**Research Paper** **Engineering**

# Ratio-Product Type Estimators for Population Mean in Heterogeneous Population

*** R. K. Gangele ** S. K. Choubey**

* Department of Mathematics and Statistics, Dr. H.S. Gour Central University, Sagar – 470003 (M.P.) INDIA

** Department of Mathematics and Statistics, Dr. H.S. Gour Central University, Sagar – 470003 (M.P.) INDIA

**ABSTRACT**

*In this paper we have suggested an exponential Ratio product type estimators for population mean in heterogeneous population stricture. The bias and mean squired error are obtained to the first degree approximation we observed that , proposed estimator found better than usual ratio and product estimators*

**Keywords:**

### 1.1. Introduction

In modern surveys the scientific technique for selecting a sample is that of selecting a probability sample that is usually based upon a stratification of the population. It is well known that stratification is one of the design tools that gives increased precision. In stratified design, the population under investigation is divided into different strata so as to obtain the homogeneity within each stratum, and sample observations are drawn within each stratum generally by well known procedure of simple random sampling. It is known that the proper use of supplementary information on auxiliary variable(s) may lead to more efficient estimators. Out of many ratio, regression and product methods of estimation are good examples in this context. It is noted that the regression estimator of population mean is the most efficient estimator. The ratio (product) estimator of population mean is equally good if the regression line is a straight line and passes through the origin. However in many practical situations the regression line does not pass through the origin. This led various authors including Srivastava (1967), Walsh (1970), Reddy (1974), Gupta (1978), Vos (1980), Naik and Gupta (1991), Upadhyaya and Singh (1999), Singh (2003), Singh and Ruiz Espejo (2003), Singh and Vishwakarma (2006a, b) and Sharma and Tailor (2010) to modify ratio (product) estimator to get better alternatives. Keeping this in view we have made an effort to propose separate and combined ratio-product estimators for estimating population mean $\bar{Y}$ using auxiliary information in stratified random sampling and analyze their properties.

Consider a fi·i·e population $U=(U_1,U_2,U_3,\dots,U_N)$ of size $N$ and let $y$ and $x$ respectively, be the study and auxiliary variables on each unit $U_j$ $(j=1,2,3,\dots,N)$ of the population *U.* Let the population be divided into $L$ strata with the $h^{th}$-stratum containing $N_h$ units, *h=1,2,3,…, L* so that $\sum_{h=1}^{L} N_h = N$

.Suppose that a simple random sample of size $\sum_{h=1}^{L} n_h = n$ is drawn without

replacement (SRSWOR) from the $h^{th}$-stratum such that

. Let $(y_{hi}, x_{hi})$ denote the observed values of $y$ and $x$ on $i^{th}$-unit of the $h^{th}$-stratum $(i=1,2,3,\dots,N_h;\ h=1,2,3,\dots,L)$. Moreover let us denote by

$$\left(\bar{y}_h=\sum_{i=1}^{n_h} y_{hi}/n_h,\ \bar{y}_{st}=\sum_{h=1}^{L} W_h\bar{y}_h\right),\ \left(\bar{x}_h=\sum_{i=1}^{n_h} x_{hi}/n_h,\ \bar{x}_{st}=\sum_{h=1}^{L} W_h\bar{x}_h\right)$$

and

$$\left(\bar{Y}_h=\sum_{i=1}^{N_h} y_{hi}/N_h,\ \bar{Y}=\sum_{h=1}^{L} W_h\bar{Y}_h\right),\ \left(\bar{X}_h=\sum_{i=1}^{N_h} x_{hi}/N_h,\ \bar{X}=\sum_{h=1}^{L} W_h\bar{X}_h\right)$$

the sample and population means of $y$ and $x$ , where $W_h=N_h/N$.

In order to have a survey estimate of the population mean $\bar{Y}$ of the main variable *y*, assuming the knowledge of population mean $\bar{X}_h$ of the $h^{th}$-stratum $(h=1,2,3,\dots,L)$ of the auxiliary variable *x,* we mention the following well-known estimators. The separate ratio estimator

$$\bar{y}_{RS}=\sum_{h=1}^{L} W_h\hat{R}_h\bar{X}_h,\quad x_h\neq 0 \qquad 1.1.1$$

where $\hat{R}_h=\bar{y}_h/\bar{x}_h$ , is the estimate of the ratio $R_h=\bar{Y}_h/\bar{X}_h$ , $\bar{X}_h\neq 0$ of the $h^{th}$-stratum in the population. This estimator is only efficient if the variables are strongly positively correlated. The separate product estimator

$$\bar{y}_{PS}=\sum_{h=1}^{L} W_h\frac{\hat{P}_h}{\bar{X}_h} \qquad 1.1.2$$

where $\hat{P}_h=\bar{y}_h.\bar{x}_h$ is the estimate of the product $P_h=\bar{Y}_h.\bar{X}_h$ of the means of the $h^{th}$-stratum in the population. This estimator will often be used if the two variables are supposed to be strongly negatively correlated.

When the population mean $\bar{X}$ of the auxiliary variate *x* is known, *Hansen, Hurwitz, and Gurney (1946)* suggested a "combined ratio estimator"

$$\bar{y}_{RC}=\bar{y}_{st}\left(\frac{\bar{X}}{\bar{x}_{st}}\right) \qquad 1.1.3$$

The "combined product estimator" for $\bar{Y}$ is defined by

$$\bar{y}_{PC}=\bar{y}_{st}\left(\frac{\bar{x}_{st}}{\bar{X}}\right) \qquad 1.1.4$$

To the first degree of approximation, the biases and variances of $\bar{y}_{RS}$ , $\bar{y}_{PS}$ and $\hat{\bar{Y}}_{RS}^{(c)}$ are respectively given by

$$B(\bar{y}_{RS})=\frac{1}{\bar{X}}\sum_{h=1}^{L} W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left(R_h S_{xh}^2-S_{hxy}\right) \qquad 1.1.5$$

$$V(\bar{y}_{RS})=\sum_{h=1}^{L} W_h^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)\left(S_{yh}^2+R^2S_{xh}^2-2R_h S_{hxy}\right) \qquad 1.1.6$$

$$B(\bar{y}_P) = \frac{1}{\bar{X}}\sum_{h=1}^{L} W_h^2\left(\frac{1}{n_h} - \frac{1}{N_h}\right) S_{hxy} \quad 1.1.7$$

$$V(\bar{y}_P) = \sum_{h=1}^{L} W_h^2\left(\frac{1}{n_h} - \frac{1}{N_h}\right)\left(S_{hy}^2 + R^2 S_{hx}^2 + 2R S_{hxy}\right) \quad 1.1.8$$

where $C_{hx}^2 = \frac{S_{hx}^2}{\bar{X}_h^2}, C_{hy}^2 = \frac{S_{hy}^2}{\bar{Y}_h^2}, R = \frac{\bar{Y}}{\bar{X}}, K_h = \rho_{hxy}\frac{C_{hy}}{C_{hx}}, \; S_{hy}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y}_h)^2,$

$S_{hx}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(x_{hi} - \bar{X}_h)^2, S_{hxy} = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y}_h)(x_{hi} - \bar{X}_h)$ and $\rho_{hxy} = \frac{S_{hxy}}{S_{hx}S_{hy}}.$

In stratified random sampling, the total sample size is $n = \sum_{i=1}^{L} n_i$.

If an equivalent simple random sample of size *n* were selected without replacement directly from the population size *N*, the variance of the mean $\left(\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j\right)$ per unit and variance of the usual ratio estimator $\left(\bar{y}_R = \bar{y}.\bar{X}/\bar{x}\right)$ to the first degree of approximation are respectively given by,

$$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2 C_y^2 \quad 1.1.9$$

and

$$V(\bar{y}_R) = \left(\frac{1}{n} - \frac{1}{N}\right)\bar{Y}^2\left[C_y^2 + C_x^2(1 - 2K)\right] \quad 1.1.10$$

where $C_x^2 = \frac{S_x^2}{\bar{X}^2}, C_y^2 = \frac{S_y^2}{\bar{Y}^2}, K = \rho\frac{C_y}{C_x}, \; S_y^2 = \frac{1}{(N-1)}\sum_{h=1}^{L}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y})^2,$

$S_x^2 = \frac{1}{(N-1)}\sum_{h=1}^{L}\sum_{i=1}^{N_h}(x_{hi} - \bar{X})^2, S_{xy} = \frac{1}{(N-1)}\sum_{h=1}^{L}\sum_{i=1}^{N_h}(y_{hi} - \bar{Y})(x_{hi} - \bar{X})$ and $\rho = \frac{S_{xy}}{S_x S_y}.$

In this paper we have suggested exponential ratio-product type estimators for estimating the population mean using auxiliary information in stratified sampling and analyses their properties. These estimators are compared for their precision with simple mean per unit in stratified sampling, usual ratio, usual product estimator. An empirical study is also carried out to compare the performance of the suggested estimator with that of others.

**1.2. Proposed ratio-product estimator**

We propose the following combined ratio-product estimator for estimating population mean $\bar{Y}$ as

$$\hat{\bar{Y}}_G^{(c)} = \bar{y}_s\left\{\alpha\exp\left(\frac{\bar{X} - \bar{x}_s}{\bar{X} + \bar{x}_s}\right) + (1-\alpha)\exp\left(\frac{\bar{x}_s - \bar{X}}{\bar{x}_s + \bar{X}}\right)\right\} \quad 1.2.1$$

where $\alpha$ is a real constant to be determined such that the variance of $\hat{\bar{Y}}_G^{(c)}$ is a minimum. To obtain the variance of $\hat{\bar{Y}}_G^{(c)}$ to the first degree of approximation, we write

$$\bar{y}_s = \sum_{h=1}^{L} W_h\bar{y}_h = \bar{Y}(1 + e_0)$$

and

$$\bar{x}_s = \sum_{h=1}^{L} W_h\bar{x}_h = \bar{X}(1 + e_1)$$

such that,

$$E(e_0) = E(e_1) = 0$$

,

under SRSWOR, we have

$$\left.\begin{aligned} E(e_0^2) &= \frac{1}{\bar{Y}^2}\sum_{h=1}^{L} W_h^2\gamma_h S_{hy}^2 \\ E(e_1^2) &= \frac{1}{\bar{X}^2}\sum_{h=1}^{L} W_h^2\gamma_h S_{hx}^2 \\ E(e_0 e_1) &= \frac{1}{\bar{Y}\bar{X}}\sum_{h=1}^{L} W_h^2\gamma_h S_{hyx} \end{aligned}\right\} \quad 1.2.2$$

Expressing equation (1.2.1) in terms of e's we have

$$\hat{\bar{Y}}_G^{(c)} = \bar{Y}(1 + e_0)\left[\alpha\exp\left\{\frac{-e_1}{2}\left(1 + \frac{e_1}{2}\right)^{-1}\right\} + (1-\alpha)\exp\left\{\frac{e_1}{2}\left(1 + \frac{e_1}{2}\right)^{-1}\right\}\right] \quad 1.2.3$$

We now assume that $|e_1| < 1$ so that we may expand $\left(\frac{e_1}{2 + e_1}\right)$ as a series in powers of $e_1$. Expanding multiplying out retaining terms of e's to the second degree, we obtain

$$\hat{\bar{Y}}_G^{(c)} = \bar{Y}\left\{1 + e_0 + \frac{e_1}{2} - \frac{e_1^2}{8} + \frac{e_0 e_1}{2} - \alpha e_1 - \alpha e_0 e_1 + \alpha\frac{e_1^2}{2}\right\}$$

$$\hat{\bar{Y}}_G^{(c)} - \bar{Y} = \bar{Y}\left\{e_0 + \frac{e_1}{2} - \frac{e_1^2}{8} + \frac{e_0 e_1}{2} - \alpha\left(e_1 + e_0 e_1 - \frac{e_1^2}{2}\right)\right\} \quad 1.2.4$$

Taking expectations of both sides of (1.2.4) and using the result (1.2.2), we obtain the bias of $\hat{\bar{Y}}_G^{(c)}$ to order $o(n^{-1})$ as

$$Bias\left(\hat{\bar{Y}}_G^{(c)}\right) = \frac{1}{\bar{X}}\left[\frac{1}{2}\sum_{h=1}^{L} W_h^2\gamma_h\left(S_{hyx} - \frac{R}{4}S_{hx}^2\right) - \alpha\sum_{h=1}^{L} W_h^2\gamma_h\left(S_{hyx} - \frac{R}{2}S_{hx}^2\right)\right] \quad 1.2.5$$

Squaring both sides of equation (1.2.5) and again retaining terms to the second degree, we have

$$\left(\hat{\bar{Y}}_G^{(c)} - \bar{Y}\right)^2 = \bar{Y}^2\left[e_0 - \frac{e_1}{2}(2\alpha - 1)\right]^2$$

$$\left(\hat{\bar{Y}}_G^{(c)} - \bar{Y}\right)^2 = \bar{Y}^2\left[e_0^2 + \frac{e_1^2}{4}(2\alpha - 1)^2 + e_0 e_1(2\alpha - 1)\right] \quad 1.2.6$$

Taking expectations of both sides of (1.2.6) and using the result (1.2.2), we obtain the MSE of $\hat{\bar{Y}}_G^{(c)}$ to order $o(n^{-1})$ as

$$MSE\left(\hat{\bar{Y}}_G^{(c)}\right) = \bar{Y}^2\left[\frac{1}{\bar{Y}^2}\sum_{h=1}^{L} W_h^2\gamma_h S_{hy}^2 + \frac{(2\alpha - 1)^2}{4\bar{X}^2}\sum_{h=1}^{L} W_h^2\gamma_h S_{hx}^2 - \frac{(2\alpha - 1)}{\bar{Y}\bar{X}}\sum_{h=1}^{L} W_h^2\gamma_h S_{hxy}\right]$$

$$= \left[\sum_{h=1}^{L} W_h^2\gamma_h S_{hy}^2 + (2\alpha - 1)^2\frac{R^2}{4}\sum_{h=1}^{L} W_h^2\gamma_h S_{hx}^2 - (2\alpha - 1)R\sum_{h=1}^{L} W_h^2\gamma_h S_{hxy}\right] \quad 1.2.7$$

which is minimized for

$$\alpha = \frac{1}{2}\left[1 + \frac{\sum_{h=1}^{L} W_h^2\gamma_h S_{hxy}}{R\sum_{h=1}^{L} W_h^2\gamma_h S_{hx}^2}\right] = \alpha_{opt} \quad 1.2.8$$

**1.3. Comparisons**

**It is well known under stratified random sampling that**

$$V(\bar{y}_s) = \sum_{h=1}^{L} W_h^2\gamma_h S_{hy}^2 \quad 1.3.1$$

From (1.3.1) and (1.2.7) we have

$$Var(\bar{y}_s) - Var\left(\hat{\bar{Y}}_G^{(c)}\right) > 0 \quad 1.3.2$$

From (1.1.6) and (1.2.7) we have

$$Var\left(\hat{\bar{Y}}_{Rst}^{(c)}\right) - Var\left(\hat{\bar{Y}}_G^{(c)}\right) > 0 \quad 1.3.3$$

From (1.1.8) and (1.2.7) we have

$$Var\left(\hat{\bar{Y}}_{Pst}^{(c)}\right) - Var\left(\hat{\bar{Y}}_G^{(c)}\right) > 0 \quad 1.3.4$$

**1.4. Empirical Study**

To see the performance of various estimators of population mean $\bar{Y}$, with respect to usual unbiased estimator $\bar{y}_s$, we have considered three data sets. Summaries of the data are given below:

**Population I-** Source: *Singh and Chaudhary (1986, p. 162)*

*y* : total number of trees,
*x* : area under orchards in ha.

| Total | Stratum → | 1 | 2 | 3 |
|---|---|---|---|---|
| $N = 25$ | $N_h$ | 6 | 8 | 11 |
| $n = 10$ | $n_h$ | 3 | 3 | 4 |
| $\bar{X} = 8.3792$ | $\bar{X}_h$ | 6.813 | 10.12 | 7.967 |
| $\bar{Y} = 410.84$ | $\bar{Y}_h$ | 417.33 | 503.375 | 340.00 |
| $S_x^2 = 59.737$ | $S_{xh}^2$ | 15.9712 | 132.66 | 38.438 |
| $S_y^2 = 123769.57$ | $S_{yh}^2$ | 74775.467 | 259113.70 | 65885.60 |
| $\rho = 0.9285341$ | $\rho_h$ | 0.9215191 | 0.9737715 | 8826909 |
| $R = 49.0309$ | $\gamma_h$ | 0.1666667 | 0.2083333 | 0.1590909 |
| $\rho_c = 0.9409238$ | $W_h^2$ | 0.0576 | 0.1024 | 0.1936 |

**Population I1: [Source: Singh and Mangat (1996), p. 212]**
y : leaf area for the newly developed strain of wheat,
x : weight of leaves.

| Total | Stratum → | 1 | 2 | 3 |
|---|---|---|---|---|
| $N = 39$ | $N_h$ | 12 | 13 | 14 |
| $n = 14$ | $n_h$ | 4 | 5 | 5 |
| $\bar{X} = 106.2307$ | $\bar{X}_h$ | 103.41667 | 110.92308 | 104.28571 |
| $\bar{Y} = 26.8479$ | $\bar{Y}_h$ | 25.754167 | 28.94 | 25.842857 |
| $S_x^2 = 124.1129$ | $S_{xh}^2$ | 133.90152 | 66.24359 | 154.98901 |
| $S_y^2 = 38.9952$ | $S_{yh}^2$ | 40.147572 | 30.333567 | 45.445837 |
| $\rho = 0.9385614$ | $\rho_h$ | 0.9202367 | 0.9154022 | 0.9668189 |
| $R = 0.25273$ | $\gamma_h$ | 0.1666667 | 0.1230769 | 0.1285714 |
| $\rho_c = 0.936302$ | $W_h^2$ | 0.0946746 | 0.1111111 | 0.1288626 |

**Population III- Source: Singh and Mangat (1996, p. 219)**
y : juice quantity,
x : weight of cane

| Total | Stratum → | 1 | 2 | 3 |
|---|---|---|---|---|
| $N = 25$ | $N_h$ | 6 | 12 | 7 |
| $n = 10$ | $n_h$ | 3 | 4 | 3 |
| $\bar{X} = 326$ | $\bar{X}_h$ | 366.666 | 310.833 | 317.143 |
| $\bar{Y} = 102.6$ | $\bar{Y}_h$ | 135.00 | 99.166 | 80.714 |
| $S_x^2 = 2700.05$ | $S_{xh}^2$ | 2706.666 | 1881.06 | 2890.476 |
| $S_y^2 = 558.566$ | $S_{yh}^2$ | 80.00 | 226.515 | 120.238 |
| $\rho = 0.7314955$ | $\rho_h$ | 0.9455626 | 0.948196 | 0.7532234 |
| $R = 0.3147$ | $\gamma_h$ | 0.1666667 | 0.1666667 | 0.1904762 |
| $\rho_c = 0.8676778$ | $W_h^2$ | 0.0576 | 0.2304 | 0.0784 |

The percent relative efficiency of an estimator $t \left(= \bar{y}_{st}, \bar{y}_{RC}, \bar{y}_{PC}, \hat{\bar{Y}}_{GC}^{(c)}\right)$ with respect to usual unbiased estimator $\bar{y}_{st}$ is defined by

$$PRE(t, \bar{y}_{st}) = \frac{Var(\bar{y}_{st})}{Var(t)} \times 100 \qquad 1.4.1$$

We have computed the percent relative efficiency of $\bar{y}_{st}$, $\bar{y}_{RC}$, $\bar{y}_{PC}$ and $\hat{\bar{Y}}_{GC}^{(c)}$ with respect to $\bar{y}_{st}$ and presented in Table 1.4.1.

**Table 1.4.1**

**PRE of different estimators of $\bar{Y}$ with respect to $\bar{y}_s$**

| | Population-I | | Population-II | | Population-III | |
|---|---|---|---|---|---|---|
| *Estimator* | *MSE* | *PRE* | *MSE* | *PRE* | *MSE* | *PRE* |
| $\bar{y}_{st}$ | 8274.88 | 100.00 | 1.80 | 100.00 | 11.26 | 100.00 |
| $\bar{y}_{RC}$ | 3759.04 | 220.13 | 0.66 | 271.72 | 11.05 | 101.90 |
| $\bar{y}_{PC}$ | * | * | * | * | * | * |
| $\hat{\bar{Y}}_{GC}^{(c)}$ | 2780.33 | 297.62 | 0.61 | 295.08 | 11.01 | 103.10 |

* data not applicable

**1.5. Conclusion**

Table 1.4.1 clearly indicates the proposed combined ratio-product $\hat{\bar{Y}}_G^{(c)}$ estimator is more efficient than usual estimator , ratio estimator $\bar{y}_{Pst}$ and product estimator $\bar{y}_s$ in stratified sampling procedure. It is also shown that, if the correlation between study variate and auxiliary variate x is negative, the proposed estimator $\hat{\bar{Y}}_G^{(c)}$ is always better than others. Thus, the use of the proposed class of estimators should be preferred in practice.

**REFERENCES**

Srivastava, S.K. (1967): An estimator using auxiliary information in sample surveys. Cal. Statist. Assoc. Bull., 16, 121-132. | Walsh, J.E. (1970): Generalization of ratio estimates for population total. Sankhya, A, 32, 99-106. | Reddy, V.N. (1974): On a transformad ratio method of estimation, Sankhya, 36, C, 59-70. | Gupta, P.C. (1978): On some quadratic and higher degree ratio and product estimators. Jour. Ind. Soc. Agril. Statist. , 30, 71-80. | Vos, J.W.E. (1980): Mixing of direct, ratio and product method of estimation. Statistica Neerlandica, 34, 4, 209-218. | Naik, V.D. and Gupta, P.C. (1991): A General class of estimators for estimating population mean using auxiliary information. Metrika, 38, 11-19. | Upadhyaya, L.N. and Singh, H.P. (1999): Use of transformed auxiliary variable in estimating the finite population mean. Biom. Jour.. 41, 5, 627-636 | Singh, S. (2003): Advanced sampling theory with application. How Michael selected Amy. Vol. I and II, Kluwer Academic Publishers, Dordrecht, The Netherlands. | Singh, H.P. and Ruiz Espejo, M. (2003b): On linear regression and ratio-product estimation of a finite population mean. The Statistician, 52, 1, 59-67. | Singh, H.P. and Vishwakarma, G.K. (2006 a): An efficient variant of the product and ratio estimators in stratified random sampling. Statistics in Transition, 7, 6, 1311-1325. | Sharma, B. and Tailor, R. (2010): A new ratio-cum-dual to ratio estimator of finite population mean in simple random sampling. Global Jour. Sci. Frontier Res., 10(1), 27-31. |