Brazilian Journal of Probability and Statistics (2006), 20, pp. 93–110.

An improved estimation procedure for estimating the proportion of a population possessing sensitive attribute in unrelated question randomized response technique

Housila P. Singh and **Nidhi Mathur**
Vikram University

Abstract: This paper considers the problem of estimating π_A, the proportion of human population possessing the sensitive attribute A, using the unrelated question randomized response technique envisaged by Horvitz et. al. (1967) and Greenberg et. al. (1969). A family of estimators $\hat{\pi}_H$ of population proportion π_A is defined. The bias and mean squared error (MSE) of the proposed estimator $\hat{\pi}_H$ are obtained. "Optimum estimator" in the family of estimators $\hat{\pi}_H$ is investigated. It has been shown that the "optimum estimator" is always better than usual estimator. Since the optimum estimator depends on the "optimum value" of the scalar H, which is function of unknown population parameters, so it has little practical utility. From the practical point of view, various estimators of π_A based on different estimated optimum values of H have been proposed with their properties. Efficiencies of the proposed estimators have been worked out numerically.

Key words: Estimation of proportion, mean squared error, simple random sampling with replacement, simple random unrelated question randomized response technique.

1 Introduction

Warner(1965) introduced an ingenious interviewing technique known as randomized response (RR) technique for estimating the proportion π_A of a population possessing sensitive attribute (say) A without requiring the individual respondent to report his actual classification, where it be A or not-A, to the interviewer. His design uses two related questions each of which divides the sample/population into two mutually exclusive and complementary classes. Horvitz et. al. (1967) felt that the confidence of the respondent provided by RR technique might be further enhanced if one of the two questions is referred to a non-stigmatized attribute (say) Y which is unrelated to the sensitive attribute. The model developed by them is known as unrelated question RR model (U-model). The theoretical framework for this model was given by Greenberg et. al. (1969). If π_Y, the population proportion with the non-stigmatized attribute Y, is known in advance, only one sample is required to estimate π_A, the population proportion with the sensitive attribute. If π_Y is not known beforehand, then according this model (U-model), two samples of n_1 and n_2 respondents are selected using simple random sampling with replacement (SRSWR) such that $n_1+n_2=n$, which is the required total sample size. Both the samples are used to gather information on sensitive attribute A and the non-stigmatized (netural) attribute Y by using different RR devices with respondents in two samples. Each respondent chooses a statement randomly from the RR device provided to him and is required to report "yes" if the selected statement points to his actual status and "no" otherwise. Let the statements regarding possessing of the sensitive attribute A and the neutral attribute Y be represented with probabilities p_i and $(1-p_i)$ respectively in the RR device S_i used for the respondents in the i-th sample, $i=1,2$. Assuming that the respondents report truthfully, the probability of "yes" answer when a respondent in the i-th sample is confronted with the randomization device S_i, is given by

$$\theta_i = p_i\pi_A + (1-p_i)\pi_Y, \quad (i=1,2), \tag{1.1}$$

where the values of θ_1 and θ_2 are estimated by the proportion of "yes" responses recorded in sample 1 and sample 2 respectively.

Suppose we let the observed proportion of "yes" answers reported in the first and second samples be designated by $\hat{\theta}_1 = n_1'/n_1$ and $\hat{\theta}_2 = n_2'/n_2$ respectively, where n_1' and n_2' are the number of "yes" answers in the two corresponding samples. Then, the sample estimate of π_A is given by

$$\hat{\pi}_{As} = \frac{(1-p_2)\hat{\theta}_1 - (1-p_1)\hat{\theta}_2}{(p_1-p_2)}, \quad (p_1 \neq p_2), \tag{1.2}$$

which is due to Greenberg et. al. (1969).

The estimator $\hat{\pi}_{As}$ is unbiased for π_A and has the variance

$$\mathrm{Var}(\hat{\pi}_{As}) = \frac{1}{(p_1-p_2)^2}\left[\frac{\theta_1(1-\theta_1)(1-p_2)^2}{n_1} + \frac{\theta_2(1-\theta_2)(1-p_1)^2}{n_2}\right]. \tag{1.3}$$

To increase the efficiency of the estimator $\hat{\pi}_{As}$, one has to select the design parameters n_1, n_2, p_1 and p_2 optimally. In practice, the size of the sample $n = n_1+n_2$ will remain fixed if the cost of the survey is fixed and one has to choose n_1 and n_2 optimally subject to the constraint

$$n = n_1 + n_2. \tag{1.4}$$

Thus for a given value of n, if n_1 and n_2 are chosen so as to minimize $\mathrm{Var}(\hat{\pi}_{As})$ in (1.3), the minimizing ratio n_1/n_2 is given by

$$\frac{n_1}{n_2} = \frac{\sqrt{\theta_1(1-\theta_1)(1-p_2)^2}}{\sqrt{\theta_2(1-\theta_2)(1-p_1)^2}}.$$

For this allocation, the minimum variance of $\hat{\pi}_{As}$ is given by

$$V_o(\hat{\pi}_{As}) = \frac{[(1-p_2)\sqrt{\theta_1(1-\theta_1)} + (1-p_1)\sqrt{\theta_2(1-\theta_2)}]^2}{n(p_1-p_2)^2}. \tag{1.5}$$

A rich growth of literature on randomized response technique has been documented in several research publications, for instance, see Fox and Tracy (1986), Chaudhuri and Mukerjee (1987,1988), Hedayat and Sinha (1991), Sheers (1992), Bellhouse (1995) and Tracy and Mangat (1996) . Recent developments on randomized response technique are due to Chang and Liang (1996), Singh and Joarder (1997), Mangat et. al. (1997), Arnab (1998), Chaudhuri et. al. (1998), Lee and Hong (1998), Van Der Heijden et. al. (1998), Tracy and Mangat (1998, 1999), Singh et. al. (1998, 2000), Singh and Tracy (1999), Tracy and Oshan (1999), Bhargwa and Singh (1999), Singh and King (1999), Tracy and Singh (1999), Padmawar and Vijayan (2000),Chang and Huang (2001), Arnab and Singh (2002), Mangat and Singh (2002), Chaudhuri (2001a,b, 2002), Bhargava et. al. (2002), Javed et. al. (2002), Gupta et. al. (2002), Singh (2002), Christofides (2003), Smith and Street (2003) and Singh and Mathur (2002a,b,c,d,e, 2003a,b, 2004).

In this paper, we have proposed a family of estimators of the proportion π_A and its properties are studied. Several estimators of π_A based on 'estimated optimum values' have been also suggested with their properties.

2 The suggested family of estimators

Motivated by Goodman (1953), Searls (1964) and Searls and Intarapanich (1990), we define a family of estimators of the population proportion π_A as

$$\hat{\pi}_H = H\hat{\pi}_{As}, \tag{2.1}$$

where H is a scalar to be chosen suitably by the investigator.

The bias and mean squared error of $\hat{\pi}_H$ are respectively given by

$$B(\hat{\pi}_H) = (H-1)\pi_A \tag{2.2}$$

and

$$MSE(\hat{\pi}_H) = H^2\{\pi_A^2 + \mathrm{Var}(\hat{\pi}_{As})\} - 2H\pi_A^2 + \pi_A^2, \tag{2.3}$$

where $\mathrm{Var}(\hat{\pi}_{As})$ is given by (1.3).

From (1.3) and (2.3) we have

$$MSE(\hat{\pi}_H) < \mathrm{Var}(\hat{\pi}_{As}) \text{ if } (H^2-1)\mathrm{Var}(\hat{\pi}_{As}) + (H-1)^2\pi_A^2 < 0,$$

i.e., if

$$\frac{\pi_A^2 - \mathrm{Var}(\hat{\pi}_{As})}{\pi_A^2 + \mathrm{Var}(\hat{\pi}_{As})} < H \leq 1. \tag{2.4}$$

The range of dominance of H in which proposed estimator $\hat{\pi}_H$ is better than Greenberg et. al.'s (1969) estimator $\hat{\pi}_{As}$ has been computed using (2.4) for different values of $(p_1, p_2, \pi_A, n_1, n_2, \pi_Y)$ and the values have been displayed in Table 1(a).

Table 1(a) *Range of H for different values of $(n_1, n_2), (p_1, p_2)$, π_Y and π_A*

	$\pi_A = 0.05$					$\pi_A = 0.1$				
$(p_1, p_2) \rightarrow$	(0.7, 0.3)									
$(n_1, n_2) \rightarrow$ $\pi_Y \downarrow$	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
0.1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
0.3	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
0.5	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
0.9	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
	(0.8, 0.2)									
0.1	0~1	0~1	0~1	0~1	0.06~1	0~1	0.086~1	0.081~1	0.09~1	0.049~1
0.3	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.35~1
0.5	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.26~1
0.9	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.15~1
	(0.9, 0.1)									
0.1	0~1	0~1	0~1	0~1	0.30~1	0~1	0.34~1	0.26~1	0.27~1	0.62~1
0.3	0~1	0~1	0~1	0~1	0.15~1	0~1	0.25~1	0.18~1	0.19~1	0.57~1
0.5	0~1	0~1	0~1	0~1	0.05~1	0~1	0.18~1	0.12~1	0.12~1	0.52~1
0.9	0~1	0~1	0~1	0~1	0~1	0~1	0.11~1	0.02~1	0.03~1	0.45~1

Further, the optimum values of n_1 and n_2 which minimizes the MSE of $\hat{\pi}_H$ are same as given in (1.4). Using (1.4) and (2.3) we get the minimum MSE of $\hat{\pi}_H$ as

$$\min MSE(\hat{\pi}_H) = \pi_A^2 \left[(H-1)^2 + H^2 \frac{V_0(\hat{\pi}_{As})}{n\pi_A^2} \right], \tag{2.5}$$

where $V_0(\hat{\pi}_{As})$ is given by (1.5).

It is observed from (1.5) and (2.5) that

$$\min MSE(\hat{\pi}_H) < V_0(\hat{\pi}_{As}) \quad \text{if} \quad \frac{\pi_A^2 - V_0(\hat{\pi}_{As})}{\pi_A^2 + V_0(\hat{\pi}_{As})} < H \leq 1. \tag{2.6}$$

We have computed the range of dominance of H by using (2.6) and presented in Table 1(b).

It is observed from Table 1(a) and 1(b) that the suggested estimator $\hat{\pi}_H$ is better than Greenberg et. al.'s (1969) estimator $\hat{\pi}_{As}$ for full range of H (i.e. $0 < H \leq 1$) except in few cases like

(i) $(p_1, p_2) = (0.8, 0.2)$, $(n_1, n_2) = (50, 50)$, $\pi_A = 0.1$, $0.1 \leq \pi_Y \leq 0.9$,

(ii) $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = (50, 50)$, $\pi_A = 0.05$, $0.1 \leq \pi_Y \leq 0.9$ and

(iii) $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = \{(25, 5), (20, 20), (50, 50)\}$, $\pi_A = 0.1$, $0.1 \leq \pi_Y \leq 0.9$.

Table 1(b) *Range of H for different values of* (n_1, n_2), (p_1, p_2), π_Y *and* π_A *in optimum case.*

	$\pi_A = 0.05$					$\pi_A = 0.1$				
$(p_1, p_2) \rightarrow$	(0.7, 0.3)									
$(n_1, n_2) \rightarrow$ $\pi_Y \downarrow$	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
0.1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.28~1
0.3	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.03~1
0.5	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
0.9	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1
	(0.8, 0.2)									
0.1	0~1	0~1	0~1	0~1	0.19~1	0~1	0.09~1	0.23~1	0.33~1	0.60~1
0.3	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.03~1	0.14~1	0.45~1
0.5	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.04~1	0.37~1
0.9	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0~1	0.30~1
	(0.9, 0.1)									
0.1	0~1	0~1	0.07~1	0.18~1	0.48~1	0~1	0.36~1	0.47~1	0.56~1	0.75~1
0.3	0~1	0~1	0~1	0.004~1	0.33~1	0~1	0.25~1	0.38~1	0.47~1	0.69~1
0.5	0~1	0~1	0~1	0~1	0.23~1	0~1	0.18~1	0.32~1	0.41~1	0.65~1
0.9	0~1	0~1	0~1	0~1	0.14~1	0~1	0.13~1	0.26~1	0.36~1	0.62~1

Thus we infer that the performance of the suggested estimator $\hat{\pi}_H$ is better for wider range of H when sample size (n_1, n_2) are moderately large.

We have also computed the percent relative efficiencies of $\hat{\pi}_H$ with respect to $\hat{\pi}_{As}$ with optimum allocation, by using the formula:

$$PRE(\hat{\pi}_H, \hat{\pi}_{As}) = \frac{V_0(\hat{\pi}_{As})}{\min MSE(\hat{\pi}_H)} \times 100 = \frac{A^2}{[n(H-1)^2(p_1-p_2)^2\pi_A^2 + H^2A^2]} \times 100 \quad (2.7)$$

for different values of $\pi_1, \pi_2, p_1, p_2, \pi_A, \pi_H$ and H, where

$$A = \left[(1-p_2)\sqrt{\theta_1(1-\theta_1)} + (1-p_1)\sqrt{\theta_2(1-\theta_2)}\right].$$

The values of $PRE(\hat{\pi}_H, \hat{\pi}_{As})$ have been displayed in Table 2(a) and 2(b).

The results shown in Tables 2(a) and 2(b) indicate that

(i) when $\pi_A = 0.05$, the estimator $\hat{\pi}_H$ is more efficient than Greenberg et. al's (1969) estimator $\hat{\pi}_{As}$ with substantial gain except in very few cases, particularly, when the sample sizes (n_1, n_2) are large (i.e. $(n_1, n_2) \geq (50, 50)$);

(ii) when $\pi_A = 0.1$, the performance of the suggested estimator $\hat{\pi}_H$ is better than $\hat{\pi}_{As}$ except in some cases;

(iii) The $PRE(\hat{\pi}_H, \hat{\pi}_{As})$ decreases as π_A increases and it is stable when H approaches unity;

Table 2(a) *Percent Relative Efficiency of* $\hat{\pi}_H$ *with respect to* $\hat{\pi}_{As}$.

$\pi_A = 0.05$																
$(p_1, p_2) \rightarrow$		(0.7, 0.3)					(0.8, 0.2)					(0.9, 0.1)				
$(n_1, n_2) \rightarrow$		(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
$H \downarrow$	$\pi_Y \downarrow$															
	0.1	1684.96	632.73	482.17	389.49	198.61	773.56	271.87	205.30	164.92	83.14	407.05	139.47	104.97	84.15	42.25
0.1	0.3	2820.59	1157.93	894.34	728.50	378.02	1276.45	465.06	352.90	284.32	144.21	576.99	200.02	150.77	120.98	60.86
	0.5	3484.21	1512.80	1179.19	966.14	507.59	1623.99	607.05	462.31	373.30	190.20	705.69	246.84	186.28	149.58	75.86
	0.9	4011.16	1825.11	1434.27	1181.30	627.73	1949.72	747.00	570.91	462.01	236.47	847.33	299.35	226.21	181.79	91.73
	0.1	1033.38	604.92	501.05	427.62	246.79	688.10	321.56	253.93	209.80	112.26	442.16	180.67	139.44	113.53	58.85
0.25	0.3	1247.25	865.59	750.73	662.78	417.95	909.42	488.09	396.29	333.56	186.19	568.47	248.30	193.74	158.84	83.57
	0.5	1324.74	985.61	873.77	784.72	519.84	1017.12	588.40	485.98	413.93	237.72	649.51	296.84	233.46	192.38	102.34
	0.9	1372.34	1068.32	961.78	874.57	601.74	1096.81	673.31	564.36	485.75	286.34	727.21	347.78	275.82	228.54	123.66
	0.1	377.03	338.19	321.62	306.60	248.56	348.66	277.44	251.73	230.38	161.78	309.85	213.58	184.86	162.95	102.31
0.50	0.3	387.81	365.54	355.34	345.69	304.36	368.88	319.20	299.07	281.32	216.95	332.88	249.24	221.42	199.19	132.62
	0.5	390.97	374.09	366.18	358.60	324.97	376.05	335.85	318.80	303.40	244.39	344.06	268.85	242.37	220.63	152.32
	0.9	392.76	379.04	372.53	366.25	337.74	380.60	346.94	332.25	318.76	264.95	352.93	285.70	260.85	239.98	171.41
	0.1	176.58	174.24	173.09	171.96	166.51	174.92	169.46	166.86	164.33	152.78	172.21	164.48	157.42	153.04	134.35
0.75	0.3	177.16	175.93	175.33	174.73	171.78	176.13	172.91	171.35	169.82	162.54	173.88	168.35	163.16	159.87	145.24
	0.5	177.32	176.42	175.97	175.53	173.33	176.53	174.08	172.89	171.70	166.03	174.62	170.10	165.80	163.05	150.57
	0.9	177.41	176.69	176.33	175.98	174.21	176.78	174.81	173.84	172.88	168.25	175.18	171.43	167.83	165.52	154.84
	0.1	123.36	123.23	123.09	122.99	122.54	123.23	122.79	122.57	122.34	121.25	123.01	122.36	121.71	121.28	119.18
0.90	0.3	123.41	123.34	123.27	123.22	122.98	123.33	123.07	122.94	122.82	122.18	123.15	122.69	122.24	121.94	120.46
	0.5	123.42	123.37	123.32	123.28	123.11	123.36	123.17	123.07	122.97	122.49	123.21	122.84	122.47	122.23	121.03
	0.9	123.43	123.39	123.34	123.32	123.18	123.38	123.22	123.15	123.07	122.68	123.25	122.95	122.65	122.45	121.46

Table 2(b) *Percent Relative Efficiency of* $\hat{\pi}_H$ *with respect to* $\hat{\pi}_{As}$.

$\pi_A = 0.1$																
$(p_1, p_2) \rightarrow$		(0.7, 0.3)					(0.8, 0.2)					(0.9, 0.1)				
$(n_1, n_2) \rightarrow$		(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
$H \downarrow$	$\pi_Y \downarrow$															
0.1	0.1	649.35	226.24	170.65	136.99	68.97	299.40	101.83	76.57	61.35	30.77	170.65	57.54	43.22	34.60	17.33
	0.3	1025.13	366.78	277.63	223.34	112.93	437.64	150.26	113.12	90.70	45.56	214.83	72.65	54.59	43.72	21.91
	0.5	1277.06	465.30	353.08	284.47	144.29	535.19	185.00	139.39	111.83	56.23	246.72	83.62	62.84	50.34	25.23
	0.9	1475.93	545.67	414.91	334.71	170.20	618.97	215.20	162.28	130.24	65.55	276.84	94.02	70.68	56.62	28.39
0.25	0.1	615.38	275.86	216.22	177.78	94.12	347.83	135.59	103.90	84.21	43.24	216.22	79.21	60.15	48.48	24.62
	0.3	811.05	408.34	327.13	272.86	149.15	466.78	193.16	149.38	121.78	63.30	263.99	98.87	75.32	60.83	31.00
	0.5	909.64	488.28	396.45	333.70	186.28	539.63	232.05	180.59	147.81	77.48	296.72	112.86	86.16	69.68	35.62
	0.9	974.59	546.98	448.58	380.18	215.72	596.13	264.38	206.83	169.85	89.69	326.37	125.91	96.33	78.00	39.98
0.50	0.1	339.62	260.87	233.77	211.76	144.00	285.71	181.82	153.85	133.33	80.00	233.77	127.66	104.05	87.80	49.32
	0.3	360.98	302.06	279.26	259.67	192.23	315.02	221.09	192.40	170.30	108.18	256.03	148.87	123.11	104.94	60.39
	0.5	368.89	319.24	299.11	281.37	217.01	328.32	241.69	213.52	191.24	125.65	268.81	162.33	135.49	116.27	68.02
	0.9	373.38	329.52	311.23	294.88	233.51	336.95	256.19	228.78	206.66	139.32	279.02	173.85	146.28	126.26	74.96
0.75	0.1	174.33	167.83	164.76	161.80	148.45	170.21	156.86	150.94	145.45	123.08	164.76	143.71	135.08	127.43	99.31
	0.3	175.67	171.60	169.63	167.71	158.72	172.60	163.11	158.75	154.61	136.78	167.32	149.72	142.23	135.46	109.42
	0.5	176.13	172.92	171.36	169.82	162.55	173.57	165.72	162.05	158.55	143.07	168.63	152.90	146.09	139.86	115.27
	0.9	176.38	173.65	172.32	171.00	164.73	174.16	167.34	164.13	161.04	147.18	169.61	155.33	149.05	143.27	119.98
0.90	0.1	123.19	122.65	122.38	122.12	120.81	122.85	121.65	121.07	120.48	117.65	122.38	120.29	119.27	118.27	113.49
	0.3	123.29	122.96	122.80	122.64	121.83	123.05	122.24	121.83	121.43	119.48	122.61	120.94	120.12	119.32	115.44
	0.5	123.33	123.07	122.94	122.82	122.18	123.12	122.47	122.14	121.82	120.22	122.72	121.26	120.55	119.85	116.44
	0.9	123.35	123.13	123.02	122.92	122.38	123.17	122.61	122.33	122.05	120.67	122.80	121.51	120.87	120.24	117.18

(iv) Larger gain in efficiency is observed when the sample sizes (n_1, n_2) are moderately large i.e. $(n_1, n_2) < (50, 50)$.

Thus we conclude from the results of the Tables 1(a), 1(b), 2(a) and 2(b) that there is enough scope of choosing the value of scalar H in order to increase the efficiency of the suggested estimator $\hat{\pi}_H$ over Greenberg et. al's (1969) estimator $\hat{\pi}_{As}$. For larger gain in efficiency, small sample sizes are to be preferred in practice. Such sample sizes are desirable when the survey procedure like RR technique is expensive.

3 Optimum estimator in the class $\hat{\pi}_H$ at (2.1)

Differentiating (2.3) with respect to H and equating it to zero, we get the optimum value of H as

$$H_{opt} = \frac{\pi_A^2}{\pi_A^2 + \mathrm{Var}(\hat{\pi}_{As})} \tag{3.1}$$

$$= \frac{\pi_A^2}{\pi_A^2 + \frac{1}{(p_1-p_2)^2}\left\{\frac{\theta_1(1-\theta_1)(1-p_2)^2}{n_1} + \frac{\theta_2(1-\theta_2)(1-p_1)^2}{n_2}\right\}}. \tag{3.2}$$

Substitution (3.1) in (2.1) yields the 'optimum estimator' as

$$\hat{\pi}_{H_{opt}} = \frac{\pi_A^2\hat{\pi}_{As}}{\{\pi_A^2 + \mathrm{Var}(\hat{\pi}_{As})\}} = H_{opt}\hat{\pi}_{As}$$

$$= \frac{\pi_A^2\hat{\pi}_{As}}{\pi_A^2 + \frac{1}{(p_1-p_2)^2}\left\{\frac{\theta_1(1-\theta_1)(1-p_2)^2}{n_1} + \frac{\theta_2(1-\theta_2)(1-p_1)^2}{n_2}\right\}}. \tag{3.3}$$

Putting (3.1) (or (3.2) in (2.3) we get the MSE of the optimum estimator $\hat{\pi}_{H_{opt}}$ as

$$MSE(\hat{\pi}_{H_{opt}}) = \frac{\pi_A^2\mathrm{Var}(\hat{\pi}_{As})}{\{\pi_A^2 + \mathrm{Var}(\hat{\pi}_{As})\}}. \tag{3.4}$$

Thus the relative efficiency of $\hat{\pi}_{H_{opt}}$ with respect to $\hat{\pi}_{As}$ is given by

$$PRE(\hat{\pi}_{H_{opt}}, \hat{\pi}_{As}) = 1 + \frac{\mathrm{Var}(\hat{\pi}_{As})}{\pi_A^2}$$

$$= 1 + \frac{1}{(p_1-p_2)^2\pi_A^2}\left[\frac{\theta_1(1-\theta_1)(1-p_2)^2}{n_1} + \frac{\theta_2(1-\theta_2)(1-p_1)^2}{n_2}\right], \tag{3.5}$$

which clearly indicates that the optimum estimator $\hat{\pi}_{H_{opt}}$ is always more efficient than Greenberg et. al.'s (1969) estimator $\hat{\pi}_{As}$. It is to be noted that the optimum estimator $\hat{\pi}_{H_{opt}}$ in (3.3) depends on the unknown population parameters θ_i $(i = 1, 2)$ and π_A. So it is not useful in practice. However, due to past experience or data or acquaintance with the experimental material one may have the guessed value of the parameters θ_i $(i = 1, 2)$ and π_A [see Thompson (1968) and Mehta and Srinivasan (1977)] and hence the guessed value of the optimum value (H_{opt}) of H in (3.1). Let $\tilde{H}_{opt}$ be the guessed value of H_{opt} such that $\tilde{H}_{opt} = \alpha H_{opt}$, where $\alpha(> 0)$ is a departure from true optimum value H_{opt} of H. Then the estimator of π_A based on guessed value $\tilde{H}_{opt}$ is defined by

$$\tilde{\pi}_{\tilde{H}_{opt}} = \tilde{H}_{opt}\hat{\pi}_{As} = \alpha H_{opt}\hat{\pi}_{As}. \tag{3.6}$$

Such a procedure is discussed by Searls (1967), Hirano (1972) and Singh and Shukla (2002). The bias and mean squared error (MSE) of $\tilde{\pi}_{\tilde{H}_{opt}}$ are respectively given by

$$B(\tilde{\pi}_{\tilde{H}_{opt}}) = \left[\frac{\alpha}{\{1 + (CV(\hat{\pi}_{As}))^2\}} - 1\right]\pi_A \tag{3.7}$$

and

$$MSE(\tilde{\pi}_{\tilde{H}_{opt}}) = \left[\frac{\alpha^2}{\{1 + (CV(\hat{\pi}_{As}))^2\}} - \frac{2\alpha}{\{1 + (CV(\hat{\pi}_{As}))^2\}} + 1\right]\pi_A^2, \tag{3.8}$$

where $CV(\hat{\pi}_{As}) = \sqrt{\text{Var}(\hat{\pi}_{As})}/\pi_A$ is the coefficient of variation of the estimator $\hat{\pi}_{As}$ and $\text{Var}(\hat{\pi}_{As})$ is given by (1.3).

It follows from (1.3) and (3.8) that $MSE(\hat{\pi}_{As}) < \text{Var}(\hat{\pi}_{As})$ if $(\alpha - 1)^2 < \{CV(\hat{\pi}_{As})\}^4$ i.e. if

$$(\alpha - 1)^2 < \frac{1}{(p_1 - p_2)^4\pi_A^4}\left[\frac{\theta_1(1-\theta_1)(1-p_2)^2}{n_1} + \frac{\theta_2(1-\theta_2)(1-p_1)^2}{n_2}\right]^2. \tag{3.9}$$

One can calculate the range of dominance of α in which $\tilde{\pi}_{\tilde{H}_{opt}}$ is better than Greenberg et. al.'s (1969) estimator $\hat{\pi}_{As}$ for different values of $n_1, n_2, p_1, p_2, \pi_Y$ and π_A.

4 Estimators based on estimated optimum values of H from the sample

If the experimenter is inexperienced and unable to guess the values of unknown population parameters, in such a situation, it is advisable to replace these population parameters by their estimates available from the sample data at hand. Thus replacing (θ_i, π_A) $(i = 1, 2)$ by their estimates $(\hat{\theta}_i, \hat{\pi}_{As})$ $(i = 1, 2)$ in (3.2) we get an estimate of H_{opt} as

$$\hat{H}_{opt}^{(1)} = \frac{\hat{\pi}_{As}^2}{\hat{\pi}_{As}^2 + \frac{1}{(p_1 - p_2)^2}\left\{\frac{\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{(n_1 - 1)} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{(n_2 - 1)}\right\}}. \tag{4.1}$$

Substitution of (4.1) in (3.3) yields an estimator of π_A as

$$\hat{\pi}^{(1)}_{\hat{H}_{opt}} = \frac{\hat{\pi}^3_{As}}{\hat{\pi}^2_{As} + \frac{1}{(p_1-p_2)^2}\left\{\frac{\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{n_1} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{n_2}\right\}}. \quad (4.2)$$

Replacing π_A by $\hat{\pi}_{As}$ and $\mathrm{Var}(\hat{\pi}_{As})$ by its unbiased estimator

$$\hat{\mathrm{Var}}(\hat{\pi}_{As}) = \frac{1}{(p_1-p_2)^2}\left[\frac{\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{(n_1-1)} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{(n_2-1)}\right] \quad (4.3)$$

in (3.1), we find another estimate of H_{opt} as

$$\hat{H}^{(2)}_{opt} = \frac{\hat{\pi}^2_{As}}{\hat{\pi}^2_{As} + \frac{1}{(p_1-p_2)^2}\left\{\frac{\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{(n_1-1)} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{(n_2-1)}\right\}}. \quad (4.4)$$

Substitution of (4.4) in (3.3) yields another estimator of π_A as

$$\hat{\pi}^{(2)}_{\hat{H}_{opt}} = \frac{\hat{\pi}^3_{As}}{\hat{\pi}^2_{As} + \frac{1}{(p_1-p_2)^2}\left\{\frac{\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{(n_1-1)} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{(n_2-1)}\right\}}. \quad (4.5)$$

Following Thompson (1968) and Lemmer (1981) we define a more flexible estimator of π_A as

$$\hat{\pi}^{(h_1,h_2)}_{\tilde{H}_{opt}} = \frac{\hat{\pi}^3_{As}}{\hat{\pi}^2_{As} + \frac{1}{(p_1-p_2)^2}\left\{\frac{h_1\hat{\theta}_1(1-\hat{\theta}_1)(1-p_2)^2}{n_1} + \frac{h_2\hat{\theta}_2(1-\hat{\theta}_2)(1-p_1)^2}{n_2}\right\}}, \quad (4.6)$$

where $(h_1, h_2) > 0$ are constants.

It is to be noted that for $h_1 = h_2 = 1$, $\hat{\pi}^{(h_1,h_2)}_{\tilde{H}_{opt}}$ reduces to the estimator $\hat{\pi}^{(1)}_{\tilde{H}_{opt}}$ while for $h_i = n_i/(n_i-1)$, $i = 1, 2$, it boils down to the estimator $\hat{\pi}^{(2)}_{\tilde{H}_{opt}}$ given by (4.5).

The exact MSE of an estimator $G = \hat{\pi}^{(1)}_{\hat{H}_{opt}}, \hat{\pi}^{(2)}_{\hat{H}_{opt}}$, and $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ is given by

$$MSE(G) = \sum_{n_1'=0}^{n_1}\sum_{n_2'=0}^{n_2}(G-\pi_A)^2\, {}^{n_1}C_{n_1'}\, {}^{n_2}C_{n_2'}\, \theta_1^{n_1'}\theta_2^{n_2'}(1-\theta_1)^{(n_1-n_1')}(1-\theta_2)^{(n_2-n_2')}, \quad (4.7)$$

where ${}^{n_i}C_{n_i'} = \frac{n_i!}{n_i'!(n_i-n_i')!}$, $i = 1, 2$, and $n_i!$ stands for factorial n_i $(i = 1, 2)$.

The percent relative efficiency (PRE) of an estimator G with respect to $\hat{\pi}_{As}$ is given by

$$PRE(G, \hat{\pi}_{As}) = \frac{\left[\left\{\theta_1(1-\theta_1)(1-p_2)^2/n_1\right\} + \left\{\theta_2(1-\theta_2)(1-p_1)^2/n_2\right\}\right]/(p_1-p_2)^2}{\left[\sum_{n_1'=0}^{n_1}\sum_{n_2'=0}^{n_2}(G-\pi_A)^2\, {}^{n_1}C_{n_1'}\, {}^{n_2}C_{n_2'}\theta_1^{n_1'}\theta_2^{n_2'}(1-\theta_1)^{(n_1-n_1')}(1-\theta_2)^{(n_2-n_2')}\right]} \times 100. \quad (4.8)$$

The $PRE(G, \hat{\pi}_{As})$ have been computed for different values of $(n_1, n_2, p_1, p_2, h_1, h_2, \pi_A$ and $\pi_Y)$ and displayed in Tables 3(a) and 3(b) respectively.

It is observed from the results shown in Tables 3(a) and 3(b) that:

(a) when $\pi_A = 0.05$, the proposed estimator $\hat{\pi}^{(1)}_{\hat{H}_{opt}}$ and $\hat{\pi}^{(2)}_{\hat{H}_{opt}}$ are better than Greenberg et. al.'s (1969) estimator $\hat{\pi}_{As}$ with substantial gain in efficiency. The estimator $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ with $(h_1, h_2) = (5, 6), (8, 9)$ is better than $\hat{\pi}_{As}$ except in the case $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = (50, 50)$ and $\pi_Y = 0.10$;

(b) when $\pi_A = 0.10$, the estimator $\hat{\pi}^{(1)}_{\hat{H}_{opt}}$ and $\hat{\pi}^{(2)}_{\hat{H}_{opt}}$ are better than $\hat{\pi}_{As}$ with considerable gain in efficiency except in the cases:
(i) $(p_1, p_2) = (0.8, 0.2)$, $(n_1, n_2) = (50, 50)$, $\pi_Y = 0.10$;
(ii) $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = (50, 50)$, $\pi_Y = 0.10$.
The performance of the estimator $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ with $(h_1, h_2) = (5, 6), (8, 9)$ is also better than $\hat{\pi}_{As}$ except in few cases:
(i) $(p_1, p_2) = (0.8, 0.2)$, $(n_1, n_2) = (50, 50)$, $0.1 \le \pi_Y \le 0.5$;
(ii) $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = (50, 50)$, $0.1 \le \pi_Y \le 0.9$.
We also note that the performance of the estimator $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ with $(h_1, h_2) = (8, 9)$ is not appreciable in the case $(p_1, p_2) = (0.9, 0.1)$, $(n_1, n_2) = (25, 5), (20, 20), (20, 30)$ and $\pi_Y = 0.3$. It is further observed that $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ with $(h_1, h_2) = (5, 6), (8, 9)$ is more efficient than all the estimators $\hat{\pi}_{As}$, $\hat{\pi}^{(1)}_{\hat{H}_{opt}}$ and $\hat{\pi}^{(2)}_{\hat{H}_{opt}}$ for $(p_1, p_2) = (0.7, 0.3), (0.8, 0.2)$; $(n_1, n_2) = (5, 5), (25, 5), (20, 20), (20, 30)$ and $\pi_Y \in [0.1, 0.9]$;

(c) The estimator $\hat{\pi}^{(2)}_{\hat{H}_{opt}}$ is better than $\hat{\pi}^{(1)}_{\hat{H}_{opt}}$ when $\pi_A = 0.05$. Leaving few cases it is also true with $\pi_A = 0.10$;

(d) The PRE decreases as π_A increases;

(e) Larger gain in efficiency is seen by using the proposed estimators $\hat{\pi}^{(1)}_{\hat{H}_{opt}}$, $\hat{\pi}^{(2)}_{\hat{H}_{opt}}$ and $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ over Greenberg et. al's (1969) estimator $\hat{\pi}_{As}$ for smaller sample sizes.

Table 3(a) *Percent Relative Efficiency of* $\hat{\pi}_H^{(1)}$, $\hat{\pi}_H^{(2)}$ *and* $\hat{\pi}_H^{(h_1,h_2)}$.

$\pi_A = 0.05$																
	$(p_1, p_2) \rightarrow$	(0.7, 0.3)					(0.8, 0.2)					(0.9, 0.1)				
π_Y	$(n_1, n_2) \rightarrow$ Estimator $\downarrow$	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
0.1	$\hat{\pi}_H^{(1)}$	227.66	199.46	180.10	173.10	152.31	226.42	169.87	156.00	151.45	122.26	223.16	138.67	142.67	141.12	105.56
	$\hat{\pi}_H^{(2)}$	259.97	210.90	183.77	176.02	153.10	256.22	175.61	158.34	153.40	122.55	250.42	140.84	144.14	142.42	105.61
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	777.18	530.09	383.77	348.78	221.38	610.81	283.64	241.85	229.09	130.19	491.10	163.87	174.85	171.46	93.06
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	1073.6	650.88	203.58	395.17	225.82	741.87	293.17	249.35	236.45	122.97	546.91	156.61	169.88	166.84	83.74
0.3	$\hat{\pi}_H^{(1)}$	190.91	183.53	187.25	183.99	172.80	193.69	173.27	164.91	161.63	141.77	205.56	145.15	138.39	136.05	115.10
	$\hat{\pi}_H^{(2)}$	212.04	192.91	191.22	187.41	174.00	214.44	179.87	167.55	163.90	142.37	227.57	148.11	139.92	137.41	115.27
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	572.41	490.31	452.68	414.85	323.64	512.66	354.74	291.78	270.64	185.38	464.99	200.79	189.32	181.26	113.32
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	824.86	683.19	282.99	512.59	360.80	675.77	411.70	324.52	298.20	184.99	551.72	202.15	193.84	185.17	105.30
0.5	$\hat{\pi}_H^{(1)}$	174.57	181.69	190.84	188.97	180.70	171.68	175.27	173.26	171.42	152.51	185.72	150.46	145.40	143.99	122.84
	$\hat{\pi}_H^{(2)}$	190.53	190.94	195.04	192.75	182.08	186.72	181.95	176.34	174.26	153.30	202.19	153.46	147.03	145.48	123.13
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	443.67	462.48	486.91	459.58	378.05	397.58	373.67	329.33	311.82	224.50	387.05	221.13	201.99	194.15	132.28
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	612.83	640.64	334.97	594.32	442.87	517.36	450.26	378.37	353.21	232.29	466.06	228.83	210.29	200.67	125.65
0.9	$\hat{\pi}_H^{(1)}$	165.26	185.14	192.89	192.00	186.24	152.61	179.24	179.74	179.02	162.91	156.23	156.85	158.27	157.96	133.82
	$\hat{\pi}_H^{(2)}$	177.19	194.09	197.22	196.07	187.75	161.83	185.02	183.23	182.39	163.91	165.67	159.30	160.50	160.15	134.28
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	325.84	479.10	510.39	496.23	422.42	262.07	389.86	375.98	368.70	270.52	251.56	240.55	234.26	231.45	162.24
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	405.29	662.40	383.20	677.40	519.09	308.19	477.51	454.65	443.38	292.67	282.56	253.41	245.37	241.29	159.27

Table 3(b) *Percent Relative Efficiency of* $\hat{\pi}_H^{(1)}$, $\hat{\pi}_H^{(2)}$ *and* $\hat{\pi}_H^{(h_1,h_2)}$.

$\pi_A = 0.1$																
	$(p_1, p_2) \rightarrow$	(0.7, 0.3)					(0.8, 0.2)					(0.9, 0.1)				
π_Y	$(n_1, n_2) \rightarrow$ Estimator ↓	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)	(5,5)	(25,5)	(20,20)	(20,30)	(50,50)
0.1	$\hat{\pi}_H^{(1)}$	195.96	163.08	145.21	140.75	112.11	183.98	125.67	119.53	117.43	94.67	174.77	104.88	108.66	108.74	91.33
	$\hat{\pi}_H^{(2)}$	217.52	168.38	146.96	142.17	112.25	200.31	126.87	120.21	118.00	94.56	187.54	104.97	108.84	108.33	91.10
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	470.05	273.36	202.07	189.56	108.64	333.70	136.21	126.86	123.48	72.31	263.23	89.49	97.86	97.19	61.09
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	561.56	285.51	437.72	190.30	100.59	360.79	128.04	119.63	116.44	62.65	269.88	79.67	88.63	88.04	50.27
0.3	$\hat{\pi}_H^{(1)}$	175.37	164.07	159.82	156.62	128.47	166.85	135.84	130.43	128.60	101.01	165.85	108.77	109.44	109.76	91.39
	$\hat{\pi}_H^{(2)}$	191.97	170.01	162.25	158.72	128.84	180.40	137.96	131.49	129.52	101.01	177.24	109.06	109.76	108.88	90.87
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	425.87	315.23	263.68	247.37	148.88	327.30	174.99	155.94	150.24	87.30	264.50	101.75	106.43	104.88	64.91
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	547.93	370.21	571.25	262.81	144.45	379.20	174.98	152.89	146.82	78.63	281.86	93.20	98.92	97.44	54.93
0.5	$\hat{\pi}_H^{(1)}$	165.90	166.41	166.77	164.24	137.94	154.30	141.41	138.33	136.76	111.35	154.76	112.36	114.25	115.05	92.63
	$\hat{\pi}_H^{(2)}$	179.49	172.98	169.55	166.73	138.47	165.29	144.09	139.72	138.04	106.97	164.19	112.79	114.66	114.10	91.91
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	369.11	327.93	300.77	285.42	176.50	291.80	195.44	179.16	173.61	99.86	246.11	110.68	115.09	113.51	69.02
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	473.86	393.17	641.29	314.62	175.90	344.90	200.66	179.78	173.11	91.92	268.55	103.05	108.40	106.74	59.50
0.9	$\hat{\pi}_H^{(1)}$	161.15	170.63	171.65	170.04	145.85	144.79	145.84	146.28	145.43	114.42	137.13	116.58	121.63	122.82	95.86
	$\hat{\pi}_H^{(2)}$	172.15	177.41	174.70	172.87	146.51	152.37	148.57	148.03	147.11	114.59	143.51	117.11	122.34	122.12	94.88
	$\hat{\pi}_H^{(h_1,h_2)=(5,6)}$	299.68	348.87	331.33	320.35	203.17	227.21	209.22	207.82	204.55	116.64	194.68	120.76	131.62	130.99	76.52
	$\hat{\pi}_H^{(h_1,h_2)=(8,9)}$	361.16	418.20	701.69	368.20	207.94	256.89	216.92	215.38	211.50	109.95	209.12	114.10	125.93	125.20	67.54

5 Conclusions

Several modified estimators of the population proportion π_A have been suggested. It has been shown that the suggested estimators are better than Greenberg et. al's (1969) estimator $\hat{\pi}_{As}$ under certain conditions. It is observed that the efficiency of the estimator $\hat{\pi}^{(h_1,h_2)}_{\hat{H}_{opt}}$ can be increased considerably for different choices of $(n_1, n_2, p_1, p_2, \pi_Y)$ through suitable selections of h_1 and h_2. The suggested estimators are recommended for their use in practice for moderately large/smaller sample sizes.

Acknowledgements

Authors are thankful to the referee for his valuable suggestions, which have improved the earlier draft of the paper considerably.

(Received February, 2004. Accepted February, 2006.)

References

Arnab, R. (1998). Randomized response surveys: Optimum estimation of a finite population total. *Statistical Papers*, **39**, 405–408.

Arnab, R. and Singh, S. (2002). Estimation of the size and mean value of a stigmatized characteristic of a hidden gang in a finite population: A unified approach. *Annals of the Institute of Statistical Mathematics*, **54**, 659–666.

Bellhouse, D. R. (1995). Estimation of correlation in randomized response. *Survey Methodology*, **21**, 13–19.

Bhargava, M. and Singh, R. (1999). On the relative efficiency of certain randomized response strategies. *Journal of the Indian Society of Agricultural Statistics*, **52**, 245–253.

Bhargava, M., Singh, R. and Rizvi, S. E. H. (2002). On comparison of some randomized response strategies in relation to privacy protection. *Calcutta Statistical Association Bulletin*, **53**, 117–128.

Chang, H. J. and Huang, K. C. (2001). Estimation of proportion and sensitivity of a quantitative character. *Metrika*, **53**, 269–280.

Chang, H. J. and Liang, D. H. (1996). A two stage unrelated randomized response procedure. *Australian Journal of Statistics*, **38**, 43–52.

Chaudhuri, A. (2001a). Using randomized response from a complex survey to estimate a sensitive proportion in a dichotomous finite population. *Journal of Statistical Planning and Inference*, **94**, 37–42.

Chaudhuri, A. (2001b). Estimating sensitive proportion from unequal probability samples using randomized responses. *Pakistan Journal of Statistics*, **17**, 259–270.

Chaudhuri, A. (2002). Estimating sensitive proportion from a randomized responses in unequal probability sampling. *Calcutta Statistical Association Bulletin*, **52**, 315–322.

Chaudhuri, A., Adhikary, A. K. and Maiti, T.(1998). A note on non-negative mean square error estimation of regression estimators in randomized response surveys. *Statistical Papers*, **39**, 409–415.

Chaudhuri, A. and Mukerjee, R. (1987). Randomized response techniques: A review. *Statistica Neerlandica*, **41**, 27–44.

Chaudhuri, A. and Mukerjee, R. (1988). *Randomized response: theory and techniques.* New York: Marcel Dekker.

Christofides, T. C. (2003). A generalized randomized response technique. *Metrika*, **57**, 195–200.

Fox, J. and Tracy, P. (1986). *Randomized response: a method for sensitive surveys.* Beverly Hills: Sage Publication.

Goodman, L. A. (1953). A simple method for improving some estimators. *Annals of Mathematical Statistics*, **24**, 114–127.

Greenberg, B. G., Abul-Ela, A. L. A., Simmons, W. R. and Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, **64**, 520–539.

Gupta, S., Gupta, B. and Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference*, **100**, 239–247.

Hedayat, A. S. and Sinha, B. K. (1991). *Design and Inference in Finite Population Sampling.* New York: Wiley.

Hirano, K. (1972). Using some approximately known coefficient of variation in estimating the mean. *Research Memorandom No. 49.* Institute of Statistical Mathematics, Tokyo, Japan.

Horvitz, D. G., Shah, B. V. and Simmons, W. R. (1967). The unrelated question randomized response model. *Proceedings of Social Statistics Section, American Statistical Association*, 65–72.

Javed, M. Gupta, J. P. and Singh, S. (2002). An improved randomized response strategy using unrelated question. *Aligarh Journal of Statistics*, **22**, 17–25.

Lee, G. S. and Hong, K. H. (1998). An improved unrelated question model. *Korean Journal of Applied Statistics*, **11**, 415–421.

Lemmer, H. H. (1981). Note on shrinkage estimators for the binomial distribution. *Communications in Statistics- Theory Method,* **A(10)10**, 1017–1027.

Mangat, N. S., Singh, S. (2002). A modified quantitative randomized response model. *Journal of Indian Society of Agricultural Statistics*, **55**, 209–219.

Mangat, N. S., Singh, R. and Singh, S. (1997). Violation of respondent's privacy in Moor's model-its rectification through a random group strategy. *Communications in Statistics-Theory Method*, **26(3)**, 743–754.

Mehta, J. S. and Srinivasan, R. (1971). Estimation of the mean by shrinkage to a point. *Journal of the American Statistical Association*, **66**, 86–90.

Padmawar, V. R. and Vijayan, K. (2000). Randomized response revisited. *Journal of Statistical Planning and Inference*, **90**, 293–304.

Searls, D. T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association*, **59**, 1225–1226.

Searls, D. T. (1967). A note on the use of an approximately known coefficient of variation. *The American Statistician*, **21**, 20–21.

Searls, D. T. and Intarapanich, P. (1990). A note on an estimator for the variance that utilizes the kurtosis. *The American Statistician*, **44**, 295–296.

Sheers, N. (1992). A review of randomized response technique. *Measurement and Evaluation in Counseling and Development*, **25**, 27–41.

Singh, S. (2002). A new stochastic randomized response model. *Metrika*, **56**, 131–142.

Singh, S., Horn, S. and Chowdhury, S. (1998). Estimation of stigmatized characteristics of a hidden gang in finite population. *Australian and New Zealand Journal of Statistics*, **40**, 291–298.

Singh, S. and Joarder, A. H. (1997). Unknown repeated trials in randomized response sampling. *Journal of the Indian Society of Agricultural Statistics*, **50**, 103–105.

Singh, S. and King, M. L. (1999). Estimation of the coefficient of determination using scrambled responses. *Journal of the Indian Society of Agricultural Statistics*, **52**, 338–343.

Singh, H. P. and Mathur, N. (2002a). An alternative to an improved randomized response strategy. *Statistics in Transition*, **5**, 873–886.

Singh, H. P. and Mathur, N. (2002b). A revisit to alternative estimators for randomized response technique. *Journal of the Indian Society of Agricultural Statistics*, **55**, 79–87.

Singh, H. P. and Mathur, N. (2002c). On alternative estimators in randomized response technique. *Journal of the Indian Society of Agricultural Statistics*, **55**, 189–196.

Singh, H. P. and Mathur, N. (2002d). On Mangat's improved randomized response strategy. *Statistica*, **LXII** , 397–403.

Singh, H. P. and Mathur, N. (2002e). An alternative randomized response technique using inverse sampling. *Calcutta Statistical Association Bulletin*, **53**, 233–244.

Singh, H. P. and Mathur, N. (2003a). Modified optional randomized response sampling. *Journal of Indian Society of Agricultural Statistics*, **56**, 199–206.

Singh, H. P. and Mathur, N. (2003b). An optionally randomized response technique. *Aligarh Journal of Statistics*, **23**, 1–5.

Singh, H. P. and Mathur, N. (2004). Unknown repeated trials in the unrelated question randomized response model. *Biometrical Journal*, **46**, 375–378.

Singh, H. P. and Shukla, S. K. (2002). A class of shrinkage estimators for the variance of exponential distribution with type-I censoring. *IAPQR Transitions*, **27**, 119–141.

Singh, S. Singh, R. and Mangat, N.S. (2000). Some alternative strategies to moor's model in randomized response sampling. *Journal of Statistical Planning and Inference*, **83**, 243–255.

Singh, S. and Tracy, D. S. (1999). Ridge regression using scrambled responses. *Metron*, **57**, 147–157.

Smith, N. F. and Street, D. J. (2003). The use of balanced incomplete block designs in designing randomized response surveys. *Australian and New Zealand Journal of Statistics*, **45**, 18–29.

Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association*, **63**, 113–123.

Tracy, D. S. and Mangat, N. S. (1996). Some developments in randomized response sampling during the last decade-A follow up of review by Chaudhuri and Mukerjee. *Journal of Applied Statistical Science*, **4**, 147–158.

Tracy, D. S. and Mangat, N. S. (1998). Comparsion of distinct units based estimators in unrelated questions randomized response model. *International Journal of Mathematical Statistical Science*, **7**, 229–240.

Tracy, D. S. and Mangat, N. S. (1999). A modified quantitative randomized response model. *Pakistan Journal of Statistics*, **15**, 19–26.

Tracy, D. S. and Oshan, S. S.(1999). An improved randomized response technique. *Pakistan Journal of Statistics*, **15(1)**, 1–6.

Tracy, D. S. and Singh, S. (1999). Calibration estimators for randomized response sampling. *Metron*, **57**, 47–68.

Van Der Heijden P. G. M., Van Gils G., Bouts J. and Hox, J. (1998). A comparison of randomized response, CASAQ, and direct questioning, eliciting sensitive information in the context of social security fraud. *Kwantiatieve Methoden*, **59**, 15–34.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, **60**, 63–69.

Housila P. Singh and **Nidhi Mathur**
School of Studies in Statistics
Vikram University
Ujjain-456010, India
E-mails: nidhimathur_29@yahoo.co.in and hpsujn@rediffmail.com