Jour. Ind. Soc. Ag. Statistics
50(1), 1997 : 103–105

Unknown Repeated Trials in Randomized Response Sampling

Sarjinder Singh and Anwar H. Joarder
Monash University, Clayton, Victoria, Australia 3168
(Received : July, 1996)

SUMMARY

In this paper, an alternative randomization device for eliciting information on the sensitive issues is proposed. The estimator based on the proposed strategy is found to be unbiased for population proportion and is always more efficient than the Warner's [2] usual estimator.

Key words: Randomized response technique, Warner's model, Estimation of proportion.

1. *Introduction*

In sample surveys, it is often difficult to elicit truthful information from the respondents regarding the stigmatized characters. Warner [2] suggested an ingenious method of collecting information on sensitive characters. According to the method, for estimating the population proportion π possessing the sensitive character A, a simple random sample of n persons is drawn with replacement from the population. Each interviewee in the sample is furnished with an identical randomization device where the outcome 'I possesses character A' occurs with probability p while its compliment 'I does not possess character A' occurs with probability $(1 - p)$. The respondent answers 'yes' if the outcome of the randomization device tallies with his/her actual status otherwise he/she answers 'no'. An unbiased estimator of π, the proportion of population belonging to the sensitive group A, considered by Warner is given by

$$\hat{\pi} = \frac{(n'/n - 1 + p)}{2p - 1}, \; p \neq 0.5 \tag{1.1}$$

with variance

$$V(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1)^2} \tag{1.2}$$

where n' is the number of "yes" responses from the sample of n individuals.

Subsequently many other researchers have modified and suggested alternative randomized response procedures applicable to different situations and are cited by Singh [1]. The present paper is an attempt to present an interesting modification of Warner's [2] model.

2. *Proposed Strategy*

In this method, if a respondent belongs to group A, then he/she is requested to repeat the trial in the Warner's [2] randomization device if in the first trial he/she does not get the statement according to his/her status. The rest of the procedure remains the same. The repetition of the trial is known to the interviewee but remains unknown to the interviewer. Thus the model may be called unknown repeated trial model. Assuming completely truthful reporting by the respondents, the probability of "yes" answer is given by

$$\theta_1 = \pi[p + (1-p)p] + (1-\pi)(1-p) \tag{2.1}$$

Then we consider the estimator of π as follows:

$$\hat{\pi}_s = \frac{\hat{\theta}_1 - (1-p)}{2p - 1 + p(1-p)} \tag{2.2}$$

where $\hat{\theta}_1$ is the sample proportion of "yes" responses in the proposed procedure. Since $\hat{\theta}_1$ is a binomial variable with parameters (n, θ_1), we have the following theorems:

Theorem 2.1: The estimator $\hat{\pi}_s$ is unbiased for population proportion π.

The proof being straightforward is omitted.

Theorem 2.2: The variance of the estimator $\hat{\pi}_s$ is given by

$$V(\hat{\pi}_s) = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1+p(1-p))^2} - \frac{\pi p(1-p)}{n(2p-1+p(1-p))} \tag{2.3}$$

Proof: We have

$$V(\hat{\pi}_s) = \frac{\theta_1(1-\theta_1)}{n(2p-1+p(1-p))^2} \tag{2.4}$$

On using the value of θ_1 from (2.1), the above expression reduces to one given in (2.3). Hence the theorem.

Theorem 2.3: An unbiased estimator of variance $V(\hat{\pi}_s)$ is given by

$$\upsilon(\hat{\pi}_s) = \frac{\hat{\theta}_1(1-\hat{\theta}_1)}{(n-1)[2p-1+p(1-p)]^2} \tag{2.5}$$

We are now in a position to look into the efficiency aspect of the proposed estimator $\hat{\pi}_s$ with Warner's usual estimator $\hat{\pi}$. The efficiency of the proposed estimator $\hat{\pi}_s$ relative to $\hat{\pi}$ is defined as Percent Relative Efficiency PRE $= V(\hat{\pi}) \times 100/V(\hat{\pi}_s)$. Thus the estimator $\hat{\pi}_s$ is more efficient if

$$\text{PRE} \geq 100 \tag{2.6}$$

On using (1.2) and (2.3), one can easily see that inequality (2.6) holds good for $p > 0.5$. Thus the proposed estimator $\hat{\pi}_s$ remains always more efficient than the Warner's [2] model. In order to have an idea regarding the magnitude of PRE, we resort to an empirical investigation. Table 1 shows that the PRE varies from 118.64 to 726.07 for practicable choices of p.

Table 1 : Percent Relative Efficiency (PRE) of the proposed estimator $\hat{\pi}_s$ relative to $\hat{\pi}$

π	p			
	0.6	0.7	0.8	0.9
0.1	477.60	226.23	154.48	119.18
0.2	477.31	224.63	153.06	118.64
0.3	482.88	226.87	158.40	119.55
0.4	494.65	232.88	158.14	121.54
0.5	513.58	243.29	164.66	124.78
0.6	541.48	259.70	175.19	129.93
0.7	581.54	285.44	192.69	138.68
0.8	639.46	372.98	225.05	155.99
0.9	726.07	407.08	301.04	204.80

REFERENCES

[1] Singh, Sarjinder, 1994. Unrelated question randomized response sampling using continuous distributions. *J. Indian Soc. Agric. Statist.*, **46**, 349-361.

[2] Warner, S.L., 1965. Randomized response: A survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.*, **60**, 63-69.