DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 7 | Issue 2 Article 13

11-1-2008

An Optimum Allocation with a Family of Estimators Using Auxiliary Information in Sample Survey

Gajendra K. Vishwakarma
Vikram University, vishwagk@rediffmail.com

Housila P. Singh
Vikram University, Ujjain, India

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Recommended Citation

Vishwakarma, Gajendra K. and Singh, Housila P. (2008) "An Optimum Allocation with a Family of Estimators Using Auxiliary Information in Sample Survey," *Journal of Modern Applied Statistical Methods*: Vol. 7 : Iss. 2 , Article 13.
DOI: 10.22237/jmasm/1225512720
Available at: http://digitalcommons.wayne.edu/jmasm/vol7/iss2/13

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Journal of Modern Applied Statistical Methods
November, 2008, Vol. 7, No. 2, 478-487
Copyright © 2008 JMASM, Inc.
1538 – 9472/08/$95.00

An Optimum Allocation with a Family of Estimators Using Auxiliary Information in Sample Survey

Gajendra K. Vishwakarma Housila P. Singh
Vikram University, India

The problem of obtaining optimum allocation using auxiliary information in stratified random sampling. An optimum allocation with a family of estimators is obtained and its efficiency is compared with that of Neyman allocation based on Srivastava (1971) class of estimators and the optimum allocation suggested by Zaidi et al., (1989). It is shown that the proposed allocation is better in the sense having smaller variance compared to other optimum allocation.

Key words: Auxiliary variate, study variate, variance, optimum allocation, stratified random sampling.

Introduction

When a population contains heterogeneity among units in terms of value, survey users are advised to form several homogeneous groups, and the sampling design is known as stratified sampling. All designs, other than these, are generated as a further modification of simple random sampling and stratified sampling. Stratification is one of the most widely used techniques in sample survey design due to its dual purposes of providing samples that are representative of major sub-groups of the population and increasing the precision of estimators. It is also well established that the auxiliary information may lead to more efficient estimators: ratio, product and regression methods of estimation are examples in this context. This article suggests a class of estimators using auxiliary information in stratified random sampling and discusses its properties.

Let y be the study variate and x be the auxiliary variate, let the population $U=\left(U_1,U_2,U_3,\ldots,U_N\right)$ of size N be divided into L stratum, and let N_h and n_h be the total number of units and sample size respectively in h^{th} stratum, such that $\sum_{h=1}^{L} N_h = N$ and $\sum_{h=1}^{L} n_h = n$. Next, let $\left(y_{hj},\, x_{hj}\right)$ be the pair of values according to the variate under study y and the auxiliary variate x respectively for j^{th}-unit $\left(j=1,2,3,\ldots,N_h\right)$ in the h^{th} sample of size n_h selected by simple random sampling from the h^{th} stratum $\left(j=1,2,3,\ldots,N_h;\, h=1,2,3,\ldots,L\right)$. For simplicity, assume that N_h is large enough compared to n_h so that $f_h=\dfrac{n_h}{N_h}\approx 0$. Denote

$$\bar{Y}=\sum_{h=1}^{L} W_h\bar{Y}_h,\ \bar{X}=\sum_{h=1}^{L} W_h\bar{X}_h,$$

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

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

$$\bar{y}_{st}=\sum_{h=1}^{L} W_h\bar{y}_h,\ \bar{x}_{st}=\sum_{h=1}^{L} W_h\bar{x}_h,$$

Address correspondence to Gajendra K. Vishwakarma, School of Studies in Statistics, Vikram University, Ujjain - 456010, M.P., India. E-mail: vishwagk@rediffmail.com.

$$W_h = \frac{N_h}{N},\ S_{yh}^2 = \frac{1}{(N_h - 1)}\sum_{j=1}^{N_h}\left(y_{hj} - \overline{Y}_h\right)^2,$$

$$S_{xh}^2 = \frac{1}{(N_h - 1)}\sum_{j=1}^{N_h}\left(x_{hj} - \overline{X}_h\right)^2,$$

$$S_{xyh} = \frac{1}{(N_h - 1)}\sum_{j=1}^{N_h}\left(y_{hj} - \overline{Y}_h\right)\left(x_{hj} - \overline{X}_h\right)$$

$$s_{yh}^2 = \frac{1}{(n_h - 1)}\sum_{j=1}^{n_h}\left(y_{hj} - \overline{y}_h\right)^2,$$

$$s_{xh}^2 = \frac{1}{(n_h - 1)}\sum_{j=1}^{n_h}\left(x_{hj} - \overline{x}_h\right)^2,$$

$$s_{xyh} = \frac{1}{(n_h - 1)}\sum_{j=1}^{n_h}\left(x_{hj} - \overline{x}_h\right)\left(y_{hj} - \overline{y}_h\right)$$

$$\mu_{rsh} = \frac{1}{N_h}\sum_{j=1}^{N_h}\left(y_{hj} - \overline{Y}_h\right)^r\left(x_{hj} - \overline{X}_h\right)^s,$$

$$\rho_h = \frac{S_{xyh}}{S_{yh}\,S_{xh}},\ r_h = \frac{s_{xyh}}{s_{yh}\,s_{xh}}$$

$$C_{yh}^2 = \frac{S_{yh}^2}{\overline{Y}_h^2},\ C_{xh}^2 = \frac{S_{xh}^2}{\overline{X}_h^2},\ \lambda_{rsh} = \frac{\mu_{rsh}}{\left(\mu_{20h}^r\,\mu_{02h}^s\right)^{1/2}}$$

$$a_h = \frac{\overline{x}_h}{\overline{X}_h}, b_h = \frac{s_{xh}^2}{S_{xh}^2} \text{ and } c_h = \frac{r_h}{\rho_h}.$$

Writing,

$$e_{0h} = \frac{\overline{y}_h}{\overline{Y}_h} - 1 = (t_h - 1),$$

$$e_{1h} = \frac{\overline{x}_h}{\overline{X}_h} - 1 = (a_h - 1),$$

$$\eta_{0h} = \frac{s_{yh}^2}{S_{yh}^2} - 1,\ \eta_{1h} = \frac{s_{xh}^2}{S_{xh}^2} - 1 = (b_h - 1)$$

and $\delta_h = \dfrac{r_h}{\rho_h} - 1 = (c_h - 1)$

results in,

$$E(e_{0h}) = E(e_{1h}) = E(\eta_{0h}) = E(\eta_{1h}) = 0,$$

$$E(\delta_h) = \frac{1}{n_h}\big[3\rho_h(\lambda_{40h} + \lambda_{04h}) - 4(\lambda_{31h} + \lambda_{13h}) + 2\rho_h\lambda_{22h}\big]8\rho_h,$$

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

$$E\left(\eta_{0h}^2\right) = \frac{1}{n_h}(\lambda_{40h} - 1), E\left(\eta_{1h}^2\right) = \frac{1}{n_h}(\lambda_{04h} - 1),$$

$$E\left(\delta_h^2\right) = \frac{1}{n_h}[D_h], E(e_{0h}e_{1h}) = \frac{\lambda_{30h}}{n_h}C_{yh},$$

$$E(e_{0h}\eta_{1h}) = \frac{\lambda_{12h}}{n_h}C_{yh},\ E(e_{1h}\eta_{0h}) = \frac{\lambda_{21h}}{n_h}C_{xh},$$

$$E(e_{1h}\eta_{1h}) = \frac{\lambda_{03h}}{n_h}C_{xh},\ E(e_{0h}\delta_h) = \frac{A_{0h}}{n_h}C_{yh},$$

$$E(e_{1h}\delta_h) = \frac{A_{1h}}{n_h}C_{xh},\ E(\eta_{0h}\delta_h) = \frac{B_{0h}}{n_h},$$

$$E(\eta_{1h}\delta_h) = \frac{B_{1h}}{n_h}, E(\eta_{0h}\eta_{1h}) = \frac{1}{n_h}(\lambda_{22h} - 1),$$

where,

$$D_h = \big[\rho_h^2(\lambda_{40h} + \lambda_{04h}) - 4\rho_h(\lambda_{31h} + \lambda_{13h}) + 2\left(2 + \rho_h^2\right)\lambda_{22h}\big]/4\rho_h^2$$

$$A_{0h} = [2\lambda_{21h} - \rho_h(\lambda_{12h} + \lambda_{30h})]/2\rho_h$$

$$A_{1h} = [2\lambda_{12h} - \rho_h(\lambda_{21h} + \lambda_{03h})]/2\rho_h$$

$$B_{0h} = [2\lambda_{31h} - \rho_h(\lambda_{40h} + \lambda_{22h})]/2\rho_h$$

$$B_{1h} = [2\lambda_{13h} - \rho_h(\lambda_{04h} + \lambda_{22h})]/2\rho_h.$$

Using this background and following Srivastava (1971) a family of estimators of population mean $\overline{Y}$ may be defined as

$$\hat{\overline{Y}}_q = \sum_{h=1}^{L} W_h\overline{y}_h q_h(a_h), \qquad (1)$$

where $q_h(.)$ is a function of (a_h) such that $q_h(1) = 1$ and satisfies certain regularity conditions similar to those given by Srivastava (1971).

To the first degree of approximation, the variance of $\hat{\overline{Y}}_q$ is given by

$$V\left(\hat{\overline{Y}}_q\right) = \sum_{h=1}^{L} W_h^2\,\overline{Y}_h^2\,\frac{1}{n_h}\left[C_{yh}^2 + C_{xh}^2\,q_{h1}^2(1) + 2\rho_h C_{xh}C_{yh}q_{h1}(1)\right] \qquad (1.2)$$

which is minimized for

$$q_{h1}(1) = -\rho_h \frac{C_{yh}}{C_{xh}} \tag{1.3}$$

Thus, the resulting minimum variance of $\hat{\bar{Y}}_q$ is given by

$$\min.V\left(\hat{\bar{Y}}_q\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h} S_{yh}^2 (1-\rho_h^2) \tag{1.4}$$

Following Srivasrava and Jhajj (1981), Zaidi et. al. (1989) suggested a class of estimators of population mean $\bar{Y}$ as

$$\hat{\bar{Y}}_t = \sum_{h=1}^{L} W_h \bar{y}_h t_h\left(a_h, b_h\right) \tag{1.5}$$

where $t_h(.)$ is a function of (a_h, b_h) such that $t_h(1, 1) = 1$, which satisfies certain regularity conditions similar to those given by Srivastava and Jhajj (1981).

To the first degree of approximation the variance of $\hat{\bar{Y}}_t$ is given by

$$\begin{aligned} V\left(\hat{\bar{Y}}_t\right) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \frac{1}{n_h} \big[& C_{yh}^2 + C_{xh}^2 t_{h1}^2(1, 1) \\ & + (\lambda_{04h} - 1) t_{h2}^2(1, 1) \\ & + 2\rho_h C_{xh} C_{yh} t_{h1}(1, 1) \\ & + 2\lambda_{12h} C_{yh} t_{h2}(1, 1) \\ & + 2C_{xh}\lambda_{03h} t_{h1}(1, 1) t_{h2}(1, 1) \big] \end{aligned} \tag{1.6}$$

which is minimized for

$$\left.\begin{aligned} t_{h1}(1, 1) &= \frac{C_{yh}\left[\lambda_{12h}\lambda_{03h} - \rho_h(\lambda_{04h} - 1)\right]}{C_{xh}\left[\lambda_{04h} - \lambda_{03h}^2 - 1\right]} \\ t_{h2}(1, 1) &= \frac{C_{yh}\left[\rho_h \lambda_{03h} - \lambda_{12h}\right]}{\left[\lambda_{04h} - \lambda_{03h}^2 - 1\right]} \end{aligned}\right\} \tag{1.7}$$

and the minimum variance of $\hat{\bar{Y}}_t$ is given by

$$\min.V\left(\hat{\bar{Y}}_t\right) = \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^2}{n_h} \left[(1-\rho_h^2) - \frac{(\rho_h\lambda_{03h} - \lambda_{12h})^2}{(\lambda_{04h} - \lambda_{03h}^2 - 1)} \right] \tag{1.8}$$

The crux of this article is to suggest an optimum allocation with a family of estimators considered by Srivastava and Jhajj (1983) and compares its efficiency with that of Neyman allocation and others. It is seen that the proposed allocation is better in the sense of having lesser variance than other.

The Suggested Family of Estimators

Whatever the sample chosen, let $\left(a_h, b_h, c_h\right)$ assume values in a bounded closed convex subset, R of the three dimensional real space containing the point $(1, 1, 1)$. Let $g_h\left(a_h, b_h, c_h\right)$ be the function of a_h, b_h and c_h, such that $g_h(1, 1, 1) = 1$, and satisfies the following conditions:

1. In R, the function $g_h\left(a_h, b_h, c_h\right)$ is continuous and bounded.
2. The first and second partial derivatives of $g_h\left(a_h, b_h, c_h\right)$ exist and are continuous and bounded.

Define a family of estimators for population mean $\bar{Y}$ as

$$\hat{\bar{Y}}_g = \sum_{h=1}^{L} W_h \bar{y}_h g_h\left(a_h, b_h, c_h\right) \tag{2.1}$$

Expanding $g_h\left(a_h, b_h, c_h\right)$ about the point $(1, 1, 1)$ in a second order Taylor's series and noting that the second partial derivatives of *g* are bounded. We have

$$E\left(\hat{\bar{Y}}_g\right) = \bar{Y} + 0\left(n^{-1}\right),$$

so that bias of $\hat{\bar{Y}}_g$ is of the order of n^{-1}. Thus, to the first degree of approximation the variance of $\hat{\bar{Y}}_g$ is given by

$$V\left(\hat{\bar{Y}}_g\right) = E\left(\hat{\bar{Y}}_g - \bar{Y}\right)^2$$

$$= \sum_{h=1}^{L} W_h^2 \, \bar{Y}_h^2 \, \frac{1}{n_h} \Big[C_{yh}^2 + C_{xh}^2 \, g_{h1}^2(1,1,1) + (\lambda_{04h}-1) g_{h2}^2(1,1,1) + D_h \, g_{h3}^2(1,1,1) + 2\rho_h C_{xh} C_{yh} g_{h1}(1,1,1) + 2C_{yh}\,\lambda_{12h} g_{h2}(1,1,1) + 2C_{yh} A_{0h} g_{h3}(1,1,1) + 2C_{xh}\lambda_{03h} g_{h1}(1,1,1) g_{h2}(1,1,1) + 2C_{xh} A_{1h} g_{h1}(1,1,1) g_{h3}(1,1,1) + 2B_{1h} g_{h2}(1,1,1) g_{h3}(1,1,1) \Big] \quad (2.2)$$

where $g_{h1}(1,1,1)$, $g_{h2}(1,1,1)$ and $g_{h3}(1,1,1)$ denote the first order partial derivates of $g_h(a_h, b_h, c_h)$ at the point $(1,1,1)$. Differentiating (2.2) partially with respect to $g_{h1}(.)$, $g_{h2}(.)$ and $g_{h3}(.)$, and equating them to zero the following equations

$$\begin{bmatrix} C_{xh}^2 & C_{xh}\lambda_{03h} & C_{xh}A_{1h} \\ C_{xh}\lambda_{03h} & (\lambda_{04h}-1) & B_{1h} \\ C_{xh}A_{1h} & B_{1h} & D_h \end{bmatrix} \begin{bmatrix} g_{h1}(.) \\ g_{h2}(.) \\ g_{h3}(.) \end{bmatrix} = -C_{hy} \begin{bmatrix} \rho_h C_{xh} \\ \lambda_{12h} \\ A_{0h} \end{bmatrix} \quad (2.3)$$

Solving (2.3), the optimum values of $g_{h1}(.)$, $g_{h2}(.)$ and $g_{h3}(.)$ were obtained respectively as

$$g_{1h}(1,1,1) = \frac{C_{yh}}{K_h C_{xh}} \Big[\{\lambda_{12h}\lambda_{03h} - \rho_h(\lambda_{04h}-1)\} D_h + \{(\lambda_{04h}-1)A_{1h} - \lambda_{03h}B_{1h}\} A_{0h} - (\lambda_{12h}A_{1h} - \rho_h B_{h1}) B_{1h} \Big]$$

$$g_{h2}(1,1,1) = \frac{C_{yh}}{K_h} \Big[(\rho_h\lambda_{03h} - \lambda_{12h}) D_h + (\lambda_{03h}A_{1h} - B_{1h}) A_{0h} + (\lambda_{12h}A_{1h} - \rho_h B_{h1}) A_{1h} \Big]$$

$$g_{h3}(1,1,1) = \frac{C_{yh}}{K_h} \Big[\{\rho_h(\lambda_{04h}-1) - \lambda_{12h}\lambda_{03h}\} A_{h1} - (\lambda_{04h} - \lambda_{03h}^2 - 1) A_{0h} - (\rho_h\lambda_{03h} - \lambda_{1h2}) B_{1h} \Big]$$

where,

$$K_h = \Big[(\lambda_{04h} - \lambda_{03h}^2 - 1) D_h - (\lambda_{04h}-1) A_{1h}^2 + 2\lambda_{03h} A_{1h} B_{1h} - B_{1h}^2 \Big]$$

Thus, the minimum variance of $\left(\hat{\bar{Y}}_g\right)$ is given by

$$\min .V\left(\hat{\bar{Y}}_g\right) = \sum_{h=1}^{l} W_h^2 \frac{S_{yh}^2}{n_h} \Bigg[(1-\rho_h^2) - \frac{(\rho_h\lambda_{03h} - \lambda_{12h})^2}{(\lambda_{04h} - \lambda_{03h}^2 - 1)} - \frac{\{G_h\}^2}{K_h(\lambda_{04h} - \lambda_{03h}^2 - 1)} \Bigg] \quad (2.4)$$

where,

$$G_h = (\lambda_{12h}\lambda_{03h} - \rho_h\lambda_{04h} + \rho_h) A_{1h} + (\lambda_{04h} - \lambda_{03h}^2 - 1) A_{0h} + (\rho_h\lambda_{03h} - \lambda_{12h}) B_{1h}$$

In (2.4), the first term on the right hand side gives the minimum asymptotic variance of the family when only $\bar{X}_h$ is used, and the first two terms give the minimum asymptotic variance when both $\bar{X}_h$ and S_{hx}^2 are used. The third term gives the reduction in asymptotic variance when ρ_h is also used along with $\bar{X}_h$ and S_{hx}^2.

Efficiency Comparisons

It is known that the variance of usual unbiased estimators in stratified sampling under SRSWOR is

$$V(\bar{y}_{st}) = \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^2}{n_h} \quad (3.1)$$

From (1.4) and (3.1) the following results

$$V(\bar{y}_{st}) - \min .V\left(\hat{\bar{Y}}_q\right) = \sum_{h=1}^{L} W_h^2 \rho_h^2 \frac{S_{yh}^2}{n_h} \geq 0 \quad (3.2)$$

which, in turn, yields the inequality

$$\min .V\left(\hat{\bar{Y}}_q\right) \leq V(\bar{y}_{st}) \quad (3.3)$$

From (1.4) and (1.8)

$$\min.V\left(\hat{\hat{Y}}_q\right)-\min.V\left(\hat{\hat{Y}}_t\right)$$
$$=\sum_{h=1}^{L}W_h^2\frac{S_{yh}^2}{n_h}\frac{\{\rho_h\lambda_{03h}-\lambda_{12h}\}^2}{K_h(\lambda_{04h}-\lambda_{03h}^2-1)}\geq 0 \quad (3.4)$$

which gives the inequality

$$\min.V\left(\hat{\hat{Y}}_t\right)\leq \min.V\left(\hat{\hat{Y}}_q\right) \quad (3.5)$$

Further from (1.8) and (2.4)

$$\min.V\left(\hat{\hat{Y}}_t\right)-\min.V\left(\hat{\hat{Y}}_g\right)$$
$$=\sum_{h=1}^{L}W_h^2\frac{S_{yh}^2}{n_h}\frac{\{G_h\}^2}{K_h(\lambda_{04h}-\lambda_{03h}^2-1)}\geq 0 \quad (3.6)$$

which gives the inequality

$$\min.V\left(\hat{\hat{Y}}_g\right)\leq \min.V\left(\hat{\hat{Y}}_t\right) \quad (3.7)$$

Thus from (3.3), (3.5) and (3.7) we have

$$\min.V\left(\hat{\hat{Y}}_g\right)\leq \min.V\left(\hat{\hat{Y}}_t\right)\leq \min.V\left(\hat{\hat{Y}}_q\right)\leq V(\bar{y}_{st}) \quad (3.8)$$

It follows from (3.8) that the proposed estimator $\hat{\hat{Y}}_g$ is better than $\bar{y}_{st}$, $\hat{\hat{Y}}_q$ and $\hat{\hat{Y}}_t$ at its optimum conditions.

Optimum Allocation

The variance of $\bar{y}_{st}$ under the Neyman allocation

$$n_h=n\frac{W_hS_{yh}}{\sum_{h=1}^{L}W_hS_{yh}} \quad (4.1)$$

$$V(\bar{y}_{st})_N=\frac{1}{n}\left(\sum_{h=1}^{L}W_hS_{yh}\right)^2 \quad (4.2)$$

To minimize $\min.V\left(\hat{\hat{Y}}_q\right)$, $\min.V\left(\hat{\hat{Y}}_t\right)$ and $\min.V\left(\hat{\hat{Y}}_g\right)$, consider the cost function

$$C^*=C_0+\sum_{h=1}^{L}C_hn_h\,, \quad (4.3)$$

where C_0 and C_h are the overhead cost and cost per unit within h^{th} stratum respectively, for the given cost restriction

$$C_1n_1+C_2n_2+...+C_Ln_L=C^*-C_0 \quad (4.4)$$

Using Lagrange's method of multipliers, the optimum allocation in order to minimize $\min.V\left(\hat{\hat{Y}}_q\right)$, $\min.V\left(\hat{\hat{Y}}_t\right)$ and $\min.V\left(\hat{\hat{Y}}_g\right)$ respectively is

$$n_h=n\frac{W_hS_{yh}(1-\rho_h^2)^{1/2}\Big/\sqrt{C_h}}{\sum_{h=1}^{L}W_hS_{yh}(1-\rho_h^2)^{1/2}\Big/\sqrt{C_h}} \quad (4.5)$$

$$n_h=n\frac{W_hS_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}\right\}^{1/2}\Big/\sqrt{C_h}}{\sum_{h=1}^{L}W_hS_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}\right\}^{1/2}\Big/\sqrt{C_h}} \quad (4.6)$$

and

$$n_h=n\frac{W_hS_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}-\dfrac{\{G_h\}^2}{(\lambda_{04h}-\lambda_{03h}^2-1)K_h}\right\}^{1/2}\Big/\sqrt{C_h}}{\sum_{h=1}^{L}W_hS_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}-\dfrac{\{G_h\}^2}{(\lambda_{04h}-\lambda_{03h}^2-1)K_h}\right\}^{1/2}\Big/\sqrt{C_h}} \quad (4.7)$$

In particular, if $C_h=C$ for the given cost function $C^*=C_0+nC$, the optimum allocation (4.5), (4.6) and (4.7) respectively reduce to

$$n_h=n\frac{W_hS_{yh}(1-\rho_h^2)^{1/2}}{\sum_{h=1}^{L}W_hS_{yh}(1-\rho_h^2)^{1/2}} \quad (4.8)$$

$$n_h = n\frac{W_h S_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}\right\}^{1/2}}{\sum_{h=1}^{L} W_h S_{yh}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}\right\}^{1/2}} \tag{4.9}$$

and

$$n_h = n\frac{W_h S_{hy}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}-\dfrac{\{G_h\}^2}{(\lambda_{04h}-\lambda_{03h}^2-1)K_h}\right\}^{1/2}}{\sum_{h=1}^{L} W_h S_{hy}\left\{(1-\rho_h^2)-\dfrac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}-\dfrac{\{G_h\}^2}{(\lambda_{04h}-\lambda_{03h}^2-1)K_h}\right\}^{1/2}} \tag{4.10}$$

Substituting the values of n_h from (4.8), (4.9) and (4.10) respectively in (1.4), (1.8) and (2.4) the resulting variances of $\hat{\bar{Y}}_q$, $\hat{\bar{Y}}_t$ and $\hat{\bar{Y}}_g$ are

$$\min.V\left(\hat{\bar{Y}}_q\right)_O = \frac{1}{n}\left[\sum_{h=1}^{L} W_h S_{yh}(1-\rho_h^2)^{1/2}\right]^2 \tag{4.11}$$

$$\min.V\left(\hat{\bar{Y}}_t\right)_O = \frac{1}{n}\left[\sum_{h=1}^{L} W_h S_{yh}\left\{(1-\rho_h^2)-\frac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}\right\}^{1/2}\right]^2 \tag{4.12}$$

and

$$\min.V\left(\hat{\bar{Y}}_g\right)_O = \frac{1}{n}\left[\sum_{h=1}^{L} W_h S_{yh}\left\{(1-\rho_h^2)-\frac{(\rho_h\lambda_{03h}-\lambda_{12h})^2}{(\lambda_{04h}-\lambda_{03h}^2-1)}-\frac{\{G_h\}^2}{(\lambda_{04h}-\lambda_{03h}^2-1)K_h}\right\}^{1/2}\right]^2 \tag{4.13}$$

From (4.2), (4.11), (4.12) and (4.13) it can be easily proved that

$$\min.V\left(\hat{\bar{Y}}_g\right)_O \le \min.V\left(\hat{\bar{Y}}_t\right)_O \le \min.V\left(\hat{\bar{Y}}_q\right)_O \le V(\bar{y}_{st})_N, \tag{4.14}$$

which clearly indicates that the proposed optimum allocation is better than Neyman allocation $(\bar{y}_{st})$ and the optimum allocation based on Srivastava (1971) family of estimators and the optimum allocation envisaged by Zaidi et al., (1989) in the sense of having smaller variance.

Empirical Study

The performance of various families of estimators of the population mean $\bar{Y}$ through six natural population data sets has been illustrated.

To examine the performance of the estimators $\hat{\bar{Y}}_q$, $\hat{\bar{Y}}_t$ and $\hat{\bar{Y}}_g$ with respect to $\bar{y}_{st}$ under optimum allocation we have computed the percent relative efficiencies of *t* with respect to $\bar{y}_{st}$ using the formula,

$$PRE(t,\ \bar{y}_{st}) = \frac{V(\bar{y}_{st})_N}{\min V(t)_O}\times 100,$$

where t = $\hat{\bar{Y}}_q$, $\hat{\bar{Y}}_t$, $\hat{\bar{Y}}_g$; results are presented in Table 5.1.

Conclusion

Table 5.1 clearly indicates that the proposed family of estimator $\hat{\bar{Y}}_g$ is more efficient than the usual unbiased estimator $\bar{y}_{st}$, $\hat{\bar{Y}}_q$ and the Zaidi, et al. (1989) estimator, $\hat{\bar{Y}}_t$. Thus the proposed family of estimator $\hat{\bar{Y}}_g$ would be preferred over $\bar{y}_{st}$, $\hat{\bar{Y}}_q$ and $\hat{\bar{Y}}_t$.

Table 5.1: Percent Relative Efficiencies of $\hat{\hat{Y}}_q$, $\hat{\hat{Y}}_t$, and $\hat{\hat{Y}}_g$ with respect to $\bar{y}_{st}$

Population	$PRE\left(\hat{\hat{Y}}_q, \bar{y}_{st}\right)$	$PRE\left(\hat{\hat{Y}}_t, \bar{y}_{st}\right)$	$PRE\left(\hat{\hat{Y}}_g, \bar{y}_{st}\right)$
I	872.12	879.51	2308.29
II	351.30	367.04	690.30
III	420.66	496.89	571.88
IV	856.61	984.67	1746.53
V	615.88	727.70	1003.45
VI	147.64	242.84	362.15

Population I: Singh and Chaudhary (1986, p. 162)
y: total number of trees, x: area under orchards in ha.

$N = 25$, $L = 3$, $N_1 = 6$, $N_2 = 8$, $N_3 = 11$

Stratum — Values of parameters for h^{th} stratum

No.	S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	273.45103	0.9215191	-0.2276668	-0.071714	-0.2400887	0.138323
2	509.03212	0.9737715	1.6980145	1.6304126	1.7646005	1.576411
3	256.6819	0.8826909	1.0289035	0.8472329	1.2344161	0.5897102

Stratum — Values of parameters for h^{th} stratum (continued)

No.	λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	1.2773905	1.3483853	1.5310737	1.239425	1.3741684
2	4.4920977	4.7537207	4.2700966	4.6186087	4.3727487
3	3.264646	4.3492128	2.684855	3.7646968	2.8334168

For illustration take $n = 10$, $n_1 = 3$, $n_2 = 3$, $n_3 = 4$

Population II: Singh and Mangat (1996, p. 194)
y: pocket money, x: annual income

$N = 27$, $L = 3$, $N_1 = 4$, $N_2 = 10$, $N_3 = 13$

Stratum No.	S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	225.46249	0.9527907	0.9817665	0.9616631	0.9637509	0.906753
2	108.14085	0.8074107	0.1045162	0.0851702	0.0745106	-0.0097243
3	98.871841	0.7621946	-0.1720774	-0.0129786	-0.0879664	-0.1103153

Values of parameters for h^{th} stratum (continued)

Stratum No.	λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	2.1256188	2.1872063	2.1224402	2.1470526	2.1142848
2	1.4455092	1.7719919	2.1393301	1.484715	1.5986642
3	1.6145628	1.9933334	1.5608654	1.6582907	1.3338932

For illustration take $n = 10$, $n_1 = 2$, $n_2 = 4$, $n_3 = 5$

Population III: Singh and Mangat (1996, p. 207)
y: no. refrigerators sold in current year, x: no. refrigerators sold last summer

$N = 42$, $L = 4$, $N_1 = 14$, $N_2 = 9$, $N_3 = 12$, $N_4 = 7$

Stratum No.	S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	12.911576	0.7929927	-0.019159	0.3665704	-0.3717353	0.8009986
2	13.201431	0.8697081	0.4460543	0.402637	0.4681387	0.3062423
3	15.05344	0.9191256	-0.1618712	-0.2565663	-0.128619	-0.4344209
4	13.062123	0.9055795	0.2273419	-0.0915551	0.5905558	-0.3916206

Values of parameters for h^{th} stratum (continued)

Stratum No.	λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	1.8121436	2.2006301	3.3060221	1.7701281	2.263858
2	1.5135141	2.2975185	1.6129147	1.7937746	1.4355898
3	1.928372	1.9632339	2.7733335	1.815768	2.2420385
4	1.7822884	2.4742281	1.9126016	2.0034381	1.7549122

For illustration take $n = 16$, $n_1 = 5$, $n_2 = 3$, $n_3 = 5$, $n_4 = 3$

Population IV: Singh and Mangat (1996, p. 212)
y: leaf area for newly developed strain of wheat, x: weight of leaves

$N = 39$, $L = 3$, $N_1 = 12$, $N_2 = 13$, $N_3 = 14$

Stratum No.	S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	6.3362112	0.9202367	0.429305	0.5097853	0.23599	0.5031633
2	5.5075918	0.9154022	0.9960984	0.815551	1.0341649	0.5847596
3	6.7413528	0.9668189	0.2057622	0.2971175	0.083846	0.3360654

Values of parameters for h^{th} stratum (continued)

Stratum No.	λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	1.9123464	2.2748233	1.9394547	2.0257975	1.879711
2	2.970998	3.436904	2.9819269	3.0966741	2.9303901
3	2.5134376	2.8955496	2.3448986	2.6759523	2.3988602

For illustration take $n = 14$, $n_1 = 4$, $n_2 = 5$, $n_3 = 5$

Population V: Singh and Mangat (1996, p. 218)
y: juice quantity, x: weight of cane

$N = 25$, $L = 3$, $N_1 = 6$, $N_2 = 12$, $N_3 = 7$

Stratum No.	S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	8.9442719	0.9455626	0.576173	0.6492226	0.4598407	0.688919
2	15.05042	0.948196	0.9857208	0.9738854	0.9465183	0.9187277
3	10.965313	0.7532234	1.0354011	0.8915649	0.8581802	0.727283

Values of parameters for h^{th} stratum (continued)

Stratum No.	λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	2.2641624	2.2865633	2.3437501	2.2586791	2.2886912
2	3.379509	3.2689734	3.792407	3.2777466	3.5484598
3	2.3117711	3.1306353	2.3294286	2.487514	2.2170337

For illustration take $n = 10$, $n_1 = 3$, $n_2 = 4$, $n_3 = 3$

Population VI: Singh and Mangat (1996, p. 219)
y: total number of milch cows 1993, x: total number of milch cows 1990

$N=24$, $L=3$, $N_1=7$, $N_2=12$, $N_3=5$

Stratum No.	Values of parameters for h^{th} stratum S_{yh}	ρ_h	λ_{12h}	λ_{21h}	λ_{03h}	λ_{30h}
1	4.197505	0.7654592	-0.4418403	-0.4494459	0.0382842	-0.324885
2	4.0778411	0.4066542	-0.2762718	-0.2448949	0.1507925	-0.6181979
3	3.6469165	0.4945774	-0.8119799	-0.2847418	-0.569229	-0.0912794

Stratum No.	Values of parameters for h^{th} stratum (continued) λ_{22h}	λ_{04h}	λ_{40h}	λ_{13h}	λ_{31h}
1	1.1348072	1.8497596	1.6555367	1.0929828	1.3169373
2	0.5695984	2.312027	2.7509735	0.8349021	0.6748404
3	1.3461457	1.8333916	1.5925434	1.1123488	1.0704605

For illustration take $n=10$, $n_1=3$, $n_2=5$, $n_3=2$

References

Singh, D., & Chaudhary, F. S. (1986). *Theory and analysis of sample survey designs*. New Delhi, India: Wiley Eastern Ltd.

Singh, R., & Mangat, N. S. (1996). *Element of survey sampling.* London, England: Kluwer Academic Publishers.

Srivastava, S. K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *Journal of the American Statistical Association*, *66*, 404-407.

Srivastava, S. K., & Jhajj, H. S. (1981). A class of estimators of the population mean in survey sampling using auxiliary information. *Biometrika*, *68*, 341-343.

Srivastava, S. K., & Jhajj, H. S. (1983). Class of estimators of mean and variance using auxiliary information when correlation coefficient is known. *Biometrical Journal*, *24*(4), 401-409.

Zaidi, S., Shanker, U., & Singh, R. K. (1989). An optimum allocation with a class of estimators using auxilary information. *Journal of Statistical Research*, *23*, 1-4.