Jour. Ind. Soc. Ag. Statistics
45(2), 1993 : 177-186

A General Class of Product-type Estimators Under Super-population Model

V.K. Singh, G.N. Singh* and D. Shukla[□]
Banaras Hindu University, Varanasi - 221 005
(Received : August, 1989)

Summary

The problem of constructing classes of estimators for population mean has been widely discussed by various authors under design approach in sample surveys. An attempts by Upadhyaya *et al* [9] has been made to combine the usual mean and product estimator with suitable weights in order to define a general class. This paper is an attempt to study properties of the same estimator under super-population model. Consequently, optimum choice of weights has theoretically been obtained. Results have been supported with some numerical examples.
Key Words : Product estimation, Super-population, Optimisation, Bias, Mean Square Error.

Introduction

The product method of estimation is generally used when the study variable Y is negatively correlated with an auxiliary characteristics X whose population mean is assumed to be known. In order to improve the efficiency of product estimation, sometimes product-type estimators are used which are developed by mixing product estimator with usual mean estimator. Some of the important works in this direction are Ray *et al* [3], Srivenkataramana [7], Vos [10], etc. It is to note that such estimators generally fall in the most general class of product-type estimators given by $T_p = w_1 \bar{y} + w_2 \bar{y}_p$; where w_1 and w_2 are unknown weights which are either specified or estimated and $\bar{y}$ and $\bar{y}_p$ are respectively mean estimator and usual product estimator. Although T_p has been observed to be more efficient than $\bar{y}$ and $\bar{y}_p$ under different situations in design approach, no concised study of its properties has been done under super-population model approach.

The present work is devoted to the study of the estimator T_p under super-population model with uncorrelated errors and a gamma distributed auxiliary characteristic X. The Bias and Mean Squared Error (MSE) of T_p are obtained. Further, minimising the MSE of the

* Punjab University, Chandigarh.

□ Devi Ahilya Vishwavidyalaya, Indore.

estimator, optimum choices of weights w_1 and w_2 are derived. For a few combinations of the parametric values, relative efficiencies of T_p with respect to $\bar{y}$ and $\bar{y}_p$ have been obtained.

2. *Bias and MSE of* T_p.

Let a sample of size n be drawn from a finite population of size N using simple random sampling without replacement strategy. Let $(\bar{Y}, \bar{X})$ and $(\bar{y}, \bar{x})$ denote the population and sample mean of the study variable Y and the auxiliary characteristic X based on N and n units respectively. The usual product estimator is then defined as

$$\bar{y}_p = \bar{y}\,\frac{\bar{x}}{\bar{X}} \tag{1}$$

We consider the following general class of product-type estimators proposed by Upadhyaya *et al* [9]:

$$T_p = w_1\,\bar{y} + w_2\,\bar{y}_p \tag{2}$$

with $w_1 + w_2 \neq 1$.

Let us consider that the finite population of size N is a sample from a super-population and the relation between Y and X of the form

$$y_i = \alpha + \beta x_i + e_i \qquad (i = 1, 2, \ldots, N) \tag{3}$$

where α and β are unknown real constants and e_i's are random errors such that

$$E_c\,(e_i \mid x_i) = 0 \tag{4}$$

$$E_c\,(e_i\, e_j \mid x_i, x_j) = 0 \quad \text{for } i \neq j \tag{5}$$

and $$E_c\,(e_i^2 \mid x_i) = \delta\, x_i^g\,; \qquad \delta > 0,\ 0 \leq g \leq 2 \tag{6}$$

E_c denotes the conditional expectation given x_i (i=1, 2, . . ., N). We assume that x_i's are independently and identically distributed gamma variates with common density

$$f(x) = \frac{1}{\Gamma\theta}\, e^{-x}\, x^{\theta-1}\,; \quad x > 0,\ \theta \geq 1 \tag{7}$$

Let us denote the expectation with respect to the common distribution of x_i by E_x, model expectation by E_m (= E_xE_c) and design expectation by E_d.

It is to be mentioned here that the model (3) to (6) and the density

(7) are those taken by Durbin [2], Tin [8], Rao and Webster [4], Shah and Gupta [6] and Sahoo [5].

In order to evaluate the model expectation E_m make use of the lemma 3.2 given by Chaubey *et al* [1], which is as follows:

Let $X_1, X_2, \ldots, X_n$ be N independently and identically distributed gamma variates with parameter θ, then for given non-negative numbers $m_1, m_2, \ldots, m_p$ and k, we have

$$E\left[\frac{X_{i_1}^{m_1}\; X_{i_2}^{m_2} \ldots X_{i_p}^{m_p}}{\bar{X}^k}\right] = \frac{X_{i_1}^{m_1}\; X_{i_2}^{m_2} \ldots X_{i_p}^{m_p}}{E[T^s]}\; E[T^{s-k}]\, N^k \qquad (8)$$

where $\{i_1, i_2, \ldots, i_p\}$ is a subset of p distinct elements from $\{1, 2, \ldots, N\}$;

$$S = \sum_{j=1}^{p} m_j, \quad T = \sum_{j=1}^{N} X_j \quad \text{and} \quad \bar{X} = \frac{T}{N}$$

The bias of T_p is given by

$$B(T_p) = E_m\, E_d\, [w_1\, (\bar{y}-\bar{Y}) + w_2(\bar{y}_p-\bar{Y}) + (w_1+w_2-1)\, \bar{Y}\,] \qquad (9)$$

$$= w_2\, \frac{\beta\theta\, (N-n)}{n\, (N\theta+1)} + R\, (\alpha+\beta\theta) \qquad (10)$$

where $R = (w_1 + w_2 - 1)$

Similarly MSE of the estimator will be

$$M(T_p) = E_m\, E_d\, [w_1\, (\bar{y}-\bar{Y}) + w_2(\bar{y}_p-\bar{Y}) + R\bar{Y}\,]^2$$

Now since $E_m = E_x E_c$, we have

$$M(T_p) = E_x E_c E_d\, [w_1^2\, (\bar{y}-\bar{Y})^2 + w_2^2(\bar{y}_p-\bar{Y})^2 + R^2\bar{Y}^2 +$$

$$2w_1w_2\, (\bar{y}-\bar{Y})\, (\bar{y}_p-\bar{Y}) + 2w_1R\, (\bar{y}-\bar{Y})\bar{Y} + 2w_2R\, (\bar{y}_p-\bar{Y})\bar{Y}]$$

Remembering that $E_d\, (\bar{y}-\bar{Y})\, \bar{Y} = 0$ and writing

$$(\bar{y}-\bar{Y}) = \beta\, (\bar{x}-\bar{X}) + (\bar{e}_n - \bar{e}_N)\ ;$$

$$(\bar{y}_p-\bar{Y}) = \alpha\left(\frac{\bar{x}}{\bar{\bar{X}}} - 1\right) + \beta\left(\frac{\bar{x}^2}{\bar{\bar{X}}} - \bar{X}\right) + \left(\frac{\bar{e}_n\, \bar{x}}{\bar{\bar{X}}} - \bar{e}_N\right)$$

where $$\bar{e}_n = \frac{1}{n}\sum_{i=1}^{n} e_i\ , \quad \bar{e}_N = \frac{1}{N}\sum_{i=1}^{N} e_i$$

we have, due to the result (8)

$$M(T_p) = w_1^2 A + w_2^2 B + 2w_1 w_2 C + R^2 D + 2w_2 RE \tag{11}$$

where
$$A = \frac{N-n}{Nn}\left[\beta^2\theta + \frac{\delta\,\Gamma(\theta+g)}{\Gamma\theta}\right], \tag{12}$$

$$B = \alpha^2 B_1 + \beta^2 B_2 + 2\alpha\beta B_3 + \delta B_4 \tag{13}$$

with
$$B_1 = \frac{N-n}{n(N\theta+1)}, \tag{14}$$

$$B_2 = \frac{N^2\theta(n\theta+1)(n\theta+2)(n\theta+3)}{n^3(N\theta+2)(N\theta+3)} - \frac{2\theta(n\theta+1)}{n} + \frac{\theta(N\theta+1)}{N}, \tag{15}$$

$$B_3 = \frac{N^2\theta(n\theta+1)(n\theta+2)}{n^2(N\theta+1)(N\theta+2)} - \frac{N\theta(n\theta+1)}{n(N\theta+1)}, \tag{16}$$

$$B_4 = \frac{\Gamma(\theta+g)}{\Gamma\theta}\left[\frac{N^2(n\theta+g)(n\theta+g+1)}{n^3(N\theta+g)(N\theta+g+1)} - \frac{2(n\theta+g)}{n(N\theta+g)} + \frac{1}{N}\right], \tag{17}$$

$$C = \frac{\alpha\beta(N-n)\theta}{n(N\theta+1)} + \beta^2\left[\frac{N\theta}{n^2(N\theta+2)}\{n\theta(n\theta+3)+2\} - \frac{\theta(n\theta+1)}{n}\right] + \frac{\delta\,(N-n)}{n^2}\,\frac{\Gamma(\theta+g)}{\Gamma\theta}\,\frac{(n\theta+g)}{(N\theta+g)}, \tag{18}$$

$$D = \alpha^2 + \frac{\beta^2(N\theta+1)\theta}{N} + \frac{\delta\,\Gamma(\theta+g)}{\Gamma\theta\, N} + 2\alpha\beta\theta \tag{19}$$

and
$$E = \frac{\alpha\beta(N-n)\theta}{n(N\theta+1)} + \frac{\beta^2(N-n)\theta}{nN} + \frac{\delta\,\Gamma(\theta+g)}{\Gamma\theta} - \frac{(N-n)g}{nN(N\theta+g)}. \tag{20}$$

It can be seen that for $w_1=0$, $w_2=1$; $T_p = \bar{y}\,\bar{x}/\bar{X}$ which is usual product estimator. Similarly, for $w_1=N/(N-n)$, $w_2=-n/(N-n)$, T_p reduces to dual to ratio estimator considered by Srivenkataramana [7]. The bias and MSE of these estimators under the given super-population model have been obtained by Shah and Gupta [6] and Sahoo [5] respectively.

3. *Optimum Choices of w_i (i=1, 2)*

Since w_i (i=1, 2) are unknown weights and a specific choice of these yields a particular member of the class T_p, it is desirable to detect that member of the class which has minimum MSE. This can be achieved by minimising MSE expression with respect to the unknown constants w_i. Differentiating the expression (11) successively with respect to w_1 and w_2 and equating them to zero, we have the optimum choices of w_i (i=1, 2) as follows :

$$w_1 = \frac{D(B+D+2E) - (D+E)(C+D+E)}{(A+D)(B+D+2E) - (C+D+E)^2} \quad (21)$$

$$w_2 = \frac{(A+D)(D+E) - D(C+D+E)}{(A+D)(B+D+2E) - (C+D+E)^2} \quad (22)$$

These weights when substituted in the expression (11) produce the minimum MSE.

Numerical Example

In order to get an insight of the efficiency of the proposed estimator T_p under the optimality condition some numerical illustrations are presented. The example has been taken from Sahoo [5]. Here N=60, δ=2.0, θ=8.0. Tables 1–4 present relative efficiencies of T_p over $\bar{y}$ and $\bar{y}_p$ for α=0.00(0.50)1.50, β=0.5(0.5)1.5, g = 0.0(0.5)2.0 and n = 10(10)40. In the tables $E_1 = 100E_m v(\bar{y})/M(T_p)$ and $E_2 = 100M(\bar{y}_p)/M(T_p)$. Since the MSE of T_p has been minimised, substantial gain over $\bar{y}$ and $\bar{y}_p$ is expected which is apparent from the tables.

REFERENCES

[1] Chaubey, Y.P., Dwivedi, T.D. and Singh, M., 1984. An efficiency comparison of product and ratio estimator, *Communication in Statistics*, **13(6)**, 699–709.

[2] Durbin, J., 1959. A note on the application of Quenouille's method of bias reduction to the estimation of ratios, *Biometrika*, **46**, 477–480.

[3] Ray, S.K., Sahi, A. and Sahai, A., 1979. A note on ratio and product-type estimators, *Annals of the Institute of Mathematical* Statistics, **31**, 141–144.

[4] Rao, J.N.K. and Webster, J.T., 1966. On two methods of bias reduction in the estimation of ratios, *Biometrika*, **53**, 571–577.

[5] Sahoo, L.N., 1986. A note on the efficiency of a product-type estimator under a super-population model, *Journal of the Indian Society of Agricultural Statistics*, **38(3)**, 383–387.

[6] Shah, D.N. and Gupta, M.R., 1987. An efficiency comarison of dual ratio and product estimators, *Communication in Statistics*, **16(3)**.

[7] Srivenkataramana, T., 1980. A dual to ratio estimator in sample surveys, *Biometrika*, **67(1)**, 199–204.

[8] Tin, M., 1965. Comparison of some ratio estimators, *Journal of American Statistical Association* **60**, 294–307.

[9] Upadhyaya, L.N., Singh, H.P. and Vos, J.W.E., 1985. On the estimation of population means and ratios using supplementary information, *Statistica Neerlandica*, **39(3)**, 309–318.

[10] Vos, J.W.E., 1980. Mixing of direct, ratio and product method estimators, *Statistica Neerlandica*, **34**, 209–218.

Table 1. Relative efficiencies of the proposed estimator Tp with $\bar{y}$ and $\bar{y}_p$

α = 0.00									
g	β	n = 10		n = 20		n = 30		n = 40	
		E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
0.0	0.5	192.2	491.2	196.4	494.5	198.0	495.9	198.8	496.7
	1.0	450.7	1568.8	478.0	1636.8	488.1	1662.5	493.5	1676.0
	1.5	825.4	3127.5	919.4	3426.3	956.7	3545.6	976.8	3609.8
0.5	0.5	136.1	249.7	135.7	245.4	135.7	244.0	135.6	243.4
	1.0	233.6	662.6	239.1	667.1	241.1	669.0	242.1	670.0
	1.5	390.7	1317.2	408.7	1355.3	415.3	1369.6	418.8	1377.0
1.0	0.5	120.2	165.2	115.4	155.7	113.8	152.7	113.0	151.2
	1.0	149.9	308.1	149.7	302.4	149.7	300.5	149.7	299.6
	1.5	207.4	550.8	210.1	548.4	211.0	547.8	211.5	547.5
1.5	0.5	128.4	150.0	113.8	129.8	123.9	123.2	106.4	120.0
	1.0	123.4	183.3	119.3	173.5	118.0	170.3	117.3	168.7
	1.5	140.5	264.8	138.8	256.5	138.3	253.8	138.0	252.5
2.0	0.5	168.0	184.1	127.9	136.1	114.6	120.7	107.9	113.1
	1.0	125.3	152.2	113.3	133.8	109.3	127.9	107.3	125.0
	1.5	122.3	170.3	116.3	157.8	114.3	153.7	113.3	151.7

Table 2. Relative efficiencies of the proposed estimator Tp with $\bar{y}$ and $\bar{y}_p$

α = 0.5									
g	β	n = 10		n = 20		n = 30		n = 40	
		E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
0.0	0.5	193.7	545.2	197.1	547.2	198.4	548.0	199.0	548.5
	1.0	455.7	1679.3	480.4	1742.6	489.4	1766.2	494.2	1778.6
	1.5	836.5	3297.5	925.1	3588.0	959.9	3702.7	978.5	3764.1
0.5	0.5	136.1	268.5	135.8	264.0	135.7	262.6	135.7	261.9
	1.0	234.8	701.2	239.6	704.5	241.4	705.9	242.3	706.6
	1.5	393.1	1376.3	409.9	1411.8	416.0	1425.0	419.1	1431.9
1.0	0.5	118.8	170.1	114.9	161.6	113.6	158.9	112.9	157.5
	1.0	150.1	321.2	149.8	315.2	149.8	313.3	149.7	312.4
	1.5	207.9	570.8	210.3	567.8	211.2	566.9	211.6	560.5
1.5	0.5	123.8	147.2	111.9	130.1	108.0	124.5	106.0	121.7
	1.0	122.8	187.0	119.1	177.6	117.9	174.5	117.3	173.0
	1.5	140.4	271.3	138.8	263.0	138.3	260.3	138.0	258.9
2.0	0.5	155.8	172.0	123.1	131.8	112.2	118.9	106.7	112.6
	1.0	123.3	151.5	112.5	134.4	109.0	128.9	107.2	126.2
	1.5	121.7	171.8	116.1	159.7	114.2	155.7	113.2	153.7

Table 3. Relative efficiencies of the proposed estimator Tp with $\bar{y}$ and $\bar{y}_p$

α = 1.00									
g	β	n = 10		n = 20		n = 30		n = 40	
		E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
0.0	0.5	194.8	601.9	197.6	602.6	198.6	603.0	199.1	630.2
	1.0	459.9	1791.9	482.3	1850.7	490.5	1872.5	494.7	1883.9
	1.5	846.7	3469.4	930.3	3751.9	962.7	3862.2	980.0	3921.1
0.5	0.5	136.1	288.3	135.8	283.7	135.7	282.3	135.7	281.6
	1.0	235.8	740.8	240.1	742.9	241.6	743.8	242.4	744.4
	1.5	395.3	1436.1	410.9	1469.2	416.5	1481.5	419.4	1487.9
1.0	0.5	117.7	175.8	114.5	168.0	113.4	165.5	112.8	164.2
	1.0	150.1	334.6	149.9	328.4	149.8	326.4	149.7	325.4
	1.5	208.3	591.1	210.5	587.5	211.3	586.4	211.7	585.9
1.5	0.5	120.4	145.8	110.6	130.9	107.3	126.1	105.7	123.7
	1.0	122.3	190.9	118.9	181.8	117.8	178.8	117.2	177.4
	1.5	140.3	277.9	138.8	269.9	138.3	266.8	138.0	265.4
2.0	0.5	146.7	163.0	119.4	128.8	110.3	117.8	105.8	112.4
	1.0	121.6	151.1	111.9	135.2	108.6	130.0	107.0	127.5
	1.5	121.1	173.3	115.8	161.6	114.1	157.7	118.2	155.9

Table 4. Relative efficiencies of the proposed estimator Tp with $\bar{y}$ and $\bar{y}_p$

		α = 1.5							
g	β	n = 10		n = 20		n = 30		n = 40	
		E_1	E_2	E_1	E_2	E_1	E_2	E_1	E_2
0.0	0.5	195.7	661.3	198.0	661.0	198.9	661.0	199.3	661.1
	1.0	463.7	1906.8	484.0	1961.6	491.4	1981.7	495.2	1992.2
	1.5	856.0	36563.3	934.9	3917.9	965.3	4024.3	981.3	4080.7
0.5	0.5	136.2	309.3	135.9	304.6	135.8	303.1	135.8	302.3
	1.0	236.7	781.3	240.5	782.4	241.9	782.9	242.6	783.2
	1.5	397.3	1496.8	411.8	1527.6	417.0	1538.9	419.7	1544.9
1.0	0.5	117.0	182.2	114.2	174.9	113.3	172.6	112.8	171.4
	1.0	150.3	348.4	150.0	342.1	149.9	340.0	149.8	339.0
	1.5	208.8	611.8	210.7	607.6	211.4	606.3	211.8	606.7
1.5	0.5	117.9	145.6	109.7	132.4	106.9	128.1	105.5	125.9
	1.0	122.0	195.1	118.8	186.3	117.8	183.4	117.3	181.9
	1.5	140.4	284.7	138.8	276.3	138.3	273.5	138.1	272.1
2.0	0.5	139.7	156.4	116.7	126.8	109.0	117.2	105.1	112.5
	1.0	120.2	151.1	111.3	136.1	108.4	131.3	106.9	128.9
	1.5	120.7	175.1	115.6	163.6	114.0	159.9	113.2	158.0