Jour. Ind. Soc. Ag. Statistics
47(3), 1995 : 243–248

On the Alternative Estimator for Randomised Response Technique

S. Sampath, N. Uthayakumaran* and D. S. Tracy¤
Loyola College, Madras 34
(Received : November, 1993)

SUMMARY

In this paper Warner's [4] estimator is shown to be the best estimator when one tries to develop estimators better than the conventional estimator proceeding in the direction of Singh and Singh [3].

Key words : Mean square errors, gain in efficiency, generalised estimator.

Introduction

Warner [4] suggested a randomised response technique to estimate the proportion π of individuals belonging to a sensitive category using a simple random sample (with replacement) of size n. For the sake of brevity the details are omitted. The suggested estimator is

$$\hat{\pi}_w = \frac{\hat{\theta} - \bar{P}}{2P - 1}, \bar{P} = 1 - P \qquad (1.1)$$

where $\hat{\theta}$ and P are the proportion of yes answers in the sample and the probability of getting the question "Do you belong to the sensitive category A ?" respectively.

We know that $E(\hat{\theta}) = \theta = \pi p + (1 - \pi)(1 - P)$

and $$V(\hat{\theta}) = \frac{\theta(1 - \theta)}{n}$$

Therefore, $E(\hat{\pi}_w) = \pi$

and $$V(\hat{\pi}_w) = \frac{\theta(1 - \theta)}{n(2p - 1)^2}$$

Recently Singh and Singh [3] suggested a generalised estimator

* ICMR, Madras 54
¤ University Of Windsor, Canada

$$\hat{\pi}_{rs} = \frac{\lambda \hat{\theta} - \bar{P}}{(2p-1)} \tag{1.2}$$

and proved that for the optimum value of λ the mean square error of the above estimator is smaller than the variance of the conventional estimator given in (1.1). In the following sections some other classes are suggested and comparisons made.

2. *New Class of Estimators*

Searls [2] modified the conventional estimator $\bar{y}$ for the population mean by defining the estimator $\lambda\bar{y}$ and proved that for the optimum value of λ the suggested estimator is more efficient than $\bar{y}$. Motivated by this, we suggest the estimator

$$\hat{\pi}_{ss} = \frac{\lambda(\hat{\theta} - \bar{P})}{(2P-1)} \tag{2.1}$$

as an estimator of π.

The mean square error of the estimator given in (2.1) is

$$M(\hat{\pi}_{ss}) = \frac{\frac{\lambda^2\theta(1-\theta)}{n} + (\lambda-1)^2(\theta-\bar{p})^2}{(2p-1)^2} \tag{2.2}$$

It can be seen that the mean square error given above is minimum if

$$\lambda = \frac{(\theta-\bar{p})^2}{(\theta-\bar{p})^2 + \frac{\theta(1-\theta)}{n}} \tag{2.3}$$

which essentially lies between zero and one. The minimum mean square error of the estimator given in (2.1) is

$$\frac{(\theta-\bar{p})^4[n+\theta(1-\theta)]\theta(1-\theta)}{(2p-1)^2[n(\theta-\bar{p})^2+\theta(1-\theta)]^2} \tag{2.4}$$

Now we shall prove that the above mean square error is less than the variance of Warner [4] estimator.

Consider the difference

$$V(\hat{\pi}_w) - M(\hat{\pi}_{ss}) = \frac{\theta(1-\theta)}{n(2p-1)^2} - \frac{\lambda^2\theta(1-\theta)}{n(2p-1)^2} - \frac{(\lambda-1)^2(\theta-\bar{p})}{(2p-1)^2}$$

$$= \frac{\theta(1-\theta) - \lambda^2\theta(1-\theta) - n(\lambda-1)^2(\theta-\bar{p})^2}{n(2p-1)^2}$$

$$= \frac{\theta(1-\theta)(1-\lambda^2) - n(1-\lambda)^2(\theta-\bar{p})^2}{n(2p-1)^2}$$

$$= \frac{(1-\lambda)\{\theta(1-\theta)(1+\lambda) - n(1-\lambda)(\theta-\bar{p})^2\}}{n(2p-1)^2}$$

$$= \frac{(1-\lambda)}{n(2p-1)^2} \; \frac{\dfrac{\theta^2(1-\theta)^2}{n} + \theta(1-\theta)(\theta-\bar{p})^2}{\dfrac{\theta(1-\theta)}{n} + (\theta-\bar{p})^2}$$

Clearly the right hand side of the expression is non-negative, because $0 < \lambda < 1$. Hence we conclude that the mean square error given in (2.4) is less than the variance of Warner [4] estimator.

In order to assess the amount of gain in efficiency due to $\hat{\pi}_{ss}$ over $\hat{\pi}_w$ the percentage efficiency of $\hat{\pi}_{ss}$ over $\hat{\pi}_w$ has been calculated for different choices of π, P and n. Their values are listed in Table 2.1. It may be noted that when n = 5, 10 and $\pi = 0{\cdot}2$, the mean square errors of $\hat{\pi}_{ss}$ are very small if p = 0·6. Therefore, in these cases the percentage efficiencies are quite large. The Table 2.1 clearly establishes the superiority of $\hat{\pi}_{ss}$ over $\hat{\pi}_{rs}$ for optimum values of λ.

It is pertinent to note that the optimum value of λ given in (2.3) depends on θ which can be estimated unbiasedly by the sample proportion of yes answers. Hence it is suggested for the use of $\hat{\lambda}$ which is derived from λ on replacing θ by its unbiased estimator. For related results one can refer to Sampath [1] and Singh and Singh [3].

3. *Further Improvement*

Motivated by the form of the estimator $\hat{\pi}_w$ we suggest a more generalised class of estimators, namely

$$\hat{\pi}_{ab} = a\,\hat{\theta} + b \tag{3.1}$$

which reduces to $\hat{\pi}_w$ when $a = (2p-1)^{-1}, b = -\bar{p}(2p-1)^{-1}$,

$\hat{\pi}_{rs}$ when $a = \lambda(2p-1)^{-1}, b = -\bar{p}(2p-1)^{-1}$ and

Table 2.1. Efficiency of $\hat{\pi}_{ss}$ over $\hat{\pi}_{rs}$

$\pi = .2$

P	Sample size									
	5		10		50		100		500	
	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS
0.6	125.45	3180.00	112.73	1640.00	102.55	408.00	101.27	254.00	100.25	130.80
0.7	132.63	836.25	116.32	468.13	103.26	173.63	101.63	136.81	100.33	107.36
0.8	142.50	402.22	121.25	251.11	104.25	130.22	102.13	115.11	100.43	103.02
0.9	156.92	250.31	128.46	175.16	105.69	115.03	102.85	107.52	100.57	101.50

$\pi = .4$

P	Sample size									
	5		10		50		100		500	
	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS
0.6	121.67	880.00	110.83	490.00	102.17	178.00	101.08	139.00	100.22	107.80
0.7	123.48	294.06	111.74	197.03	102.35	119.41	101.17	109.70	100.23	101.94
0.8	125.45	185.56	112.73	142.78	102.55	108.56	101.27	104.28	100.25	100.86
0.9	127.62	147.58	113.81	123.79	102.76	104.76	101.38	102.38	100.28	100.48

Table 2.1. Contd.

$\pi = .6$										
P	Sample size									
	5		10		50		100		500	
	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS
0.6	118.46	446.67	109.23	273.33	101.85	134.67	100.92	117.33	100.18	103.47
0.7	117.04	186.25	108.52	143.13	101.70	108.63	100.85	104.31	100.17	100.86
0.8	115.71	138.02	107.86	119.01	101.57	103.80	100.79	101.90	100.16	100.38
0.9	114.48	121.15	107.24	110.57	101.45	102.11	100.72	101.06	100.14	100.21
$\pi = .8$										
P	Sample size									
	5		10		50		100		500	
	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS	ERS	ESS
0.6	115.71	292.50	107.86	196.25	101.57	119.25	100.79	109.62	100.16	101.93
0.7	112.26	146.02	106.13	123.01	101.23	104.60	100.61	102.30	100.12	100.46
0.8	109.41	118.89	104.71	109.44	100.94	101.89	100.47	100.94	100.09	100.19
0.9	107.03	109.39	103.51	104.70	100.70	100.94	100.35	100.47	100.07	100.09

ERS – Efficiency of the estimator suggested in Singh and Singh (1992)
ESS – Efficiency of the new estimator suggested.

$$\hat{\pi}_{ss} \text{ when } a = \lambda(2p-1)^{-1}, b = -\lambda\bar{p}(2p-1)^{-1}$$

The mean square error of the estimator given in (3.1) is

$$M(\hat{\pi}_{ab}) = a^2 \frac{\theta(1-\theta)}{n} + (a\theta + b)^2 + \frac{(\theta - \bar{p})^2}{(2p-1)^2} - \frac{2(a\theta + b)(\theta - \bar{p})}{(2p-1)} \quad (3.2)$$

The mean square error of the estimator given in (3.2) is minimum if $a = 0$ and $b = (\theta - \bar{p}) / (2p - 1)$ and the resulting mean square error is zero.

4. *Concluding Remarks*

The exact mean square error of the estimator $\hat{\pi}_{ab}$ is zero for the optimum values of a and b given in Section 3. But the optimum value of b requires the knowledge of θ for which the natural choice is $\hat{\theta}$, the proportion of yes answers in the sample. It may be noted that for such a choice the estimator $\hat{\pi}_{ab}$ reduces to $\hat{\pi}_w$ of Warner [4].

Hence it is concluded that the estimator $\hat{\pi}_w$ is the best estimator when one tries to develop estimators better than the conventional estimator proceeding in the direction of Singh and Singh [3].

ACKNOWLEDGEMENT

The authors are grateful to the referee for his comments.

REFERENCES

[1] Sampath, S., 1989. On the optimal choice of unknowns in ratio type estimators, *Jour. Ind. Soc. Ag. Stat.*, **41**, 166-172.

[2] Searls, D. T., 1965. The utilisation of a known coefficient of variation in the estimation procedure., *Jour. Amer. Stat. Assoc.*, **59**, 1225-1226.

[3] Singh, S., and Singh, R., 1992. An alternative estimator for randomised response technique, *Jour. Ind. Soc. Ag. Stat.*, **44**, 149-154.

[4] Warner, S. L., 1965. A survey technique for eliminating evasive answer bias, *Jour. Amer. Stat. Assoc.*, **60**, 63-69.