^{n_h}(y_{hj} - \bar{y}_h)^2$$

$$s_{xyh} = \frac{1}{n_h - 1}\sum_{j=1}^{n_h}(x_{hj} - \bar{x}_h)(y_{hj} - \bar{y}_h), b_h = \frac{s_{xyh}}{s_{xh}^2}$$

Assuming that $\bar{X}_h$ is known $\forall h = 1, 2, ..., L$.The proposed generalized estimator $\hat{\bar{Y}}_{\theta S}$ for estimating the population mean $\bar{Y}$ of the study variable is given by

$$\hat{\bar{Y}}_{\theta S} = \sum_{j=1}^{L} W_h \left\{ \bar{y}_h \left[1 + \frac{\theta_h \left(\hat{\sigma}_{xh}^2 - \sigma_{xh}^2 \right)}{\sigma_{xh}^2} \right] + b_h \left(\bar{X}_h - \bar{x}_h \right) \right\}$$

$$\hat{\bar{Y}}_{\theta S} = \sum_{j=1}^{L} W_h \left\{ \bar{y}_h + \theta_h \bar{y}_h \left(\frac{\hat{\sigma}_{xh}^2}{\sigma_{xh}^2} - 1 \right) + b_h \left(\bar{X}_h - \bar{x}_h \right) \right\} \quad (1.1)$$

Where θ_h are the characterizing scalars to be chosen suitably, population strata means $\bar{X}_h$ and population strata variances σ_{xh}^2 of the auxiliary variable x are assumed to be known. It should be noted that for $\theta = 0$ the proposed generalized estimator reduces to the separate linear regression estimator given by

$$\bar{y}_{LRS} = \sum_{j=1}^{L} W_h \left\{ \bar{y}_h + b_h \left(\bar{X}_h - \bar{x}_h \right) \right\} \quad (1.2)$$

BIAS AND MEAN SQUARE ERROR OF THE PROPOSED ESTIMATOR $\hat{\bar{Y}}_{\theta S}$

Let

$$\bar{y}_h - \bar{Y} = e_{0h}, \bar{x}_h - \bar{X} = e_{1h}, s_{xyh} - S_{xyh} = e_{2h}, s_{xh}^2 - S_{xh}^2 = e_{3h}, \hat{\sigma}_{xh}^2 - \sigma_{xh}^2 = e_{4h}$$

$$E(e_{0h}) = E(e_{1h}) = E(e_{2h}) = E(e_{3h}) = E(e_{4h}) = 0, \forall h = 1, 2, ...L$$

Now from (1.1.1) we have

$$\hat{\bar{Y}}_{\theta S}=\sum_{j=1}^{L}W_h\left\{(\bar{Y}_h+e_{0h})+\theta_h(\bar{Y}_h+e_{0h})\left[\frac{e_{4h}}{\sigma_{xh}^2}\right]+\beta_h\left(1+\frac{e_{2h}}{S_{xyh}}\right)\left(1+\frac{e_{3h}}{S_{xh}^2}\right)^{-1}\left(-e_{1h}\right)\right\} \qquad (2.1)$$

$$\hat{\bar{Y}}_{\theta S}=\sum_{j=1}^{L}W_h\left\{\bar{Y}_h+e_{0h}+\theta_h\bar{Y}_h\left[\frac{e_{4h}}{\sigma_{xh}^2}+\frac{e_{0h}e_{4h}}{\bar{Y}_h\sigma_{xh}^2}\right]+\beta_h\left(1+\frac{e_{2h}}{S_{xyh}}\right)\left(1-\frac{e_{3h}}{S_{xh}^2}+......\right)\left(-e_{1h}\right)\right\}$$

$$\hat{\bar{Y}}_{\theta S}=\sum_{j=1}^{L}W_h\left\{\bar{Y}_h+e_{0h}+\theta_h\bar{Y}_h\left[\frac{e_{4h}}{\sigma_{xh}^2}+\frac{e_{0h}e_{4h}}{\bar{Y}_h\sigma_{xh}^2}\right]+\beta_h\left[-e_{1h}-\frac{e_{1h}e_{2h}}{S_{xyh}}+\frac{e_{1h}e_{3h}}{S_{xh}^2}+............\right]\right\}$$

Let the sample size be so large that $|e_i|$, i=0,1,2,3,4 ; $\forall h=1,2,......,L$;becomes so small that terms of e_i's having powers greater than two may be neglected. Under this assumption, we get

$$E(\hat{\bar{Y}}_{\theta S})=\sum_{j=1}^{L}W_h\left\{\bar{Y}_h+\theta_h\frac{E(e_{0h}e_{4h})}{\sigma_{xh}^2}+\beta\left[\frac{E(e_{1h}e_{3h})}{S_{xh}^2}-\frac{E(e_{1h}e_{2h})}{S_{xyh}}\right]\right\}$$

Using the results given in Sukhatme and Sukhatme (1997) and proved in appendix

$$E(e_{1h}e_{3h})=\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\mu_{30h},\ E(e_{1h}e_{2h})=\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\mu_{21h}$$

$$E(e_{0h}e_{4h})=\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\mu_{21h};\forall h=1,2,......,L$$

we have

$$E(\hat{\bar{Y}}_{\theta S})=\sum_{h=1}^{L}W_h\left[\bar{Y}_h+\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left\{\frac{\theta_h\mu_{21h}}{\sigma_{xh}^2}+\beta_h\left(\frac{\mu_{30h}}{S_{xh}^2}-\frac{\mu_{21h}}{S_{xyh}}\right)\right\}\right] \qquad (2.2)$$

Showing that $\hat{\bar{Y}}_{\theta S}$ is a biased estimator of population mean $\bar{Y}$ and its bias is given by:

$$B(\hat{\bar{Y}}_{\theta S})=E(\hat{\bar{Y}}_{\theta S})-\bar{Y}_h=\sum_{h=1}^{L}W_h\left[\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left\{\frac{\theta_h\mu_{21h}}{\sigma_{xh}^2}+\beta_h\left(\frac{\mu_{30h}}{S_{xh}^2}-\frac{\mu_{21h}}{S_{xyh}}\right)\right\}\right] \qquad (2.3)$$

The mean square error of $\hat{\bar{Y}}_{\theta S}$ is given by:

$$MSE(\hat{\bar{Y}}_{\theta S}) = E(\hat{\bar{Y}}_{\theta S} - \bar{Y})^2 = E\left\{\sum_{h=1}^{L} W_h \left[e_{0h} + \theta_h \bar{Y}_h \left(\frac{e_{4h}}{\sigma_{xh}^2} + \frac{e_{0h}e_{4h}}{\bar{Y}_h \sigma_{xh}^2} \right) + \beta_h \left(-e_{1h} - \frac{e_{1h}e_{2h}}{S_{xyh}} + \frac{e_{1h}e_{3h}}{S_{xh}^2} + ... \right) \right] \right\}$$

$$MSE(\hat{\bar{Y}}_{\theta S}) = E\left\{\sum_{j=1}^{L} W_h \left(e_{0h} + \theta_h \bar{Y}_h \frac{e_{4h}}{\sigma_{xh}^2} - \beta_h e_{1h} \right)^2 \right\}$$

Using (2.2) to the first order of approximation, we have

$$MSE(\hat{\bar{Y}}_{\theta S}) = \sum_{j=1}^{L} W_h \left\{ \begin{array}{l} E\left(e_{0h}^2\right) + E\left(\frac{\theta_h{}^2 \bar{Y}_h{}^2}{\sigma_{xh}^2} e_{4h}^2 \right) + E(\beta_h{}^2 e_{1h}^2) + 2\frac{\theta_h \bar{Y}_h}{\sigma_{xh}^2} E(e_{0h}e_{4h}) - 2\beta_h E(e_{0h}e_{1h}) \\ -2\frac{\theta_h \bar{Y}_h \beta_h}{\sigma_{xh}^2} E(e_{1h}e_{4h}) \end{array} \right\}$$

Substituting the following results given in Sukhatme and Sukhatme (1997) and proved in appendix

$$E(e_{0h}^2) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) S_{yh}^2, E(e_{1h}^2) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) S_{xh}^2, E(e_{0h}e_{1h}) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) S_{xyh}$$

$$E(e_{4h}^2) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left(\mu_{40h} - \mu_{20h}^2 \right), E(e_{0h}e_{4h}) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{21h} E(e_{1h}e_{4h}) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \mu_{30h}$$

$; \forall h = 1, 2, ..., L$

We get,

$$MSE\left(\hat{\bar{Y}}_{\theta S}\right) = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ \begin{array}{l} S_{yh}^2 + \frac{\theta_h{}^2 \bar{Y}_h{}^2}{\sigma_X^4} . \left(\mu_{40h} - \mu_{20h}^2 \right) + \beta_h{}^2 S_{xh}^2 \\ + \frac{2\theta_h \bar{Y}_h}{\sigma_{xh}^2} \mu_{21h} - 2\beta_h S_{xyh} - \frac{2\theta_h \bar{Y}_h \beta_h}{\sigma_{xh}^2} \mu_{30h} \end{array} \right\} \quad (2.4)$$

(1.2.4) is minimum when $\theta_h \bar{Y}_h = \frac{\left(\beta_h \mu_{20h} - \mu_{21h}\right)}{\left(\mu_{40h} - \mu_{20h}^2\right)} \mu_{20h}; \forall h = 1, 2, ..., L$ (2.5)

and the minimum mean square error of $\hat{\bar{Y}}_{\theta S}$ is given by

$$MSE\left(\hat{\bar{Y}}_{\theta S}\right)_{opt} = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left\{\left(1-\rho_h{}^2\right) S_{yh}^2 - \frac{\left(\beta_h\mu_{30h} - \mu_{21h}\right)^2}{\left(\mu_{40h} - \mu_{20h}^2\right)}\right\}$$

$$MSE\left(\hat{\bar{Y}}_{\theta S}\right)_{opt} = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left\{\left(1-\rho_h{}^2\right) S_{yh}^2 - \frac{\left(\beta_h\mu_{30h} - \mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h} - 1\right)}\right\} \quad (2.6)$$

THE PROPOSED JACK-KNIFE ESTIMATOR $\hat{\bar{Y}}_j^*$

Let a simple random sample of size m=2n is drawn without replacement from the population of size N. This sample of size m=2n is then split up at random into two sub-samples each of size n. Let us define:

$$\hat{\bar{Y}}_{\theta S}^{(1)} = \sum_{j=1}^{L} W_h \left\{\bar{y}_h + \theta_h \bar{y}_h \left(\frac{\hat{\sigma}_{xh}^2}{\sigma_{xh}^2} - 1\right) + b_h\left(\bar{X}_h - \bar{x}_h\right)\right\}$$

$$\hat{\bar{Y}}_{\theta S}^{(2)} = \sum_{j=1}^{L} W_h \left\{\bar{y}_h + \theta_h \bar{y}_h \left(\frac{\hat{\sigma}_{xh}^2}{\sigma_{xh}^2} - 1\right) + b_h\left(\bar{X}_h - \bar{x}_h\right)\right\}$$

$$\hat{\bar{Y}}_{\theta S}^{(3)} = \sum_{j=1}^{L} W_h \left\{\bar{y}_h + \theta_h \bar{y}_h \left(\frac{\hat{\sigma}_{xh}^2}{\sigma_{xh}^2} - 1\right) + b_h\left(\bar{X}_h - \bar{x}_h\right)\right\} \text{Where } b^{(1)} = \frac{s_{xyh}^{(1)}}{s_{xh}^{2(1)}}, b^{(2)} = \frac{s_{xyh}^{(2)}}{s_{xh}^{2(2)}}, b^{(3)} = \frac{s_{xyh}^{(3)}}{s_{xh}^{2(2)}}$$

where $\bar{y}_n^{(1)}, \bar{y}_n^{(2)}, \bar{y}_{2n}$, be the respective sample means based on two sub-samples of size n and the entire sample of size 2n for the characteristic y under study; $\bar{x}_n^{(1)}, \bar{x}_n^{(2)}, \bar{x}_{2n}$ and $s_{xh}^{2(1)}, s_{xh}^{2(2)}, s_{xh}^2$ be the corresponding mean and sample variances for the auxiliary variable x. Also $\hat{\sigma}_{xh}^{2(1)}, \hat{\sigma}_{xh}^{2(2)}, \hat{\sigma}_{xh}^2$, are given by:

$$\hat{\sigma}_{xh}^{2(1)} = \frac{1}{n}\sum_{i=1}^{n}(x_i^{(1)} - \bar{X})^2, \hat{\sigma}_{xh}^{2(2)} = \frac{1}{n}\sum_{i=1}^{n}(x_i^{(2)} - \bar{X})^2, \hat{\sigma}_{xh}^2 = \frac{1}{2n}\sum_{i=1}^{n}(x_i - \bar{X})^2$$

It can be easily seen that

$$B\left(\hat{\bar{Y}}_{\theta h}^{(1)}\right) = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left\{\frac{\theta_h\mu_{21h}}{\sigma_{xh}^2} + \beta_h\left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}}\right)\right\}, B\left(\hat{\bar{Y}}_{\theta h}^{(2)}\right) = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left\{\frac{\theta_h\mu_{21h}}{\sigma_{xh}^2} + \beta_h\left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}}\right)\right\}$$

$$B\left(\hat{\bar{Y}}_{\theta h}^{(3)}\right) = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left\{\frac{\theta_h\mu_{21h}}{\sigma_{xh}^2} + \beta_h\left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}}\right)\right\} = B_1 \quad (3.1)$$

Let us define:

$$Bias(\hat{\hat{Y}}_{gh}^{(1)}) = \left\{ \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left[\frac{g'(1)}{\sigma_{xh}^2} \mu_{21h} + \frac{\bar{Y} g''(1)}{2\sigma_{xh}^4} \left(\mu_{40h} - \mu_{20h}^2 \right) + \beta_h \left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right] \right\}$$

$$Bias(\hat{\hat{Y}}_{gh}^{(1)}) = \left\{ \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left[\frac{g'(1)}{\sigma_{xh}^2} \mu_{21h} + \frac{\bar{Y} g''(1)}{2\sigma_{xh}^4} \left(\mu_{40h} - \mu_{20h}^2 \right) + \beta_h \left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right] \right\} \tag{2.2.3}$$

$$Bias(\hat{\hat{Y}}_{gh}^{(3)}) = \left\{ \left(\frac{1}{2n_h} - \frac{1}{N_h} \right) \left[\frac{g'(1)}{\sigma_{xh}^2} \mu_{21h} + \frac{\bar{Y} g''(1)}{2\sigma_{xh}^4} \left(\mu_{40h} - \mu_{20h}^2 \right) + \beta_h \left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right] \right\}$$

Be an alternative estimator of the population mean ($\bar{Y}$). The bias of $\hat{\hat{Y}}_\theta'$ is given by:

$$B\left(\hat{\hat{Y}}_{\theta h}'\right) = \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ \frac{\theta_h \mu_{21_h}}{\sigma_{x_h}^2} + \beta \left(\frac{\mu_{30h}}{S_{xh}^2} - \frac{\mu_{21h}}{S_{xyh}} \right) \right\} \quad = \left(B_2 \right), \text{say}$$

With $R = \left(\frac{B_1}{B_2} \right)$ the following jack-knifed estimator is proposed for estimating the population mean $\bar{Y}$ as:

$$\hat{\hat{Y}}_j^* = \frac{\hat{\hat{Y}}^{(3)} - R\left(\hat{\hat{Y}}_{\theta h}'\right)}{(1-R)} = \frac{\hat{\hat{Y}}^{(3)} - \left\{ \frac{N-2n}{2(N-n)} \right\} \hat{\hat{Y}}_{\theta h}'}{(1-R)} \tag{3.2}$$

Taking expectation of (2.2.4) we have

$$E(\hat{\hat{Y}}^*) = \frac{E(\hat{\hat{Y}}^{(3)}) - RE\left(\hat{\hat{Y}}_\theta'\right)}{(1-R)}$$

$$E(\hat{\hat{Y}}_h^*) = \bar{Y} \tag{3.3}$$

showing that $\left(\hat{\hat{Y}}_j^*\right)$ is an unbiased estimate of population mean $\bar{Y}$ to the first order of approximation.

MEAN SQUARE ERROR OF $\hat{\hat{Y}}_j^*$

$$MSE\left(\left(\hat{\hat{Y}}_j^*\right)\right) = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h} \right) \left\{ \begin{array}{l} S_{yh}^2 + \frac{\theta_h^{\,2} \bar{Y}_h^{\,2}}{\sigma_X^4} . \left(\mu_{40h} - \mu_{20h}^2 \right) + \beta_h^{\,2} S_{xh}^2 \\ + \frac{2\theta_h \bar{Y}_h}{\sigma_{xh}^2} \mu_{21h} - 2\beta_h S_{xyh} - \frac{2\theta_h \bar{Y}_h \beta_h}{\sigma_{xh}^2} \mu_{30h} \end{array} \right\} \tag{.4.1}$$

(1.2) is minimum when $\theta_h \overline{Y}_h = \frac{\left(\beta_h \mu_{20h} - \mu_{21h}\right)}{\left(\mu_{40h} - \mu_{20h}^2\right)} \mu_{20h}; \forall h = 1, 2, ..., L$ (4.2)

and the minimum mean square error of $\hat{\bar{Y}}_{\theta S}$ is given by

$$MSE\left(\left(\hat{\bar{Y}}_j^*\right)\right)_{opt} = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\left(\mu_{40h} - \mu_{20h}^2\right)} \right\}$$

$$MSE\left(\left(\hat{\bar{Y}}_j^*\right)\right)_{opt} = \sum_{j=1}^{L} W_h \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\mu_{20h}^2 \left(\beta_{2h} - 1\right)} \right\} \quad (4.3)$$

OPTIMUM ALLOCATION WITH THE PROPOSED CLASS

Consider the cost function $C = C_0 + \sum_{h=1}^{L} c_h n_h$, where C_0 is fixed cost and c_h be the cost of drawing per unit sample within h^{th} stratum respectively, we have

$$V\left(\hat{\bar{Y}}_{\theta S}\right)_{\min} = \sum_{j=1}^{L} \frac{W_h^2}{n_h} \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\left(\mu_{40h} - \mu_{20h}^2\right)} \right\} \quad (5.1)$$

we wish to minimize this variance for fixed cost to find out n_h. For this we apply Lagrange's method of multipliers. Accordingly, we define:

$$\phi = V\left(\hat{\bar{Y}}_{\theta S}\right)_{\min} + \lambda \left(\sum_{h=1}^{L} c_h n_h - C + C_0 \right)$$

where λ is constant known as Lagrange's multiplier.

$$\text{Or } \phi = \sum_{j=1}^{L} \frac{W_h^2}{n_h} \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\left(\mu_{40h} - \mu_{20h}^2\right)} \right\} + \lambda \left(\sum_{h=1}^{L} c_h n_h - C + C_0 \right) \quad (5.2)$$

Differentiating (4.3.2) with respect to n_h and equating to zero, we get

$$-\frac{W_h^2}{n_h^2} \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\mu_{20h}^2 \left(\beta_{2h} - 1\right)} \right\} + \lambda c_h = 0$$

$$n_h = \frac{1}{\sqrt{\lambda}} \frac{W_h}{\sqrt{c_h}} \left\{ \left(1 - \rho_h^{\,2}\right) S_{yh}^2 - \frac{\left(\beta_h \mu_{30h} - \mu_{21h}\right)^2}{\mu_{20h}^2 \left(\beta_{2h} - 1\right)} \right\}^{\frac{1}{2}}; \forall h = 1, 2, ..., L \quad (5.3)$$

Summing over all strata we have

$$n=\frac{1}{\sqrt{\lambda}}\sum_{j=1}^{L}\frac{W_h}{\sqrt{c_h}}\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}} \tag{5.4}$$

Taking ratio of (1.3.3) and (1.3.4) we obtain

$$n_h=n\frac{\frac{1}{\sqrt{\lambda}}\frac{W_h}{\sqrt{c_h}}\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}}}{\frac{1}{\sqrt{\lambda}}\sum_{j=1}^{L}\frac{W_h}{\sqrt{c_h}}\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}}}\ \forall h=1,2,.....,L \tag{5.5}$$

Assuming cost of drawing per unit sample in each stratum is same. The optimum allocation (4.3) reduces to

$$n_h=n\frac{W_h\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}}}{\sum_{j=1}^{L}W_h\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}}}\ \forall h=1,2,.....,L \tag{5.6}$$

Substituting the value from (1.3.6) in (1.3.1) we have

$$V\left(\hat{\hat{Y}}_{\theta S}\right)_{\min.opt}=\frac{1}{n}\sum_{h=1}^{L}\left[W_h\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\}^{\frac{1}{2}}\right]^2=V_{opt}(say) \tag{5.7}$$

CONCLUDING REMARKS

The mean square error of the separate linear regression estimator is given by

$$MSE(\bar{y}_{LRS})=\sum_{h=1}^{L}W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left(1-\rho_h^2\right)S_{yh}^2 \tag{6.1}$$

Also the minimum mean square error of the proposed generalized regression type estimator $\hat{\hat{Y}}_{\theta S}$ is given by

$$MSE\left(\hat{\hat{Y}}_{\theta S}\right)_{\min}=\sum_{h=1}^{L}W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left\{\left(1-\rho_h^{\,2}\right)S_{yh}^2-\frac{\left(\beta_h\mu_{30h}-\mu_{21h}\right)^2}{\mu_{20h}^2\left(\beta_{2h}-1\right)}\right\} \tag{6.2}$$

Therefore the proposed class of estimators $\hat{\hat{Y}}_{\theta S}$ and $\left(\hat{\hat{Y}}_j^*\right)$ may be preferred to the separate linear regression estimator, separate ratio estimator, separate product estimator and the usual stratified sample mean in the sense of smaller mean square error. Further the parameter involved θ_h may be estimated by the corresponding sample value in order to get a class of estimators depending upon estimated optimum value.

The variance of stratified sample mean $\bar{y}_{st}$ under Neyman allocation $n_h = n \frac{W_h S_{yh}}{\sum_{h=1}^{L} W_h S_{yh}}$

Is given by $V\left(\bar{y}_{st}\right)_{Ney} = \frac{1}{n}\left(\sum_{h=1}^{L} W_h S_{yh}\right)^2$ (ignoring f.p.c)

It is evident that V_{opt} is always smaller than $V\left(\bar{y}_{st}\right)_{Ney}$ except for the case when ρ_h =0 and $\beta_h \mu_{30h} = \mu_{21h}$ simultaneously.

REFERENCES

[1] Jack – knife technique of Gray and Schucany and Sukhatme et al.

[2] W.G.Cochran, Sampling Techniques, third ed. John Wiley &sons, New York, 1977.

[3] R.S.Jessen, Statistical Survey Techniques, John Wiley &sons, New York, 1978.

[4] D.Singh,F.S.Chaudhary, Theory and Analysis of sample Survey Design, New Age Publication, New Delhi, India 1986.

[5] S.K.Srivastava, An estimator using auxiliary information in sample surveys, Calcutta Statistical Association Bulletin 16(1967) 121-132.

[6] Bhushan. S. (2013)" Improved sampling strategies in finite population, Scholars Press, Germany.

[7] Bahli, S. And Tuteja, R.K. (1991): Ratio and Product type exponential estimator, Information and Optimization sciences, Vol.XII, I, 159-163.

[8] Gray H. L. and Schucany W.R. (1972). The Generalized Jack-knife Statistic, Marcel Dekker, New York.

[9] Naik, V.D. and Gupta, P.C. (1996): A note on estimation of mean with known population proportion of an auxiliary character. Jour. Ind. Soc. Agr. Stat., 48(2),151-158.

[10] Sukhatme, P.V. andSukhatme, B.V. (1970): Sampling theory of surveys with applications. Iowa State University Press, Ames, U.S.A.