AN ESTIMATOR USING A PRIORI VALUE OF THE PARAMETER IN SURVEY SAMPLING[1]

BY

M. P. SINGH AND A. S. ROY

Dominion Bureau of Statistics and Indian Statistical Institute

(Received in February, 1970)

1. INTRODUCTION

It is well known that the use of supplementary information in a suitable manner at the stage of sample selection and/or estimation stage generally improves the estimators of the population parameters. The usual techniques in this respect assume that the values of one or more supplementary variables related to the characteristic of interest are known or can be known without much difficulty for each unit of the population. In many cases, however, such detailed *a priori* information may not be available or may be quite costly to collect. On the other hand, some summary information, for instance, an *a priori* value of the parameter θ, quite close to its true value, may be known to the experimenter. For instance, such an information may be available from census, surveys, or even from expert guesses by the specialists in the concerned field. It may also happen that the upper and the lower limits of θ may be known (Dalenius 1965) in which case a simple or modified average (depending on the expected skewness of the distribution of θ) may provide a good approximation to θ.

It seems worthwhile, therefore, to develop an estimator utilising this *a priori* value of θ, so that its mean square error is considerably small than the variance of the usual unbiased estimator. In this paper we propose such an estimator and discuss its practical applicability. The proposed estimator is a weighted average of the *a priori* value and the unbiased estimator of θ obtained from the survey. Since the optimum weight becomes a function of unknown

[1] Originally issued as a Technical Paper of National Sample Survey, Indian Statistical Institute (No. 30/68/16, December, 1968).

parameters, the estimator is modified so as to use approximate optimum weights and it is compared with the usual unbiased estimator in Section 3. Some special cases are also mentioned.

In section 4, the case where an unbiased estimator of θ is not available (such as rates, products, etc.) is considered and the efficiency of the proposed estimator is discussed in brief.

2. The Estimator, its Bias and Mean Square Error

Suppose $\hat{\theta}$ is an unbiased estimator of the parameter θ (say >0 without loss of generality) obtained from a probability sample drawn from a given population and that θ_o is an *a priori* value of the parameter which the statistician believes to be quite close to θ. Let D be the difference between θ and θ_o (*i.e.* $D=\theta-\theta_o$). Then the proposed estimator of θ is a weighted average of $\hat{\theta}$ and θ_o, and is given by

$$(\hat{\theta}_c)=k\hat{\theta}+(1-k)\theta_o, \qquad ...(1)$$

where k is the weight (a constant) the optimum value of which is obtained by minimizing the mean square error of $\hat{\theta}_c$. Obviously $\hat{\theta}_c$ is biased. The bias and variance of $\hat{\theta}_c$ are given by

$$B(\hat{\theta}_c)=E(\hat{\theta}_c)-\theta=(k-1)D \qquad ...(2)$$

and

$$V(\hat{\theta}_c)=E(\hat{\theta}_c{}^2)-E^2(\hat{\theta}_c)=k^2V(\hat{\theta}) \qquad .. (3)$$

respectively, where $V(\hat{\theta})=E(\hat{\theta}^2)-E^2(\hat{\theta})$ denotes the variance of $\hat{\theta}$.

The mean square error (mse) of $\hat{\theta}_c$ thus becomes

$$M(\hat{\theta}_c)=k^2V(\hat{\theta})+(k-1)^2D^2 \qquad ...(4)$$

The optimum value of k which minimizes this mse ean be obtained by differentiating $M(\hat{\theta}_c)$ with respect to k and setting the derivative equal to zero, This gives the optimum weight as

$$k_o=\frac{D^2}{D^2+V(\hat{\theta})}. \qquad ...(5)$$

Suppose $\partial=(\theta-\theta_o)/\theta$ is the relative difference between θ and θ_o, and $e(\theta)=\sqrt{V(\hat{\theta})}/\theta$, denoted by simply $e\,(\geqslant 0)$, is the relative standard error (rse) of $\hat{\theta}$. Then k_o can be expressed as

$$k_o=\frac{\partial^2}{\partial^2+e^2}. \qquad \ldots(6)$$

Now putting this value of k_o, $0 \leqslant k_o \leqslant 1$, the minimum value of $M(\hat{\theta}_c)$ in (4) is seen to be

$$\text{Min } M(\hat{\theta}_c)=\frac{\partial^2}{\partial^2+e^2}\; V(\hat{\theta})=k_o\;V(\hat{\theta}), \qquad \ldots(7)$$

since $V(\hat{\theta})=e^2\theta^2$ and $D^2=\partial^2\theta^2$.

The relative efficiency of $\hat{\theta}_c$ as compared to the usual unbiased estimate $\hat{\theta}$ is given by

$$\text{Eff. } (\hat{\theta}_c)=V(\hat{\theta})/M(\hat{\theta}_c)=\frac{1}{k_o},$$

which exceeds unity provided $e\neq 0$. If $e=0$, then $\hat{\theta}$ and $\hat{\theta}_c$ are identical, which must be the case since $\hat{\theta}$ is the best then.

3. Use of Approximate Optimum Weight

The exact optimum weight k_o cannot be determined since it requires an exact knowledge of the values of e and ∂. Of these two quantities, the value of e may be known in many cases especially when the survey is planned to achieve a prespecified precision. But the exact value of ∂ is always unknown in practice. Hence we can obtain only an approximate value of the optimum weight k_o, using some idea about the magnitude of ∂, and also of e. In this context we discuss below the two possible cases, *viz.*, Case (*i*) : An approximate value ∂_1 is used in place of ∂ ; e is known. Case (*ii*) : Approximate values ∂_1 and e_1 are used in place of both ∂ and e.

Case (*i*) : The proposed estimator in this case is

$$\hat{\theta}_{c1}=\theta_o+k_{o1}(\hat{\theta}-\theta_o) \qquad \ldots(9)$$

where

$$k_{o1} = \frac{\partial_1^2}{\partial_1^2 + e^2} \, . \qquad \ldots(10)$$

The mse of $\hat{\theta}_{c1}$ may be obtained by substituting k_{o1} in (10) for k in the expression for $M(\hat{\theta}_c)$ in (4). After some simplification, we get

$$\frac{M(\hat{\theta}_{c1})}{\theta^2} = e^2 \cdot \frac{\partial_1^2(\partial_1^2 + e'^2)}{(\partial_1^2 + e^2)^2} \qquad \ldots(11)$$

where $e' = e \mid \partial \mid / \mid \partial_1 \mid$.

Now $\hat{\theta}_{c1}$ will be more efficient than $\hat{\theta}$ if $M(\hat{\theta}_{c1})/\theta^2 < e^2$. That is, if

$$\partial_1^2 > \frac{(\partial^2 - e^2)}{2} \text{, provided, of course, } e \neq 0. \ldots(12)$$

Hence if $e \neq 0$ and

$$e^2/\partial^2 = \frac{V(\hat{\theta})}{D^2} \geqslant 1,$$

then for every value of $\mid \partial_1 \mid \neq 0$ estimator $\hat{\theta}_c$ will be more efficient then $\hat{\theta}$. But ∂ is unknown and hence we do not know whether $e^2/\partial^2 \geqslant 1$ in practice. From (12), it is clear that a sufficient condition for $\hat{\theta}_c$' to be more efficient than $\hat{\theta}$ is

$$\mid \partial_1 \mid > \frac{\mid \partial \mid}{\sqrt{2}} \, . \qquad \ldots(13)$$

Thus as long as (13) is satisfied, $\hat{\theta}_{c1}$ is more efficient that $\hat{\theta}$ even if ∂_1 differs from ∂. However too much departure of ∂_1 from ∂ will reduce the gain in efficiency of the estimator. The expression for the efficiency of $\hat{\theta}_{c1}$ relative to $\hat{\theta}$ is

$$\text{Eff. } (\hat{\theta}_{c1}) = \frac{\partial_1^2 + e^2}{\partial_1^2} \cdot \frac{\partial_1^2 + e^2}{\partial_1^2 + e'^2} \qquad \ldots(14)$$

which is greater than or equal to

$(\partial_1^2 + e^2)/\partial_1^2$ if $e' \leqslant e$, *i.e.*, if $\mid \partial_1 \mid \geqslant \mid \partial \mid$.

It follows, therefore, that since (14) tends to (8) as $|\partial_1|$ tends to $|\partial|$, for $|\partial_1|$ exceeding $|\partial|$ but quite close to it $\hat{\theta}_c$, would be at least as good an estimate of θ as $\hat{\theta}_c$. Values of *Eff.* $(\hat{\theta}_{c1})$ have been given in Table 1·1 to 1·3.

Case (*ii*) : In this case the proposed estimator becomes

$$\hat{\theta}_{c2}=\theta_o+k_{02}\,(\hat{\theta}-\theta_o) \qquad \text{...(15)}$$

where $k_{02}=\partial_1^2/(\partial_1^2+e_1^2)$.

As in (11), we get

$$\frac{M(\hat{\theta}_{c2})}{\theta^2}=e^2\,\frac{\partial_1^2\,(\partial_1^2+e_1^2e''^2)}{(\partial_1^2+e_1^2)^2} \qquad \text{...(16)}$$

where $e''=\left(\frac{e_1}{e}\right)\frac{|\partial|}{|\partial_1|}$, e being $\neq 0$.

The corresponding condition for $\hat{\theta}_{c2}$ to be more efficient than $\hat{\theta}$ becomes

$$\partial_1^2>\left(\frac{\partial^2-e^2}{2}\right)\left(\frac{e_1}{e}\right)^2,\ \text{provided, of course, } e\neq 0. \qquad \text{...(17)}$$

The above condition will be satisfied if the condition (12) (or the modified condition (13) holds together with $e_1\leqslant e$. Thus, when anticipated values of $|\partial|$ and e are to be used, it would be safer to take a slightly larger value of $|\partial|$ and smaller value of e for calculation of k_{02}. However, as the difference between the anticipated and the true values increases the efficiency of the proposed estimator decreases.

Special cases : Suppose $\partial=\partial_1=1$. Then from (1), we get the proposed estimator as

$$\hat{\theta}_{co}=k_o\hat{\theta} \qquad \text{...(18)}$$

where $k_o=(1+e^2)^{-1}$, and the mse of $\hat{\theta}_{co}$ is

$$M(\hat{\theta}_{co})=V(\hat{\theta})/(1+e^2). \qquad \text{.. (19)}$$

The relative efficiency of $\hat{\theta}_{co}$ as compared to $\hat{\theta}$ is thus given by

$$\textit{Eff.}\ (\hat{\theta}_{co})=(1+e^2) \qquad \text{...(20)}$$

which is greater than unity.

Now suppose θ is the population mean $\overline{Y}$ and $\hat{\theta}$ is the corresponding sample mean $\bar{y}$ based on a simple random sample of size n selected with replacement, then $\hat{\theta}_{co}$ in (18) is given by

$$\hat{\overline{Y}}_{co}=\frac{n\bar{y}}{n+c_y^2} \qquad ...(21)$$

where c_y is the population coefficient of variation. The estimator $\hat{\overline{Y}}_{co}$ was suggested by Searls (1964). Efficiency of Searls' estimator is given by the entry corresponding to $|\partial|=|\partial_1|=100$ in table 1·1—1·3.

In the present case the estimator in (21) has been arrived at by considering $\partial=\partial_1=1$; it may however be mentioned that Searls suggested this estimator irrespective of the value of ∂ as he did not consider the use of knowledge of ∂. If $\partial\neq\partial_1\neq 1$, then an alternative estimator which utilizes the knowledge of $\overline{Y}_o$ and ∂_1 is given by

$$\hat{\overline{Y}}_{c1}=\overline{Y}_o+\frac{n\partial_1^2}{n\partial_1^2+c_y^2}(\bar{y}-\overline{Y}_o) \qquad ..(22)$$

where $\overline{Y}_o$ is *a priori* value of $\overline{Y}$ and $|\partial_1|$ is used as an anticipated value of $\partial=(1-\overline{Y}_o/\overline{Y})$.

4. Use of a Biased $\hat{\theta}$

So far we have assumed $\hat{\theta}$ to be unbiased estimator of θ. However, in many situations, a simple unbiased estimator of θ may not be available in general. For instance, the parameter θ may be birth-rate, death-rate, per capita consumer expenditure, total crop production etc., where the usual estimator of θ is biased. On the other hand in some other cases a biased estimator of a ratio or regression type may be deliberately used though a simple unbiased estimator exists. In such cases the suggested estimator is

$$\hat{\theta}_c'=\theta_o+k\,(\hat{\theta}-\theta_o). \qquad ...(23)$$

Its bias and mean square error are given by

$$B(\hat{\theta}'_c)=(k-1)D+kB \qquad ...(24)$$

and $$M(\hat{\theta}'_c)=k^2V(\hat{\theta})+(k-1)D^2+k^2B^2+2k(k-1)BD \qquad ...(25)$$

respectively, where $B=E(\hat{\theta})-\theta$, is the bias in $\hat{\theta}$, Differentiating $M(\hat{\theta}'_c)$ in (25) with respect to k, we get optimum k as

$$k_o = \frac{D(D+B)}{V(\hat{\theta})+(D+B)^2} \qquad \text{...(26)}$$

which on substitution in (25) gives the minimum mse as

$$M_o(\hat{\theta}_c') = \frac{V(\hat{\theta})}{V(\hat{\theta})+(D+B)^2}\, D^2. \qquad \text{...(27)}$$

In this case the proposed estimator will be more efficient than $\hat{\theta}$ if $M(\hat{\theta}'_c)$ in (27) is less than $M(\hat{\theta})=V(\hat{\theta})+B^2$. That is if

$$V(\hat{\theta})\left[1-\frac{D^2}{(D+B)^2+V(\hat{\theta})}\right]+B^2>0, \qquad \text{...(28)}$$

which is always true.

Here again the exact value of k_o will not be known in practice. The efficiency of this estimator, using approximate optimum weight may be studied as in section 3. It is believed that with a reasonably good approximation to the optimum weight the proposed estimator will be more efficient than $\hat{\theta}$ as in the unbiased case.

The authors are grateful to the referee for helpful comments.

SUMMARY

In this paper an estimation procedure is suggested which utilizes the knowledge of an *a priori* value of the population parameter θ. The *a priori* value may be available from previous censuses or surveys or even expert guesses. The proposed estimator is given by $\hat{\theta}_c=k\hat{\theta}+(1-k)\theta_0$, where θ_0 is the *a priori* value, k is some constant and $\hat{\theta}$ is the usual unbiased estimator of θ. The optimum value of k which minimizes the mean square error of $\hat{\theta}_c$ is found to be

$$k_0=\partial^2/(\partial^2+e^2) \quad \text{where} \quad |\,\partial\,| =(1-\theta_0/\theta)$$

and e is the relative standard error of $\hat{\theta}$. In many cases, e may be known in practice, especially when the survey is planned to achieve a

specified precision, but $|\partial|$ is always unknown. Hence an approximately optimum $\hat{\theta}_{c1}$ is obtained by using $k_{01}=\partial_1^2/(\partial_1^2+e^2)$ where ∂_1 is an *a priori* value of ∂. $\hat{\theta}_{c1}$ is compared with $\hat{\theta}$ for estimating the true parametric value. A table showing the relative efficiency of $\hat{\theta}_{c1}$ as compared to $\hat{\theta}$ has been given for various values of e, ∂ and ∂_1. The case when approximate values of both ∂ and e are used, have also been discussed. Further some special cases of $\hat{\theta}_c$ have been mentioned. Lastly, the case when $\hat{\theta}$ is biased for θ has been briefly discussed.

REFERENCES

1. Dalenius, T. (1965) : Current trends in the development of Sample Survey Theory and methods.

2. Searls, Donald, T. (1964) : The utilization of a known coefficient of variation in the estimation procedure, Jour. Amer. Stat. Asson , 59.

Appendix

Efficiency of $\hat{\theta}_{c1}$ compared to $\hat{\theta}$

for different values of ∂, ∂_1 and e

(all expressed in percentage)

TABLE (1·1) : e=15%

$\partial_1 \backslash \partial$	100	50	20	15	10	5
(0)	(1)	(2)	(3)	(4)	(5)	(6)
100	102·2	104·0	104·4	104·5	104·5	104·5
50	87·4	109·0	117·1	117·8	118·4	118·7
20	16·2	54·1	156·2	185·4	214·0	235·8
15	8·8	33·0	144·0	200·0	276·9	360·0
10	4·7	18·4	105.6	174·2	325·0	676·0
5	2·8	11·1	69·0	122·0	270·3	1000·0

TABLE (1·2) : e=10%

$\partial_1 \backslash \partial$	100	50	20	15	10	5
(0)	(1)	(2)	(3)	(4)	(5)	(6)
100	101·0	101·8	102·0	102·0	102·0	102 0
50	93·2	104·0	107·5	107·8	108·0	108·1
20	21·6	61·0	125·0	137·0	147·0	153·8
15	10·0	35·2	116 6	144·4	174·2	198·8
10	4·0	15·4	80 0	123·1	200·0	320·0
5	1·6	6·2	38·5	67·6	147·0	500·0

TABLE (1·3) : e=5%

$\partial_1 \backslash \partial$	100	50	20	15	10	5
(0)	(1)	(2)	(3)	(4)	(5)	(6)
100	100·2	100·4	100·5	100·5	100·5	100·5
50	98·1	101·0	101·8	101·9	102·0	102·0
20	44·1	81·2	106·2	109·7	112·2	112·5
15	20·8	55·1	103·7	111·1	117·6	121·9
10	6·0	21·6	78·1	100·0	125·0	147·0
5	1·0	4·0	23·5	40·0	80·0	200·0