# Ratio cum Product Method of Estimation

By M. P. Singh, Calcutta[1])

*Summary:* In this paper methods of estimation which may be considered as combination of ratio and product methods have been suggested. The mean square errors of these estimators utilizing two supplementary variables are compared with (i) simple unbiased estimator ($p = 0$), (ii) usual ratio and product methods of estimation ($p = 1$) and (iii) multivariate ratio and multivariate product estimators ($p = 2$), where $p$ is the number of supplementary variables utilized. Conditions for their efficient use have been obtained for each case. Extension to general case of $p$-variables has been briefly discussed. A new criteria for the efficient use of product estimator have been obtained.

## 1. Introduction

The problem of improving upon the usual unbiased estimator by the use of single suitable supplementary variable has received considerable attention in sampling theory and practice. Ratio, product and regression estimators are the typical example of such methods, of these, ratio and product estimators being easily obtainable are more prevalent in practice. Murthy [4] obtained the specific conditions under which any of the three estimators namely ratio, product and unbiased estimators can be efficiently used. Assuming that both the parameters are positive and coefficient of variations are equal these conditions are such that ratio or product estimator is to be preferred according as the estimation variable is highly positively or negatively correlated with that of supplementary variable.

However, in large scale sample surveys we often collect data on more than one supplementary character without any increase in cost and some of these characters may be correlated with the characteristics under study. Suppose the information on $k$-characters $X_i$ ($i = 1, 2, \dots, k$) correlated with the character under study ($Y$) are available in the survey instead of only $X_1$ and for simpleness in the estimation procedure we are allowed to use any two of them. We meet the problem 'which two should be chosen' and how to utilize them so as to yield more efficient estimator of the popula-

---

[1]) M. P. Singh, Indian Statistical Institute, Calcutta.

tion mean (or total) than the estimators when either one of them (or none) is used. Let us suppose $X_1$ and $X_2$ are the two chosen characteristics, then we may face two situations (i) $X_1$ and $X_2$ are such that one of them (say $X_1$) is suitable to be used as ratio estimator and the other ($X_2$) as product estimator, then the question arises which one of the two to choose or on the other hand how to use both $X_1$ and $X_2$ simultaneously to yield estimators better than the corresponding ratio or product estimators (ii) $X_1$ and $X_2$ are such that simple unbiased estimator happens to be more efficient than either ratio or product estimator utilizing information either on $X_1$ or $X_2$ respectively. In such cases we have to examine under what circumstances the estimators which utilize both $X_1$ and $X_2$ yield efficient results. For the purpose of constructing such estimators we may use the concept of multiple correlation coefficient of $Y$ on $X_1$ and $X_2$ defined by

$$\rho^2_{0.12} = \frac{\rho^2_{01} + \rho^2_{02} - 2\,\rho_{01}\,\rho_{02}\,\rho_{12}}{1 - \rho^2_{12}} \tag{1.1}$$

instead of the value of $\rho_{01}$ or $\rho_{02}$ which is used for usual ratio and product methods of estimation. Where $\rho_{01}$, $\rho_{02}$ and $\rho_{12}$ are the simple correlation coefficient between $Y$ and $X_1$; $Y$ and $X_2$ and $X_1$ and $X_2$. By some suitable transformation we examine for what values of $\rho_{01}$, $\rho_{02}$ and $\rho_{12}$ the value $\rho^2_{0.12}$ attains its maximum and minimum values. Putting $\rho_{01} = f\cos\theta$ and $\rho_{02} = f\sin\theta$, we get

$$\rho^2_{0.12} = f^2\,\frac{1 - A\,\rho_{12}}{1 - \rho^2_{12}} \tag{1.2}$$

where $A = \text{Sin}\,2\,\theta$ and that range of $A$ is $-1$ to $+1$; it is observed that value of $\rho^2_{0.12}$ can be obtained for different values of $A$ and $\rho_{12}$. We note that $A = \pm 1$ implies $\rho_{01} = \pm\,\rho_{02}$. For different values of $A$ and $\rho_{12}$ the value of $\rho^2_{0.12}$ can be tabulated. It will be observed that $\rho^2_{0.12}$ becomes larger for

(a) $\rho_{01} = -\,\rho_{02}$ and $\rho_{12} \to +1$

and

(b) $+\,\rho_{01} = +\,\rho_{02}$ and $\rho_{12} \to -1$.

Now from this increasing tendency of $\rho^2_{0.12}$ under the above conditions the characteristics $X_1$ and $X_2$ can be used together on an analogy of ratio (using the positively correlated characters in the form of ratio) and product estimators (using negatively correlated characters in the form of product) we get the following three forms of the ratio cum product estimators.

$$\hat{Y}_1 = \left(\frac{y\,x_2}{x_1}\right)\left(\frac{X_1}{X_2}\right) \tag{1.3}$$

$$\hat{Y}_2 = \left(\frac{y}{x_1\,x_2}\right)(X_1\,X_2) \tag{1.4}$$

$$\hat{Y}_3 = \frac{(y\,x_1\,x_2)}{(X_1\,X_2)}\,. \tag{1.5}$$

Where $y$, $x_1$ and $x_2$ are the unbiased estimate of the parameters $Y$, $X_1$ and $X_2$, based on any sample design. Perhaps it may be more appropriate to call the second and third type estimators as ratio-type and product-type respectively.

## 2. Bias and Mean Square Error

For considering the efficiency of these estimators we shall derive the expression for their bias and the mean square error to the second degree of approximation. Writing

$$y = Y(1 + e_0), \qquad x_i = X_i(1 + e_i), \qquad i = 1, 2 \tag{2.1}$$

where $E(e_i) = 0$ and $|e_i| < 1$, for $i = 0, 1, 2$.

Let us denote $B_j$ and $M_j$ ($j = 1, 2, 3$) as the bias and m.s.e. of $\hat{Y}_1$, $\hat{Y}_2$ and $\hat{Y}_3$. Then we get

$$B_1 = B_R + B_P - Y\,C_1\,C_2\,\varrho_{12} \tag{2.2}$$

$$B_2 = B_R - B_P + Y(C_1^2 + C_1\,C_2\,\varrho_{12}) \tag{2.3}$$

$$B_3 = B_P - Y(C_0\,C_1\,\varrho_{01} + C_1\,C_2\,\varrho_{12}) \tag{2.4}$$

$$M_1 = M_R + Y^2(C_2^2 + 2\,C_0\,C_2\,\varrho_{02} - 2\,C_1\,C_2\,\varrho_{12}) \tag{2.5}$$

$$M_2 = M_R + Y^2(C_2^2 - 2\,C_0\,C_2\,\varrho_{02} + 2\,C_1\,C_2\,\varrho_{12}) \tag{2.6}$$

$$M_3 = M_P + Y^2(C_1^2 + 2\,C_0\,C_1\,\varrho_{01} + 2\,C_1\,C_2\,\varrho_{12})\,. \tag{2.7}$$

Where $C_i$ ($i = 0, 1, 2$) is coefficient of $y$, $x_1$ and $x_2$. $B_R$, $B_P$, $M_R$ and $M_P$ are the bias and m.s.e. of the usual ratio and product estimators defined by $\hat{Y}_R = (y/x_1)\,X_1$ and $\hat{Y}_P = (y\,x_2)/X_2$.

## 3. Comparison of the Estimators

In this section comparisons of ratio cum product estimators have been made under three different categories of estimators. (i) Estimator with no supplementary character (simple unbiased), (ii) Estimators which utilize one supplementary character (ratio and product), (iii) Estimators which utilize more than one supplementary characters (multivariate ratio and multivariate product). It is interesting to note that in the second case the conditions for efficient use of ratio cum product estimators as against either ratio or product estimators turns out to be the same giving similar configurations $S_R$ and $S_P$ as obtained by the author [8]. These configurations therefore are being omitted here, we only state these conditions for preference of either of the estimators in this case.

### 3.1 *Ratio cum product vs unbiased estimator*

In case no additional information is used in the estimation procedure the variance of the unbiased estimate $y$ is given by

$$V_u = Y^2 C_0^2. \tag{3.1.1}$$

The equation (3.1.1) is compared with eqs. (2.5), (2.6) and (2.7) under the assumption (ii) of Sec. 1 to yield the conditions for preference of $\hat{Y}_1$, $\hat{Y}_2$ and $\hat{Y}_3$. The respective conditions for $C_i = C$ ($i = 0, 1, 2$) can be written as

$$(\varrho_{01} - \varrho_{02}) > (1 - \varrho_{12}) \tag{3.1.2}$$

$$(\varrho_{01} - \varrho_{02}) > (1 + \varrho_{12}) \tag{3.1.3}$$

and

$$(\varrho_{01} + \varrho_{02}) < -(1 + \varrho_{12}). \tag{3.1.4}$$

These regions for preference can much better be effected with the help of Fig. 1 given below.

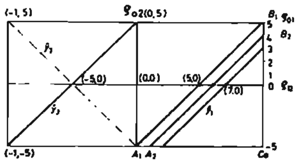


Fig. 1

From the above figure it is observed that for all plausible pair of values of $\varrho_{01}$, $\varrho_{02}$, $\hat{Y}_1$ will be more efficient than the usual unbiased estimator in the triangular regions $A_1 B_1 C_0$, for $\varrho_{01} \leqslant 0.5$, in region $A_2 B_2 C_0$, for $\varrho_{01} \leqslant 0.4$ and so on.

3.2 *Ratio cum product vs ratio and product estimator*

Following the assumptions and notations of Sec. 2 we have the m.s.e. of ratio and product estimators as

$$M_R = Y^2(C_0^2 + C_1^2 - 2\,C_0\,C_1\,\varrho_{01}) \tag{3.2.1}$$

$$M_P = Y^2(C_0^2 + C_2^2 + 2\,C_0\,C_2\,\varrho_{02}). \tag{3.2.2}$$

Comparing the eq. (3.2.1) with (2.5) and (2.6) we get the following conditions (3.2.3) and (3.2.4) under which $\hat{Y}_1$ and $\hat{Y}_2$ respectively are more efficient than $\hat{Y}_R$. These conditions are

$$\varrho_{12} > \varrho_{02}\left(\frac{C_0}{C_1}\right) + \frac{1}{2}\left(\frac{C_2}{C_1}\right) \tag{3.2.3}$$

$$\varrho_{12} < \varrho_{02}\left(\frac{C_0}{C_1}\right) - \frac{1}{2}\left(\frac{C_2}{C_1}\right). \tag{3.2.4}$$

Similarly the estimators $\hat{Y}_1$ and $\hat{Y}_3$ will be more efficient than $\hat{Y}_P$ under the following conditions (3.2.5) and (3.2.6) obtained by comparing (3.2.2) with (2.6) and (2.7) respectively. We get

$$\varrho_{12} > \frac{1}{2}\left(\frac{C_1}{C_2}\right) - \varrho_{01}\left(\frac{C_0}{C_2}\right) \tag{3.2.5}$$

$$\varrho_{12} < -\frac{1}{2}\left(\frac{C_1}{C_2}\right) - \varrho_{01}\left(\frac{C_0}{C_2}\right). \tag{3.2.6}$$

These two sets of conditions are same as obtained by SINGH [8] and hence gives rise to the similar configurations $S_R$ and $S_P$ for the preference of either of these estimators.

3.3 *Ratio cum product vs multivariate ratio estimator*

OLKIN [5] suggested the use of information on more than one supplementary characteristics in the estimation procedure in the form of multivariate ratio estimator defined by

$$\hat{Y}_{RK} = \sum_{i=1}^{K} w_i\, r_i\, X_i \tag{3.3.1}$$

where $r_i = y/x_i$ and weights $w_i$'s are such that $\sum_{i=1}^{k} w_i = 1$. The m.s.e. for this estimator in any sampling design for $k = 2$ and $w_1 = w_2$ is given by

$$M(\hat{Y}_{R2}) = Y^2\left[\frac{C^2}{2}(1+\varrho_{12}) + C_0^2 - C_0 C(\varrho_{01}+\varrho_{02})\right] \qquad (3.3.2)$$

where

$$C_1 = C_2 = C. \qquad (3.3.3)$$

For comparison of this estimator with the $\hat{Y}_1$, we put the condition (3.3.3) in eq. (2.5) and compare the resulting expression with the eq. (3.3.2). It is observed that $\hat{Y}_1$ is more efficient than $Y_{R2}$ if

$$\frac{\varrho_{01} - 3\,\varrho_{02}}{3 - 5\,\varrho_{12}} > \frac{1}{2}\left(\frac{C}{C_0}\right). \qquad (3.3.4)$$

It may be verified that these conditions is met with in a number of situations of practical interest.

### 3.4 *Ratio cum product vs multivariate product estimator*

Singh [7] has given multivariate product estimator similar to Olkin's multivariate ratio estimator and obtained the specific conditions under which either of these two may be preferred in contrast to simple unbiased estimator in any sampling design, which turned out to be similar as that obtained by Murthy [4]. The multivariate product estimator is defined by

$$\hat{Y}_{PK} = \sum_{i=1}^{K} \frac{w_i p_i}{X_i} \qquad (3.4.1)$$

where $p_i = y\, x_i$ and weights $w_i$'s are such that $\sum_{i=1}^{K} w_i = 1$. The m.s.e. for this estimator for $K = 2$ and $w_1 = w_2$ under the condition (3.3.3) is given by

$$M(\hat{Y}_{P2}) = Y^2\,[\tfrac{1}{2}C^2(1+\varrho_{12}) + C_0^2 + C\,C_0(\varrho_{01}+\varrho_{02})]. \qquad (3.4.2)$$

Again for comparison of this estimator with $\hat{Y}_1$ we obtain the equation (2.5) under the condition (3.3.3) and compare the resulting equation with (3.4.2) to yield the conditions under which $\hat{Y}_1$ is more efficient than $\hat{Y}_{P2}$, we get this condition as

$$\frac{3\,\varrho_{01} - \varrho_{02}}{3 - 5\,\varrho_{12}} > \frac{1}{2}\left(\frac{C}{C_0}\right). \qquad (3.4.3)$$

Conditions similar to (3.3.4) and (3.4.3) can be obtained for the efficient use of the $\hat{Y}_2$ and $\hat{Y}_3$ also on the similar lines.

It is interesting to note that the m.s.e. of these estimators can be expressed in terms of the intra-class correlation coefficient in the fashion similar to that of unbiased estimator for the case of systematic samples. Here we give the final expressions which can easily be derived following Singh [6]. We get

$$(M_R)\, s\, y = (M_R) \text{ ran } (1 + \overline{n-1}\, \varrho)$$

$$(M_P)\, s\, y = (M_P) \text{ ran } (1 + \overline{n-1}\, \varrho)$$

$$(M_J)\, s\, y = (M_J) \text{ ran } (1 + \overline{n-1}\, \varrho)$$

where $M_J$ ($j = 1, 2, 3$) is m.s.e. of $\hat{Y}_j$ and $\varrho_y = \varrho_{x_1} = \varrho_{x_2} = \varrho$ is the intra-class correlation coefficient for $y$, $x_1$ and $x_2$. This shows that even in case ratio cum product estimators systematic sampling can be made more efficient by arranging the units such that $\varrho$ becomes negative.

## 4. A New Criteria for Preference of Product Estimators

Let us consider the ratio cum product estimator $\hat{Y}_1$ and assume that $x_1 \equiv x_2 \equiv x$. Therefore we can write

$$\hat{Y}_1 = y = \frac{(y\, x)}{x}. \tag{4.1}$$

Following Goodman [1] and an identity derived for the ratio of two random variables by Koop [5]; and after some simplifications, we get

$$V(y) = \text{Cov}\left\{(y\,x)^2, \left(\frac{1}{x}\right)^2\right\} - \left\{\text{Cov}\left(y\,x, \frac{1}{x}\right)\right\}^2 - 2\,\text{Cov}\left(y\,x, \frac{1}{x}\right) E\left(\frac{1}{x}\right) E(y\,x) + $$
$$+ V\left(\frac{1}{x}\right) V(y\,x) + \{E(y\,x)\}^2\, V\left(\frac{1}{x}\right) + E\left(\frac{1}{x}\right)^2 V(y\,x). \tag{4.2}$$

Now for the situations where $y$ is inversely proportional to $x$ we get,

$$V(y) > \frac{V(y\,x)}{X^2}. \tag{4.3}$$

Again after some simplification, we can write

$$V(y) = \left[1 + \alpha + C_{1/x}^2(1 - \delta^2) + 2\,\frac{C_{1/x}}{C_{yx}}\left(\frac{C_{1/x}}{2\,C_{yx}} - \delta\right)\right] V(y\,x) \left\{E\left(\frac{1}{x}\right)\right\}^2 \geqslant$$
$$\geqslant [1 + \alpha + \beta + \gamma]\, \frac{V(y\,x)}{X^2}$$

where

$$\alpha = \text{Cov}\left\{(y\,x)^2, \left(\frac{1}{x}\right)^2\right\} \Big/ \left\{E\left(\frac{1}{x}\right)\right\}^2 \cdot V(y, x)$$

$$\beta = C_{1/x}^2(1 - \delta^2)$$

$$\gamma = 2\,\frac{C_{1/x}}{C_{yx}}\left(\frac{C_{1/x}}{2\,C_{yx}} - \delta\right)$$

$$\delta = \text{Cov}\left(y\,x, \frac{1}{x}\right) \Big/ \sqrt{V(y\,x)\;V\left(\frac{1}{x}\right)}$$

and

$$E\left(\frac{1}{x}\right) \geqslant \frac{1}{E(x)} \quad \text{and} \quad E(x) = X.$$

Thus if the expression in the large bracket is greater than equal to one we get other situations in which product estimator happens to be more efficient than the unbiased estimator. And that this condition can be satisfied in a variety of situations, e.g. $\beta$ always $> 0$, hence even if $(\alpha + \gamma) \geqslant 0$, the condition is fulfilled. The positivity of $\alpha$ and $\gamma$ will not be much difficult to realize in practice. It may be noted that these expressions are exact and no approximation of any order have been used.

## 5. Use of Multi-Supplementary Variables

In this section we consider the extension of ratio cum product estimator to multi-supplementary variables defined as

$$\hat{Y}_{RPK} = y \cdot \prod_{i=1}^{k_1} \frac{x_i}{X_i} \prod_{j=k_1+1}^{k} \frac{X_j}{x_j} \tag{5.1}$$

where $x_i$ and $x_j$ is an unbiased estimate of $X_i$ and $X_j$ respectively. This estimator for a particular value of $k$ can be constructed on the principle of maximization of multiple correlation coefficient $\varrho^2_{0.12\ldots k}$, similar to the case $k = 2$. For $k > 2$, there would be number of situations for which $\varrho^2_{0.12\ldots k}$ is maximized. It may be noted that for $k = k_1 = 1$ we get the usual product and for $k_1 = 0$, $k = 1$, the usual ratio estimators as particular cases of this estimator and similarly double ratio estimator also happen to be estimator of this family for $k = 4$. Mean square error of $\hat{Y}_{RPK}$ is given by

$$M(\hat{Y}_{RPK}) = Y^2\left[\sum_{i=0}^{K} C_i^2 + 2 \sum_{i(\neq i')=0}^{k_1} C_{i'} C_i \varrho_{ii'} - 2 \sum_{j(\neq j')=k_1+1}^{k} C_j C_{j'} \varrho_{jj'} - 2 \sum_{i(\neq j)=0}^{k} C_i C_j \varrho_{ij}\right]. \tag{5.2}$$

For different values of $k$ and $k_1$ this m.s.e. can be obtained and conditions of efficient use of this estimator can be given following the approach in section 3. In general it will be observed that whenever the supplementary variables available are both positively and negatively correlated with the variable under study ratio cum product estimator may result precise estimates than the other methods considered here.

The author is grateful to Dr. M. N. MURTHY, Indian Statistical Institute, for his guidance in the preparation of this paper.

**References**

[1] GOODMAN, L. A., (1960): On the exact variance of products. *Amer. Stat. Asso.* 55, 708—713.

[2] HANSEN, M. H., et al.(1953): Sample Survey Methods and Theory. Vol. I. Wiley, New York.

[3] KOOP, J. C., (1964): "On an identity for the variance of a ratio of two random variables." *Jour. Roy. Stat. Soc.* Series B, 26(3), 484—486.

[4] MURTHY, M. N., (1964): Product Method of Estimation. *Sankhya,* Series A, Vol. 26, 69—74.

[5] OLKIN, I., (1958): Multivariate ratio method of estimation for finite populations. *Biometrika,* 45, 154—65.

[6] SINGH, M. P., (1965): Multivariate product method of estimation for finite populations. (*Submitted for publication.*)

[7] SINGH, M. P., (1966): Efficient use of systematic sampling in ratio and product estimation. *Metrika,* 10, 3, 199—205.

[8] SINGH, M. P., (1965): On the estimation of ratio and product of population parameters. *Sankhya,* B, 27, 321—28.