Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 7, Issue 2 (2014): 105-118

AN ALTERNATIVE ESTIMATOR IN STRATIFIED RR STRATEGIES

Housila P. Singh[1] and Tanveer A. Tarray[2]
School of Studies in Statistics, Vikram University, Ujjain - 456010 - India.
E Mail: [1]hpsujn@gmail.com; [2]tanveerstat@gmail.com

Received May 23, 2014
Modified November 15, 2014
Accepted November 29, 2014

Abstract

This paper addresses the problem of estimating the proportion π_S of the population having some sensitive characteristics using stratified randomize response model based on Warner's model. We have suggested a class of estimators for the population proportion π_S using Searls (1965) technique. It is shown that under certain conditions the proposed class of estimators is more efficient than Hong et al. (1994) and Kim and Warde (2004) estimators. The optimum estimator in the class is investigated. It has been shown that the optimum estimator is more efficient than Hong et al. (1994) and Kim and Warde (2004) estimators. Since the optimum estimator involves the use of an unknown population parameter π_S it has therefore little practical utility. Using an estimated value of the parameter π_S in the optimum estimator, an alternative estimator has been investigated for use in practice.

Key Words: Randomized Response Technique. Stratified Random Sampling, Proportional Allocation, Optimum Allocation.

1. Introduction

In psychological surveys, a social desirability bias has been observed as a major cause of distortion in standardized personality measures. Survey researchers have similar concerns about the truth of survey results/ findings about such topics as drunk driving, use of marijuana, tax evasion, illicit drug use, induced abortion, shop lifting, child abuse, family disturbances, cheating in exams, HIV/AIDS, and sexual behavior. The most serious problem in studying certain social problems that are sensitive in nature (e.g. induced abortion, drug usage, tax evasion, etc.) is lack of reliable measure of their incidence or prevalence. Thus to obtain trustworthy data on such confidential matters, especially the sensitive ones, instead of open surveys alternative procedures are required. Such an alternative procedure known as "randomized response (RR) technique" was first introduced by Warner (1965). It provides the opportunity of reducing response biases due to dishonest answers to sensitive questions. As a result, the technique assures a considerable degree of privacy protection in many contexts. Warner (1965) himself pointed out how one may get a biased estimate in an open survey when a population consists of individuals bearing a stigmatizing character A or its complement A^c, which may or may not also be stigmatizing. Later several authors including Mangat and Singh (1990), Mangat (1994), Singh and Mangat (1996), Singh and Tarray (2012, 2014 a,b,c,d,e,f,g) etc. have modified and suggested alternative response procedures applicable to different situations.

Hong et al. (1994) envisaged RR technique that applied the same randomization device to every stratum. Stratified random sampling is generally obtained by dividing the population into non – overlapping groups called strata and selecting a simple random sample from each stratum. An RR technique using a stratified sampling gives the group characteristics related to each stratum estimator. Also, stratified samples protect a researcher from the possibility of obtaining a poor sample. Under Hong et al.'s (1994) proportional sampling assumption, it may cause a high cast because of the difficulty in obtaining a proportional sample from some stratum. To overcome this problem, Kim and Warde (2004) presented a stratified randomized response technique using an optimal allocation which is more efficient than a stratified randomized response technique using a proportional allocation.

2. Proposed model

Let the population be partitioned into strata, and a sample is selected by simple random sampling with replacement (SRSWR) in each stratum. To get the full benefit from stratification, we assume that the number of units in each stratum is known. An individual respondent in the sample of stratum 'i' is instructed to use the randomization device R_i which consists of a sensitive question (S) card with probability P_i and its negative question (S^c) with probability ($1-P_i$). The respondent should answer the question by "Yes" or "No" without reporting which question card she or he has. A respondent belonging to the sample in different strata will perform different randomization devices, each having different pre-assigned probabilities. Let n_i denote the number of units in the sample from stratum i and n denote the total number of units in samples from all stratum so that $n = \sum_{i=1}^{k} n_i$. Under the assumption that these "Yes" or "No" reports are made truthfully and P_i ($\neq 0.5$) is set by the researcher, the probability of a "Yes" answer in a stratum i for this procedure is

$$Z_i = \left[P_i \pi_{Si} + (1-P_i)(1-\pi_{Si})\right] \text{ for (i =1, 2 ..., k)}, \tag{2.1}$$

where Z_i is the proportion of "Yes" answers in a stratum i , π_{Si} is the proportion of respondents with the sensitive trait in a stratum i and P_i is the probability that a respondent in the sample stratum i has a sensitive question (S) card.

The maximum likelihood estimate of π_{Si} is shown to be

$$\hat{\pi}_{Si} = \frac{\hat{Z}_i - (1-P_i)}{2P_i - 1} \text{ , for (i =1,2,…,k)}, \tag{2.2}$$

where $\hat{Z}_i$ is the proportion of "Yes" answer in a sample in the stratum i since each $\hat{Z}_i$ is a binomial distribution $B(n_i, Z_i)$ and the selections in different strata are made independently, the maximum likelihood estimate of $\pi_S = \sum_{i=1}^{k} w_i \; \pi_{Si}$ is easily shown to be $\hat{\pi}_S = \sum_{i=1}^{k} w_i \hat{\pi}_{Si} = \sum_{i=1}^{k} w_i \left[\frac{\hat{Z}_i - (1-P_i)}{(2P_i - 1)} \right]$, (2.3)

where N being the number of units in the whole population, N_i to be the total number of units in the stratum i and $w_i = (N_i/N)$ for (i = 1, 2, ...k) so that $w = \sum_{i=1}^{k} w_i = 1$

As each estimator $\hat{\pi}_{Si}$ is unbiased for π_{Si} , the expected value for $\hat{\pi}_S$ is

$$E\left(\hat{\pi}_S\right) = E\left(\sum_{i=1}^{k} w_i \; \hat{\pi}_{Si}\right) = \sum_{i=1}^{k} w_i \; \pi_{Si} = \pi_S \tag{2.4}$$

Since each unbiased estimator $\hat{\pi}_{Si}$ has its own variance, the variance of $\hat{\pi}_S$ is

$$\mathrm{Var}(\hat{\pi}_S) = \sum_{i=1}^{k} \frac{w_i^2}{n_i} \left[\pi_{Si}(1-\pi_{Si}) + \frac{P_i(1-P_i)}{(2P_i - 1)^2} \right] \tag{2.5}$$

Under the proportional allocation (i.e. $n_i = n\,(N_i/N)$ the variance of $\hat{\pi}_S$ is given by

$$\mathrm{Var}(\hat{\pi}_S)_P = \frac{1}{n} \sum_{i=1}^{k} w_i \left(\pi_{Si}(1-\pi_{Si}) + \frac{P_i(1-P_i)}{(2P_i - 1)^2} \right) \tag{2.6}$$

Further, we assume $P_i = P$ for all i ,(4.2.6) becomes the following

$$\mathrm{Var}(\hat{\pi}_S)_P = \frac{1}{n} \sum_{i=1}^{k} w_i L_i^2 \tag{2.7}$$

where

$$L_i = \sqrt{\left\{ \pi_{Si}(1-\pi_{Si}) + \frac{P(1-P)}{(2P-1)^2} \right\}} \tag{2.8}$$

The expression (2.7) is due to Hong et al. (1994).

Information on π_{Si} is usually unavailable. But if prior information on π_{Si} is available from past experience then it helps to derive the optimal allocation formula. The optimal allocation of n to $n_1, n_2, \ldots, n_{k-1}$ and n_k to drive the minimum variance of the $\hat{\pi}_S$ subject to $n = \sum_{i=1}^{k} n_i$ is approximately given by $\frac{n_i}{n} = \frac{w_i V_i}{\sum_{i=1}^{k} w_i V_i}$ (2.9)

where

$$V_i = \sqrt{\left\{\pi_{Si}(1-\pi_{Si}) + \frac{P_i(1-P_i)}{(2P_i-1)^2}\right\}}$$

Using (2.9) in (2.5), we get the minimal variance of the estimator $\hat{\pi}_S$ as

$$Var(\hat{\pi}_S)_O = \frac{1}{n}\left(\sum_{i=1}^{k} w_i V_i\right)^2 \tag{2.10}$$

which is due to Kim and Warde (2004). Further , under the assumption $P_i = P$ for all i , (2.10) becomes

$$Var(\hat{\pi}_S)_O = \frac{1}{n}\left(\sum_{i=1}^{k} w_i L_i\right)^2 \tag{2.11}$$

which is due to Kim and Warde (2004) .

In this paper , we have suggested a class of estimators for estimating the population proportion π_S .We have shown that the optimum estimator of the class is more efficient than Hong et al.(1994) and Kim and Warde (2004) estimators. An alternative estimator based on estimated optimum values has been derived along with its properties.

3. SUGGESTED ESTIMATORS

Motivated by Searls (1965), we have suggested a class of estimators of π_S as

$$\hat{\pi}_{S1} = \lambda\, \hat{\pi}_S \tag{3.1}$$

Where λ is a suitably chosen constant. The bias of $\hat{\pi}_{S1}$ is given by

$$B(\hat{\pi}_{S1}) = E(\hat{\pi}_{S1} - \pi_S) = (\lambda - 1)\pi_S \tag{3.2}$$

The mean square error (MSE) of $\hat{\pi}_{S1}$ is given by

$$MSE(\hat{\pi}_{S1}) = E(\hat{\pi}_{S1} - \pi_S)^2 = \lambda^2\left(\pi_S{}^2 + Var(\hat{\pi}_S)\right) - 2\lambda\pi_S{}^2 + \pi_S{}^2 \tag{3.3}$$

Minimization of (3.3) with respect to λ yields the optimum value of λ as

$$\lambda = \frac{\pi_S{}^2}{(\pi_S{}^2 + Var(\hat{\pi}_S))} = \lambda_0 \quad \text{say} \tag{3.4}$$

So, the value of λ_0 of λ at (3.4) is the optimum value of λ which will minimize $MSE(\hat{\pi}_{S1})$ at (3.3), Putting the value of λ_0 in place of λ in (3.3), we get the minimum MSE of $(\hat{\pi}_{S1})$ as

$$\text{min. } MSE(\hat{\pi}_{S1}) = \frac{\pi_S{}^2\, Var(\hat{\pi}_S)}{\pi_S{}^2 + Var(\hat{\pi}_S)} \tag{3.5}$$

From (2.5) and (3.5) we have

$$Var(\hat{\pi}_S) - \text{min.}MSE(\hat{\pi}_{S1}) = \frac{(Var(\hat{\pi}_S))^2}{\pi_S{}^2 + Var(\hat{\pi}_S)}$$

which is always positive. Thus the proposed estimator $\hat{\pi}_{S1}$ at its optimum condition is always better than usual unbiased estimator $\hat{\pi}_S$

Proportional Allocation

Under the proportional allocation (i.e. $n_i = n\,(N_i/N)$), the MSE of $\hat{\pi}_{S1}$ in (3.3) is given by

$$MSE\left(\hat{\pi}_{S1}\right)_P = \lambda^2\left[\pi_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i V_i\right] - 2\lambda\pi_S^{\,2} + \pi_S^{\,2} \tag{3.6}$$

where V_i is same defined earlier.
For $P_i = P$ for all i , (3.6) reduces to

$$MSE\left(\hat{\pi}_{S1}\right)_P = \lambda^2\left[\pi_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i L_i^{\,2}\right] - 2\lambda\pi_S^{\,2} + \pi_S^{\,2} \tag{3.7}$$

which is minimum when

$$\lambda = \frac{\pi_S^{\,2}}{\left\{\pi_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i L_i^{\,2}\right\}} = \lambda_{0P} \tag{3.8}$$

where L_i is same as defined earlier.
Substitution of (3.8) in (3.7) yields the minimum MSE of $\hat{\pi}_{S1}$ under the proportional allocation and $P_i = P$ for all i, as

$$\text{min. } MSE\left(\hat{\pi}_{S1}\right)_P = \frac{\left\{\frac{1}{n}\left\{\sum_{i=1}^{k} w_i L_i^{\,2}\right\}\right\}\pi_S^{\,2}}{\left\{\pi_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i L_i^{\,2}\right\}} \tag{3.9}$$

From (2.7) and (3.9) we have

$$Var\,(\hat{\pi}_S)_P - \text{min.}MSE(\hat{\pi}_{S1})_P = \frac{\left\{\frac{1}{n}\left\{\sum_{i=1}^{k} w_i L_i^{\,2}\right\}\right\}^2}{\left\{\pi_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i L_i^{\,2}\right\}} > 0 \tag{3.10}$$

which shows that the proposed class of estimator $\hat{\pi}_{S1}$ is more efficient than Hong et al.'s (1994) estimator $\hat{\pi}_S$ under proportional allocation and $P_i = P$ for all i .

Optimum allocation

Under optimum allocation (2.9), the MSE of $\hat{\pi}_{S1}$ is given by

$$\text{MSE}\left(\hat{\pi}_{S1}\right)_O = \lambda^2\left[\pi_S{}^2 + \frac{1}{n}\left(\sum_{i=1}^{k} w_i V_i\right)^2\right] - 2\lambda\pi_S{}^2 + \pi_S{}^2 \tag{3.11}$$

where V_i is same as defined earlier.

The MSE $(\hat{\pi}_{S1})_O$ in (3.11) is minimized for

$$\lambda = \frac{\pi_S{}^2}{\left\{\pi_S{}^2 + \frac{1}{n}\left(\sum_{i=1}^{k} w_i V_i\right)^2\right\}} = \lambda_{0O} \quad \text{(say)} \tag{3.12}$$

Thus the resulting minimum MSE of $\hat{\pi}_{S1}$ under optimum allocation (2.9) is given by

$$\min.\text{MSE}(\hat{\pi}_{S1})_O = \frac{\left\{\left(\sum_{i=1}^{k} w_i V_i\right)^2\right\}\left(\frac{\pi_S{}^2}{n}\right)}{\left\{\pi_S{}^2 + \frac{1}{n}\left\{\sum_{i=1}^{k} w_i V_i\right\}^2\right\}} \tag{3.13}$$

For $P_i = P$ for all i, (3.11) reduces to

$$\text{MSE}\left(\hat{\pi}_{S1}\right)_O = \lambda^2\left[\pi_S{}^2 + \frac{1}{n}\left(\sum_{i=1}^{k} w_i L_i\right)^2\right] - 2\lambda\pi_S{}^2 + \pi_S{}^2 \tag{3.14}$$

where L_i is same as defined earlier.

The MSE $(\hat{\pi}_{S1})_O$ in (3.14) is minimized for

$$\lambda = \pi_S{}^2\left\{\pi_S{}^2 + \frac{1}{n}\left(\sum_{i=1}^{k} w_i L_i\right)^2\right\}^{-1} = \lambda^*{}_{0O} \quad \text{(say)} \tag{3.15}$$

Thus the resulting minimum MSE of $\hat{\pi}_{S1}$ under optimum allocation (2.9) and $P_i = P$ for all i, is given by

$$\min.\text{MSE}(\hat{\pi}_{S1})_O = \frac{\left\{\left(\sum_{i=1}^{k} w_i L_i\right)^2\right\}\left(\frac{\pi_S{}^2}{n}\right)}{\left\{\pi_S{}^2 + \frac{1}{n}\left\{\sum_{i=1}^{k} w_i L_i\right\}^2\right\}} \tag{3.16}$$

From (2.11) and (3.16) we have

$$\text{Var}(\hat{\pi}_S)_O - \min.\text{MSE}(\hat{\pi}_{S1})_O = \frac{\left\{\frac{1}{n^2}\left\{\sum_{i=1}^{k} w_i L_i\right\}^4\right\}}{\left\{\pi_S{}^2 + \frac{1}{n}\left\{\sum_{i=1}^{k} w_i L_i\right\}^2\right\}} > 0 \tag{3.17}$$

which indicates that the proposed estimator $\hat{\pi}_{S1}$ is more efficient than Kim and Warde (2004) estimator $\hat{\pi}_S$ under the optimum allocation and $P_i = P$ for all i .

4. Numerical Illustration

In order to assess the amount of gain in efficiency due to $\hat{\pi}_{S1}$ under proportional and optimum allocation over Hong et al. (1994) estimator $(\hat{\pi}_S)_P$ (under proportional allocation) and Kim and Warde (2004) estimator $(\hat{\pi}_S)_O$ (under optimum allocation) ,the percentage relative efficiency of $\hat{\pi}_{S1}$ (under proportional and optimum allocation) over $\hat{\pi}_S$ (under proportional and optimum allocation) have been computed for different values π_{S1} , π_{S2} , w_1 , w_2 , π_S , $P = P_1$ and P_2 using the following formulae:

(i) The percent relative efficiency of the estimator $\hat{\pi}_S$ (under proportional allocation) with respect to the estimator $\hat{\pi}_S$ (under optimum allocation) is defined by:

$$PRE((\hat{\pi}_S)_O,(\hat{\pi}_S)_P) = \frac{Var(\hat{\pi}_S)_P}{Var(\hat{\pi}_S)_O} \times 100 \qquad \text{(for k = 2)} \qquad (4.1)$$

(ii) The percent relative efficiency of the estimator $\hat{\pi}_{S1}$ (under optimum allocation) with respect to Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum allocation) (from (2.10) and (3.13)) is defined by:

$$PRE\ (\hat{\pi}_{S1},\hat{\pi}_S)_O = \frac{Var(\hat{\pi}_S)_O}{\min.MSE(\hat{\pi}_{S1})_O} \times 100$$

$$= \left(1 + \frac{1}{n\,\pi_S^2}\left(\sum_{i=1}^{2} w_i V_i\right)^2\right) \times 100 \quad \text{(for k = 2)} \qquad (4.2)$$

(iii) The percent relative efficiency of the estimator $\hat{\pi}_{S1}$ (under proportional allocation) with respect to Hong et al.'s (1994) estimator $\hat{\pi}_S$ (under proportional allocation) (from (2.7) and (3.9)) is defined by:

$$PRE\ (\hat{\pi}_{S1},\hat{\pi}_S)_P = \frac{Var(\hat{\pi}_S)_P}{\min.MSE(\hat{\pi}_{S1})_P} \times 100$$

$$= \left(1 + \frac{1}{n\,\pi_S^2}\sum_{i=1}^{2} w_i L_i^2\right) \times 100 \qquad \text{(for k = 2)} \qquad (4.3)$$

when n = 1000												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	140.2	160.2	127.2	174.0	123.5	134.4	107.5	112.5
0.08	0.13	0.3	0.7	0.115	245.9	383.1	184.3	283.8	168.8	213.1	118	131.5
0.28	0.33	0.7	0.3	0.295	138.8	157.1	124.2	140.1	181.1	125.3	104.8	107.8
0.28	0.33	0.3	0.7	0.315	238.6	357.7	172.8	244.7	150.5	176.5	111.6	119.5
0.48	0.53	0.7	0.3	0.495	138.4	156.3	123.5	138.4	116.9	123.5	104.3	107
0.48	0.53	0.3	0.7	0.515	236.7	351.7	170.2	236.9	147	170.3	110.6	117.6
0.68	0.73	0.7	0.3	0.695	139	157.5	124.5	140.8	118.5	126.1	105	108.1
0.68	0.73	0.3	0.7	0.715	2399.9	362.2	174.6	250.5	152.9	181.1	112.3	120.8
0.88	0.93	0.7	0.3	0.895	140.6	161.3	128.1	149.5	125.3	137.8	108.4	114.1
0.88	0.93	0.3	0.7	0.915	248.8	394.6	189.3	305.2	178.9	237.4	121.9	139.6

Table 4.1: Percent relative efficiency of the estimator $\hat{\pi}_S$ (under proportional allocation) with respect to the estimator $\hat{\pi}_S$ (under optimum allocation)

when n = 6												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	5603.59	5465.08	5465.08	5416.61	5416.61	5403.53	5408.21	5403.53
0.08	0.13	0.3	0.7	0.115	1606.44	1091.84	1091.84	911.73	911.73	863.13	880.55	863.13
0.28	0.33	0.7	0.3	0.295	684.07	669.70	669.70	664.68	664.68	663.32	663.81	663.32
0.28	0.33	0.3	0.7	0.315	311.17	242.58	242.58	218.14	218.14	214.58	214.14	212.58
0.48	0.53	0.7	0.3	0.495	309.14	304.04	304.04	302.25	302.25	301.77	301.94	301.77
0.48	0.53	0.3	0.7	0.515	180.09	154.43	154.43	145.62	145.62	143.89	143.02	143.89
0.68	0.73	0.7	0.3	0.695	205.41	202.83	202.83	201.92	201.92	201.68	201.78	201.68
0.68	0.73	0.3	0.7	0.715	140.66	127.35	127.35	122.69	122.69	121.43	121.88	121.43
0.88	0.93	0.7	0.3	0.895	162.23	160.67	160.67	160.12	160.12	160.98	159.03	160.98
0.88	0.93	0.3	0.7	0.915	123.40	115.27	115.27	112.42	112.42	111.66	111.93	111.66

when n = 100												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	443.97	435.31	435.31	432.28	432.28	431.47	431.76	431.47
0.08	0.13	0.3	0.7	0.115	194.15	161.99	161.93	150.73	150.73	147.69	148.78	147.69
0.28	0.33	0.7	0.3	0.295	136.50	135.60	135.60	135.29	135.29	135.20	135.23	135.20
0.28	0.33	0.3	0.7	0.315	113.19	108.91	108.91	107.41	107.41	107.06	107.15	107.06
0.48	0.53	0.7	0.3	0.495	113.07	112.75	112.75	112.64	112.64	112.61	112.62	112.61
0.48	0.53	0.3	0.7	0.515	105.00	103.40	103.40	102.84	102.84	102.68	102.74	102.68
0.68	0.73	0.7	0.3	0.695	106.58	106.42	106.42	106.37	106.37	106.35	106.36	106.35
0.68	0.73	0.3	0.7	0.715	102.54	101.71	101.71	101.41	101.41	101.34	101.36	101.34
0.88	0.93	0.7	0.3	0.895	103.89	103.79	103.79	103.75	103.75	103.74	103.74	103.74
0.88	0.93	0.3	0.7	0.915	101.46	100.95	100.95	100.77	100.77	100.72	100.72	100.72

when n = 1000												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	134.39	133.53	133.53	133.22	133.22	133.14	133.17	133.14
0.08	0.13	0.3	0.7	0.115	109.41	106.19	106.19	105.07	105.07	104.77	104.87	104.77
0.28	0.33	0.7	0.3	0.295	103.65	103.56	103.56	103.52	103.52	103.52	103.53	103.52
0.28	0.33	0.3	0.7	0.315	101.32	100.89	100.89	100.74	100.74	100.70	100.71	100.70
0.48	0.53	0.7	0.3	0.495	101.70	101.70	101.71	101.71	101.27	101.26	101.27	101.26
0.48	0.53	0.3	0.7	0.515	100.50	100.34	100.34	100.28	100.28	100.26	100.27	100.26
0.68	0.73	0.7	0.3	0.695	100.65	100.64	100.61	100.63	100.63	100.63	100.64	100.63
0.68	0.73	0.3	0.7	0.715	100.25	100.17	100.17	100.14	100.14	100.13	100.14	100.13
0.88	0.93	0.7	0.3	0.895	100.38	100.37	100.37	100.38	1000.38	100.37	100.37	100.37
0.88	0.93	0.3	0.7	0.915	100.14	100.09	100.09	100.07	100.07	100.07	100.07	100.07

Table 4.2: Percent relative efficiency of the estimator $\hat{\pi}_{S1}$ (under optimum allocation) with respect to the Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum allocation)

when n = 6												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	11338.13	10988.61	2678.36	2678.36	1139.42	1139.42	600.79	600.79
0.08	0.13	0.3	0.7	0.115	7789.03	7789.03	1881.66	1881.66	787.70	787.70	404.82	404.82
0.28	0.33	0.7	0.3	0.295	1288.82	1288.82	391.09	391.09	224.84	224.82	166.66	166.66
0.28	0.33	0.3	0.7	0.315	1143.96	1143.96	356.61	356.61	210.80	210.80	129.77	159.77
0.48	0.53	0.7	0.3	0.495	525.08	525.08	206.24	206.24	147.19	147.19	126.53	126.53
0.48	0.53	0.3	0.7	0.515	492.70	492.70	198.13	198.13	143.59	143.59	124.29	124.49
0.68	0.73	0.7	0.3	0.695	314.32	314.32	152.58	152.58	122.63	122.63	112.14	112.14
0.68	0.73	0.3	0.7	0.715	302.234	302.234	149.41	149.41	121.11	121.11	111.21	111.21
0.88	0.93	0.7	0.3	0.895	226.78	226.78	129.25	129.25	111.19	111.19	104.87	104.81
0.88	0.93	0.3	0.7	0.915	220.98	220.98	127.66	127.66	110.38	110.38	104.33	104.23

when n = 100												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	774.28	774.28	254.89	254.89	158.71	158.71	125.04	125.04
0.08	0.13	0.3	0.7	0.115	561.34	561.34	206.89	206.89	141.26	141.26	118.28	118.28
0.28	0.33	0.7	0.3	0.295	171.32	171.32	117.46	117.46	107.49	107.49	103.99	103.99
0.28	0.33	0.3	0.7	0.315	162.63	162.63	115.39	115.39	106.64	106.64	103.83	103.83
0.48	0.53	0.7	0.3	0.495	125.50	125.50	106.37	106.37	102.83	102.83	101.59	101.59
0.48	0.53	0.3	0.7	0.515	123.56	123.56	105.88	105.88	102.61	102.61	101.47	101.47
0.68	0.73	0.7	0.3	0.695	112.85	112.85	103.15	103.15	101.35	101.35	100.72	100.72
0.68	0.73	0.3	0.7	0.715	112.13	112.13	102.96	102.96	101.26	101.26	100.67	100.67
0.88	0.93	0.7	0.3	0.895	107.60	107.60	101.75	101.75	100.67	100.67	100.29	100.29
0.88	0.93	0.3	0.7	0.915	107.25	107.25	101.65	101.65	100.23	100.23	100.26	100.26

when n = 1000												
π_{S1}	π_{S2}	w_1	w_2	π_S	$P = P_1$							
					0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9
					P_2							
					0.7	0.8	0.8	0.9	0.9	0.95	0.93	0.95
0.08	0.13	0.7	0.3	0.095	167.42	167.42	115.48	115.48	105.87	105.87	102.50	102.50
0.08	0.13	0.3	0.7	0.115	146.13	146.13	110.69	110.69	104.12	104.12	101.82	101.82
0.28	0.33	0.7	0.3	0.295	107.13	107.13	101.74	101.74	100.74	100.74	100.4	100.4
0.28	0.33	0.3	0.7	0.315	106.26	106.26	101.53	101.53	100.66	100.66	100.35	100.35
0.48	0.53	0.7	0.3	0.495	102.55	102.55	100.63	100.63	100.28	100.28	100.15	100.15
0.48	0.53	0.3	0.7	0.515	102.35	102.35	100.58	100.58	100.26	100.26	100.14	100.14
0.68	0.73	0.7	0.3	0.695	101.28	101.28	100.31	100.31	100.13	100.13	100.07	100.07
0.68	0.73	0.3	0.7	0.715	101.21	101.21	100.29	100.29	100.12	100.12	100.06	100.06
0.88	0.93	0.7	0.3	0.895	100.76	100.76	100.07	100.17	100.06	100.06	100.02	100.02

Table 4.3: Percent relative efficiency of the estimator $\hat{\pi}_{S1}$ (under proportional allocation) with respect to Hong et al. (1994) estimator $\hat{\pi}_S$ (under proportional allocation)

Table 4.1 shows that the values of the relative efficiency are greater than 100 % for all parameter values tabled. This shows the superiority of the Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum estimator) over Hong et al. (1994) estimator $\hat{\pi}_S$ (under proportional allocation). Table 4.2 exhibits that the percent relative efficiency of the proposed estimator $\hat{\pi}_{S1}$ (under optimum allocation) with respect to Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum allocation) decreases as sample size increases. It is observed that the percent relative efficiency is almost 100 % when the sample size is large. Larger gain is observed when the sample size n and π_S are small. However, the percent relative efficiency is always greater than 100 % which establishes the superiority of the proposed estimator $\hat{\pi}_{S1}$ (under optimum allocation) over Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum allocation). Table 4.3 - exhibits that the percent relative efficiency of the proposed estimator $\hat{\pi}_{S1}$ (under proportional allocation) with respect to Hong et al. (1994) estimator $\hat{\pi}_S$ (under proportional allocation) decreases as sample size and value of P increase. Larger gain in efficiency is observed for small as well as moderately large sample sizes. However, the percent relative efficiency is more than 100 % for all parametric values considered here, therefore the proposed estimator $\hat{\pi}_{S1}$ (under proportional allocation) is better than Hong et al. (1994) estimator $\hat{\pi}_S$ (under proportional allocation).

Finally from the above discussion we conclude that the proposed class of estimators $\hat{\pi}_{S1}$ under proportional as well as optimum allocations is better than the Hong et al.(1994) estimator $\hat{\pi}_S$ (under proportional allocation) and Kim and Warde (2004) estimator $\hat{\pi}_S$ (under optimum allocation).

Remark 4.1 – It is pertinent to note that the optimum values λ_0 , λ_{0P} and λ_{0O} in (3.4), (3.8) and (3.12) respectively depend on the Z_i's which can be estimated unbiased by sample proportion $\hat{Z}_i$ of the "Yes" answers. Hence it is suggested for the use of $\hat{\lambda}_0$, $\hat{\lambda}_{0P}$ and $\hat{\lambda}_{0O}$ respectively defined as

$$\hat{\lambda}_0 = \frac{\hat{\pi}_S^{\,2}}{\left\{\hat{\pi}_S^{\,2} + \sum_{i=1}^{k} w_i^{\,2}\,\frac{\hat{Z}_i(1-\hat{Z}_i)}{n_i(2P_i-1)^2}\right\}} \tag{4.4}$$

$$\lambda_{0P} = \frac{\hat{\pi}_S^{\,2}}{\left\{\hat{\pi}_S^{\,2} + \frac{1}{n}\sum_{i=1}^{k} w_i \hat{L}_i^{\,2}\right\}} \tag{4.5}$$

and

$$\hat{\lambda}_{0O} = \frac{\hat{\pi}_S^{\,2}}{\left\{\hat{\pi}_S^{\,2} + \frac{1}{n}\left(\sum_{i=1}^{k} w_i \hat{V}_i\right)^2\right\}} \tag{4.6}$$

where

$$\hat{L}_i = \sqrt{\left\{\hat{\pi}_{Si}(1-\hat{\pi}_{Si}) + \frac{P(1-P)}{(2P-1)^2}\right\}} \tag{4.7}$$

and

$$\hat{V}_i = \sqrt{\left\{\hat{\pi}_{Si}(1-\hat{\pi}_{Si}) + \frac{P_i(1-P_i)}{(2P_i-1)^2}\right\}} \tag{4.8}$$

For further discussion on this issue one can refer to Singh and Singh (1992) and Sampath et al. (1995).

5. Further Development

Motivated by Sampth et al. (1995) we consider a more generalized class of estimators for π_S namely

$$\hat{\pi}_{Sab} = \sum_{i=1}^{k} w_i \left\{a_i \hat{Z}_i + b_i\right\} \tag{5.1}$$

which reduces to (i) $\hat{\pi}_S = \sum_{i=1}^{k} w_i \hat{\pi}_{Si}$,

when

$a_i = (2P_i - 1)^{-1}$, $b_i = -(1-P_i)(2P_i-1)^{-1}$

(ii) Singh and Singh (1992) type estimator $\hat{\pi}_{SS} = \sum_{i=1}^{k} w_i \frac{\left\{h_i \hat{Z}_i - (1-P_i)\right\}}{(2P_i - 1)}$

when

$a_i = h_i (2P_i - 1)^{-1}$, $b_i = -(1-P_i)(2P_i-1)^{-1}$

(iii) the proposed estimator $\hat{\pi}_{S1} = h\hat{\pi}_S$

when

$a_i = h(2P_i - 1)^{-1}$, $b_i = -h(1-P_i)(2P_i-1)^{-1}$

The mean square error of the estimator $\hat{\pi}_{Sab}$ is

$$\text{MSE}\ (\hat{\pi}_{Sab}) = \text{Var}(\hat{\pi}_{Sab}) + \left[B(\hat{\pi}_{Sab})\right]^2 \tag{5.2}$$

Here

$$\text{Var}\ (\hat{\pi}_{Sab}) = \sum_{i=1}^{k} w_i^2 a_i^2 \frac{Z_i(1-Z_i)}{n_i} \tag{5.3}$$

$$B(\hat{\pi}_{Sab}) = E\ (\hat{\pi}_{Sab}) - \pi_S = E\left\{\sum_{i=1}^{k} w_i (a_i \hat{Z}_i + b_i) - \pi_S\right\}$$

$$= \sum_{i=1}^{k} w_i \left\{\left(a_i - \frac{1}{(2P_i - 1)}\right) Z_i + \left(b_i + \frac{(1-P_i)}{(2P_i - 1)}\right)\right\} \tag{5.4}$$

Thus the MSE of $\hat{\pi}_{Sab}$ is given by

$$MSE(\hat{\pi}_{Sab}) = \sum_{i=1}^{k} w_i^2 a_i^2 \frac{Z_i(1-Z_i)}{n_i} + \left[\sum_{i=1}^{k} w_i \left\{\left(a_i - \frac{1}{(2P_i-1)}\right) Z_i + \left(b_i + \frac{(1-P_i)}{(2P_i-1)}\right)\right\}\right]^2 \quad (5.5)$$

The mean square error of the estimator given in (5.4) is minimum if $a_i = 0$ and $b_i = \dfrac{Z_i - (1-P_i)}{(2P_i-1)}$ and the resulting mean square error is zero.

The exact mean square error of the estimator $(\hat{\pi}_{Sab})$ is zero for the optimum values of $a_i = 0$ and $b_i = \dfrac{Z_i - (1-P_i)}{(2P_i-1)}$. But the optimum value b_i needs the knowledge of Z_i for which the usual choice is $\hat{Z}_i$, the proportion of "Yes" answers in a stratum i.

Thus the substitution of $a_i = 0$ and $b_i = \dfrac{\hat{Z}_i - (1-P_i)}{(2P_i-1)}$ in $\hat{\pi}_{Sab}$ given by (5.1) yields the estimator

$$\hat{\pi}_{Sab} = \sum_{i=1}^{k} w_i \frac{\left(\hat{Z}_i - (1-P_i)\right)}{(2P_i-1)} = \sum_{i=1}^{k} w_i \hat{\pi}_{Si} = \hat{\pi}_S$$

which is the conventional unbiased estimator [see Hong et al. (1994)]. Hence it is inferred that the estimator $\hat{\pi}_S$ is the best estimator if one tries to formulate estimators better than the usual estimator $\hat{\pi}_S$ proceeding in the direction of Singh and Singh (1992), [see Sampath et al. (1995) p. 248] .

Acknowledgement

Authors are thankful to the Editor – in- chief and the learned referees for their valuable suggestions regarding improvement of the paper.

References

1. Hong, K., Yum, J. and Lee, H. (1994). A stratified randomized response technique, Korean Jour. Appl. Statist., 7, p. 141-147.
2. Kim, J. and Warde, W. (2004). A stratified Warner randomized response mode, Jour. Statist. Plan. Infer., 120, p. 155-165.
3. Mangat, N. S. (1994). An improved randomized response strategy, Jour. Roy. Statist. Soc., B, 56 (1), p. 93-95.
4. Mangat, N.S. and Singh, R. (1990). An alternative randomized procedure, Biometrika, 77, p. 439-442.

5. Sampath, S., Uthayakumaran, N. and Tracy, S. D. (1995). On the alternative estimator for randomized response technique, Jour. Ind. Soc. Agri. Statist., 47(3), p. 243-248.
6. Searles , D.T. (1965). The utilization of a known coefficient of variation in the estimation procedure, Jour. Amer. Stat. Assoc., 59, p. 1225 - 1226
7. Singh, R. and Mangat, N.S. (1996). Elements of Survey Sampling, Kluwer Academic Publishers, Dordrecht, the Netherlands.
8. Singh, S. and Singh, R. (1992). An alternative estimator for randomized response technique, Jour. Ind. Soc. Agri. Statist., 44, p. 149-154.
9. Singh H.P. and Tarray T.A. (2012). A stratified unknown repeated trials in randomized response sampling, Comm. Kor. Statist. Soc., 19, (6), p. 751-759**.**
10. Singh H.P. and Tarray T.A. (2014 a). An alternative to stratified Kim and Warde's randomized response model using optimal (Neyman) allocation, Model Assist. Statist. Appl., 9, p. 37-62.
11. Singh H.P. and Tarray T.A. (2014 b). An improvement over Kim and Elam stratified unrelated question randomized response model using Neyman allocation, Sankhya – B, DOI : 10.1007/s13571-014-0088-5.
12. Singh H.P. and Tarray T.A. (2014 c). An adroit stratified unrelated question randomized response model using neyman allocation, Sri. Jour. Appl. Statist., 15(2), p. 83-90.
13. Singh H.P. and Tarray T.A. (2014 d). An alternative to Kim and Warde's mixed randomized response model, Statist. Oper. Res. Trans. , 37 (2), p. 189 210.
14. Singh H.P. and Tarray T.A. (2014 e). A dexterous randomized response model for estimating a rare sensitive attribute using Poisson distribution, Statist. Prob. Lett., 90, p. 42-45.
15. Singh H.P. and Tarray T.A. (2014 f). An efficient alternative mixed randomized response procedure, Soc. Meth. Res., DOI: 10.1177/ 0049124114553309.
16. Singh H.P. and Tarray T.A. (2014 g). An alternative to Kim and Warde's mixed randomized response technique, Statistica, Anno, 73(3), p. 379-402.
17. Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias, Jour. Ameri. Statist. Assoc., 60, p. 63-69.
18. Warner, S. L. (1971). The Linear Randomized Response Model, Jour. Ameri. Statist. Assoc., 66, p. 884-888.