An Improved Randomized Response Strategy

Author(s): N. S. Mangat

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 56, No. 1 (1994), pp. 93-95
Published by: Wiley for the Royal Statistical Society
Stable URL: http://www.jstor.org/stable/2346030
Accessed: 29/01/2014 23:02

J. R. Statist. Soc. B (1994)
56, *No.* 1, *pp.* 93–95

An Improved Randomized Response Strategy

By N. S. MANGAT†

Punjab Agricultural University, Ludhiana, India

[Received July 1991. Final revision August 1992]

SUMMARY

A randomized response procedure is proposed which has the benefit of simplicity over that of Mangat and Singh. Conditions are obtained under which the proposed strategy is better than those of Warner or Mangat and Singh.

Keywords: EQUAL PROBABILITIES WITH REPLACEMENT; ESTIMATION OF PROPORTION; RANDOMIZED RESPONSE TECHNIQUE

1. INTRODUCTION

To procure trustworthy data for estimating π, the proportion of a population having a sensitive attribute A say, Warner (1965) introduced a procedure called the randomized response (RR) technique. Subsequently, several other workers have suggested various alternative RR strategies (see Hedayat and Sinha (1991) for a review). Recently, Mangat and Singh (1990) proposed a two-stage RR procedure. Their method, requiring the use of two randomized devices, makes the interview procedure a little cumbersome. This motivated the author to consider a relatively simple RR strategy. The situation where the respondents are not completely truthful in their answers has also been considered.

2. METHOD PROPOSED

Each of n respondents, assumed to be selected by equal probabilities with replacement sampling, is instructed to say 'yes' if he or she has the attribute A. If he or she does not have attribute A, the respondent is required to use the Warner randomization device consisting of two statements:

(a) 'I belong to attribute A' and
(b) 'I do not have attribute A',

represented with probabilities p and $1-p$ respectively. Then he or she is to report 'yes' or 'no' according to the outcome of this randomization device and the actual status that he or she has with respect to attribute A. The whole procedure is completed by the respondent unobserved by the interviewer. The probability of a yes answer for this procedure is given by

$$\alpha = \pi + (1-\pi)(1-p).$$

†*Address for correspondence*: Department of Plant Breeding, Punjab Agricultural University, Ludhiana 141004, India.

 0035–9246/94/56093

Since the yes answer may come from respondents in both group A and group not-A, the confidentiality of the person reporting yes will not be violated. However, the individuals reporting 'no' will always come from group not-A. Since it is assumed that belonging to group not-A does not carry any perceived stigma, the respondents are assumed not to object to the procedure.

The maximum likelihood estimator of π for the proposed procedure can be easily seen to be

$$\hat{\pi}_{\mathrm{m}} = (\hat{\alpha} - 1 + p)/p,$$

where $\hat{\alpha}$ is the observed proportion of yes answers obtained from the n sampled individuals.

Since $\hat{\alpha}$ follows the binomial distribution $B(n, \alpha)$, the estimator $\hat{\pi}_{\mathrm{m}}$ is unbiased for π. This leads simply to the following result.

Theorem 1. The variance $V(\hat{\pi}_{\mathrm{m}})$ and its unbiased estimator $v(\hat{\pi}_{\mathrm{m}})$ are respectively given by

$$V(\hat{\pi}_{\mathrm{m}}) = \pi(1-\pi)/n + (1-\pi)(1-p)/np \tag{2.1}$$

and

$$v(\hat{\pi}_{\mathrm{m}}) = \hat{\alpha}(1-\hat{\alpha})/(n-1)p^2.$$

2.1. *Efficiency Comparison*

On using equation (2.3) of Mangat and Singh (1990) and equation (2.1) of this paper, we arrive at a result stated in the following theorem.

Theorem 2. The proposed estimator $\hat{\pi}_{\mathrm{m}}$ will be more efficient than the Mangat and Singh (1990) estimator if

$$\pi > 1 - \frac{p(1-T)\{1-(1-T)(1-p)\}}{\{2p-1+2T(1-p)\}^2}, \qquad p \neq 0.5, \tag{2.2}$$

where T is the probability representing the sensitive attribute in their first device and p is as above. Once the investigator has decided the values of p and T, it is easy to verify inequality (2.2) by using a prior guess about the value of π from an earlier study or pilot survey. On putting $T=0$, for which Mangat and Singh's (1990) result reduces to Warner's (1965) result, in inequality (2.2), we obtain the following.

Theorem 3. The proposed strategy is more efficient than that of Warner (1965) if

$$\pi > 1 - \{p/(2p-1)\}^2,$$

which always holds for $p > \frac{1}{3}$.

3. INCOMPLETELY TRUTHFUL REPORTING

Let T_3 be the probability that the respondents belonging to the sensitive class report the truth. The respondents in the non-sensitive group have no reason to tell a lie. The probability of yes answers for the proposed procedure thus becomes

$$\beta = \pi T_3 + (1-\pi)(1-p).$$

Here, the estimator $\hat{\pi}_{\mathrm{m}}$ is biased with magnitude of bias given by

$$B(\hat{\pi}_{\mathrm{m}}) = \pi(T_3 - 1)/p.$$

Then we have the following theorem for which the proof is straightforward.

Theorem 4. The mean-square error of the estimator $\hat{\pi}_{\mathrm{m}}$ is given by

$$\mathrm{MSE}(\hat{\pi}_{\mathrm{m}}) = \left[\frac{\pi T_3(1-\pi T_3)}{n} + \frac{(1-\pi)(1-p)\{1-(1-\pi)(1-p)-2\pi T_3\}}{n} + \{\pi(T_3-1)\}^2\right] \Big/ p^2. \tag{3.1}$$

This expression can be compared with expressions (1.3) and (3.3) of Mangat and Singh (1990) for an efficiency comparison based on the minimum mean-square error.

ACKNOWLEDGEMENTS

The author is grateful to the referee and Professor Ravindra Singh, for their valuable comments and suggestions.

REFERENCES

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

Mangat, N. S. and Singh, R. (1990) An alternative randomized response procedure. *Biometrika*, **77**, 439–442.

Warner, S. L. (1965) Randomized response: a survey technique for eliminating evasive answer bias. *J. Am. Statist. Ass.*, **60**, 63–69.