AN ALTERNATIVE RANDOMIZED RESPONSE MODEL USING TWO DECK OF CARDS: A REJOINDER

Raghunath Arnab
Department of Statistics, University of Botswana, Botswana; and School of Statistics & Actuarial Sciences, University of Kwa-Zulu Natal, RSA

Sarjinder Singh
Department of Mathematics, Texas A&M University-Kingsville, USA

1. INTRODUCTION

While collecting information, directly from respondents, relating to sensitive issues such as induced abortion, drug addiction, duration of suffering from Aids and so on, the respondents very often report untrue values or even refuse to respond. Warner (1965) introduced an ingenious technique known as randomized response technique (RR) for estimating \( \pi \), the proportion of population possessing certain stigmatized character \( A \) (say) by protecting the privacy of respondents and preventing the unacceptable rate of non-response. Warner’s (1965) technique was modified by Horvitz et al. (1967), Greenberg et al. (1969), Raghavrao (1978), Kim (1978), Franklin (1989), Mangat and Singh (1990), Kuk (1990), Mangat and Singh (1991, 1992), Singh (1993), Kervliet (1994), Singh et al. (1994), Mahajan et al. (1994), Bhargava and Singh (2000), Singh et al. (2000), Singh and Mathur (2002, 2004), Arnab (2004), Espejo and Singh (2004), Sidhu and Bansal (2005/06), Pal and Sonali (2005/06), Javed and Grewal (2005/06), Zhimin (2005/06), Ryu et al. (2005/06), Zaizai (2005/06) and Singh (2010) among other researchers for improving greater co-operation and efficiency. One could also refer to Giordano and Perri (2011).

Recently Odumade and Singh (2009) proposed two decks of cards for estimating the population proportion \( \pi \) for RR surveys under SRSWR sampling. They showed on the basis of empirical investigation their model fares better than Warner (1965) and Mangat and Singh (1990) models if the value of \( \pi \to 0 \) or \( \pi \to 1 \). Arnab et al. (2012) extended Odumade and Singh (2009) RR technique for complex survey designs and wider classes of estimators while Abdelfatah et al. (2011) modified Odumade and Singh (2009) RR procedure under SRWR sampling method by using unrelated question. Abdelfatah et al. (2011) proved that the proposed method provides more efficient method of estimation \( \pi \). In this paper we have shown that Abdelfatah et al. (2011) RR strategy (combination of estimator and sampling design) cannot provide efficient estimator than Warner (1965) model. For more clarity we describe Odumade and Singh (2009) and Abdelfatah et al. (2011) RR as follows:

1.1. Odumade and Singh (2009)

Odumade and Singh (2009) selected a sample of size \( n \) by SRSWR method. Each of the selected respondents in the sample is asked to select two cards, one card from the Deck-I and the
other from the Deck-II. Each of the decks consists of two types of cards as in Warner (1965) model. The proportion of cards written “I belong to the sensitive group A” in the Deck-I and Deck-II are \( P \) and \( T \) respectively. The respondents are asked to report his/her response as \((X, Y)\) where \( X \) indicates response from the card selected from the Deck-I while \( Y \) indicates response from the card selected from the Deck-II. For example if a respondent selects a card written “I belong to the sensitive group A” from the Deck-I and selects the other card written “I do not belong to the sensitive group A” from the Deck-II, then he/she will supply with a response “Yes, No” if he/she belong to the sensitive group A. On the other hand if the respondent do not belongs to the group A, he/she will supply “No, Yes” as his/her response. Let out of the \( n \) responses \( n_{11}, n_{10}, n_{01} \) and \( n_{00} \) denote respectively the frequencies of the responses (Yes, Yes), (Yes, No), (No, Yes) and (No, No). An unbiased estimator for the population proportion \( \pi \) was obtained by Odumade and Singh (2009) as

\[
\hat{\pi}_{uw} = \frac{1}{2} + \frac{(P + T - 1)(n_{11} - n_{00})(P - T)(n_{10} - n_{00})}{2n([P + T - 1]^2 + (P - T)^2)}
\] (1)

The variance of \( \hat{\pi}_{uw} \) and an unbiased estimator of the variance of \( \hat{\pi}_{uw} \) were obtained respectively as

\[
V(\hat{\pi}_{uw}) = \frac{(P + T - 1)^2\{PT + (1 - P)(1 - T)\} + (P - T)^2\{T(1 - P) + P(1 - T)\}}{4n([P + T - 1]^2 + (P - T)^2]^2} - \frac{(2\pi - 1)^2}{4n} - (2\hat{\pi} - 1)^2
\] (2)

and

\[
\hat{V}(\hat{\pi}_{uw}) = \frac{1}{4(n - 1)}\left[\frac{(P + T - 1)^2\{PT + (1 - P)(1 - T)\} + (P - T)^2\{T(1 - P) + P(1 - T)\}}{[([P + T - 1]^2 + (P - T)^2]^2}
\] (3)

1.2. Abdelfatah (2011)

Under this RR technique each respondent is asked to draw two cards; one from the “Deck(I)” and another from “Deck(II)”. Deck (I) comprises two types of cards as in Warner (1965) viz. “I belong to the sensitive group A” with proportion \( W \) and “ I do not belong to the sensitive group A” with proportion \( 1 - W \). The respondent should answer truthfully “Yes” or “No”.

The deck two comprises also two types of cards written “YES” with proportion \( Q \) and “NO” with proportion \( 1 - Q \). Regardless of his/her actual status, the respondent have to answer “YES” if he /she receives card written “YES”. Alternatively, if the respondent receives the card written “NO” the respondent should answer “NO” as his or her response.

<table>
<thead>
<tr>
<th>Deck (I)</th>
<th>Deck (II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I \in A ) with proportion ( W )</td>
<td>“YES” with proportion ( Q )</td>
</tr>
<tr>
<td>( I \in A^c ) with proportion ( W )</td>
<td>“NO” with proportion ( 1 - Q )</td>
</tr>
</tbody>
</table>
Let the responses of the selected sample of \( n \) units by SRSWR method be classified as follows:

<table>
<thead>
<tr>
<th>Response from Deck (I)</th>
<th>Response from Deck (II)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>( n_{11} )</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>( n_{01} )</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>( n_{10} )</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>( n_{00} )</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>( n )</td>
</tr>
</tbody>
</table>

Abdelfatah et al. (2011) derived the following results:

(i) An unbiased estimator of the population \( \pi \) is

\[
\hat{\pi}_f = \frac{1}{2} + \frac{Q(n_{11} / n - n_{11} / n) + (1 - Q)(n_{01} / n - n_{01} / n)}{2(2W - 1)[Q^2 + (1 - Q)^2]}, \quad W \neq 0.5
\]

(ii) The variance of \( \hat{\pi}_f \) is

\[
V(\hat{\pi}_f) = \frac{Q^3 + (1 - Q)^3}{4n(2W - 1)^2[Q^2 + (1 - Q)^2]^2} - \frac{(2\pi - 1)^2}{4n}, \quad W \neq 0.5
\]

(iii) An unbiased estimator of the variance of \( \hat{\pi}_f \) is

\[
\hat{V}(\hat{\pi}_f) = \frac{1}{4(n - 1)} \left[ \frac{Q^3 + (1 - Q)^3}{(2W - 1)^2[Q^2 + (1 - Q)^2]^2} - (2\hat{\pi}_f - 1)^2 \right]
\]

2. PROPOSED IMPROVED ESTIMATOR

Suppose that from a finite population \( U = \{U_1, \ldots, U_N\} \) of \( N \) identifiable units a sample \( s \) of size \( n \) is selected by SRSWR method. Then each of the selected respondent in the sample \( s \) is asked to perform randomizes device suggested by Abdelfatah et al. (2011). The respondents will provide one of the answers (Yes, Yes), (Yes, No), (No, Yes) and (No, No) as his/her randomize response. Let us denote

\[
x_{11}(i) = \begin{cases} 1 & \text{if the answer from the } i\text{th unit is } (\text{Yes, Yes}) \\ 0 & \text{otherwise} \end{cases}
\]

\[
x_{10}(i) = \begin{cases} 1 & \text{if the answer from the } i\text{th unit is } (\text{Yes, No}) \\ 0 & \text{otherwise} \end{cases}
\]

\[
x_{01}(i) = \begin{cases} 1 & \text{if the answer from the } i\text{th unit is } (\text{No, Yes}) \\ 0 & \text{otherwise} \end{cases}
\]

\[
x_{00}(i) = \begin{cases} 1 & \text{if the answer from the } i\text{th unit is } (\text{No, No}) \\ 0 & \text{otherwise} \end{cases}
\]
\[
x_{00}(i) = \begin{cases} 
1 & \text{if the answer from the ith unit is (No, No)} \\
0 & \text{otherwise}
\end{cases} \quad (7)
\]

Here the variable under study \( y \) is an indicator variable and \( y_i \) the value of \( y \) obtained from the ith unit can be written as
\[
y_i = \begin{cases} 
1 & \text{if the ith unit } A \\
0 & \text{if the ith unit } \notin A
\end{cases} \quad (8)
\]

For the Abdelfatah RR technique described above, we have

- Probability of getting "yes" answer from card type I = \( W y_i + (1 - W)(1 - y_i) \)
- Probability of getting "no" answer from card type I = \( (1 - W)y_i + W(1 - y_i) \)
- Probability of getting "yes" answer from card type II = \( Q \)
- Probability of getting "no" answer from card type II = \( 1 - Q \)

Hence,

\[
\begin{align*}
\theta_{11} &= \text{Probability of getting "yes, yes" answer} = \{W y_i + (1 - W)(1 - y_i)\}Q \\
\theta_{10} &= \text{Probability of getting "yes, no" answer} = \{W y_i + (1 - W)(1 - y_i)\}(1 - Q) \\
\theta_{01} &= \text{Probability of getting "no, yes" answer} = \{(1 - W)y_i + W(1 - y_i)\}Q \\
\theta_{00} &= \text{Probability of getting "no, no" answer} = \{(1 - W)y_i + W(1 - y_i)\}(1 - Q)
\end{align*}
\]

The likelihood function of getting \( x_{11} \) "yes, yes"; \( x_{10} \) "yes, no" answer; \( x_{01} \) "no, yes" and \( x_{00} \) "No, No" answers is
\[
L = \{W y_i + (1 - W)(1 - y_i)\}^{x_{11}} \times \{(1 - W)y_i + W(1 - y_i)\}^{x_{10}} \\
\times \{(1 - W)y_i + W(1 - y_i)\}^{x_{01}} \times \{(1 - W)y_i + W(1 - y_i)\}^{x_{00}} \quad (9)
\]

Note that \( x_{11}, x_{10}, x_{01} \) and \( x_{00} \) are binary and subject to \( x_{11} + x_{10} + x_{01} + x_{00} = 1 \).

Differentiating the likelihood with respect to \( y_i \) we get
\[
\frac{\partial L}{\partial y_i} = \frac{x_{11}(2W - 1)}{\{W y_i + (1 - W)(1 - y_i)\}} + \frac{x_{10}(2W - 1)}{\{W y_i + (1 - W)(1 - y_i)\}} \\
- \frac{x_{01}(2W - 1)}{\{(1 - W)y_i + W(1 - y_i)\}} - \frac{x_{00}(2W - 1)}{\{(1 - W)y_i + W(1 - y_i)\}}
\]
An alternative randomized response model using two deck of cards: a rejoinder

\[ \frac{x_{i*}}{Wy_i + (1 - W)(1 - y_i)} - \frac{x_{0*}}{(1 - W)y_i + W(1 - y_i)} \]  

where \( x_{i*} = x_{11} + x_{10} \) and \( x_{0*} = x_{21} + x_{22} \)

\( \frac{\partial L}{\partial y_i} = 0 \) implies

\[ x_{i*} \left\{ (1 - W)y_i + W(1 - y_i) \right\} - x_{0*} \left\{ Wy_i + (1 - W)(1 - y_i) \right\} = 0 \]  

The equation (11) yields

\( (1 - 2W)y_i + W(x_{1*} + x_{2*}) - x_{0*} = 0 \)

i.e. MLE of \( y_i \) is

\[ \hat{y}_i = \frac{W - x_{0*}}{2W - 1} \]  

Denoting \( E_R \) and \( V_R \) as expectation and variance over the RR technique we have

**Theorem 1.**

(i) \( E_R(\hat{y}_i) = \frac{W - E_R(x_{0*})}{2W - 1} \)

(ii) \( V_R(\hat{y}_i) = \frac{W(1 - W)}{(2W - 1)^2} = \varphi \)

**Proof.**

(i) \( E_R(\hat{y}_i) = \frac{W - E_R(x_{0*})}{2W - 1} \)

\[ = \frac{W - (\theta_{01} + \theta_{00})}{2W - 1} \]

\[ = y_i \]

(ii) \( V_R(\hat{y}_i) = \frac{V_R(x_{0*})}{(2W - 1)^2} \)

\[ = V_R(x_{01}) + V_R(x_{02}) + 2C_R(x_{01}, x_{02}) \]  

Now
\[ V_R(x_{01}) = \theta_{01} - \theta_{01}^2 \]
\[ = \{(1 - W)y_i + W(1 - y_i)\}Q - \{(1 - W)y_i + W(1 - y_i)\}Q^2 \]
\[ = \{(1 - W)y_i + W(1 - y_i)\}Q - \{(1 - W)^2y_i + W^2(1 - y_i)\}Q^2 \]
\[ (\text{Since } y_i \text{ can take values 0 or 1 with positive probabilities}) \]
\[ = Q(1 - Q)(1 - 2W)y_i + WQ(1 - WQ) \quad (14) \]

\[ V_R(x_{00}) = \theta_{00} - \theta_{00}^2 \]
\[ = Q(1 - Q)(1 - 2W)y_i + W(1 - Q)\{1 - W(1 - Q)\} \quad (15) \]

And
\[ C_R(x_{01}, x_{00}) = -\theta_{01} \theta_{00} \]
\[ = -\{(1 - W)y_i + W(1 - y_i)\}^2Q(1 - Q) \]
\[ = -Q(1 - Q)\{(1 - 2W)y_i + W^2\} \quad (16) \]

Finally substituting (14), (15) and (16) in (13) we get
\[ R_R(\hat{\gamma}_i) = \frac{V_R(x_{01}) + V_R(x_{00}) + 2C_R(x_{01}, x_{00})}{(2W - 1)^2} \]
\[ = \frac{WQ(1 - WQ) + W(1 - Q)\{1 - W(1 - Q)\} - 2Q(1 - Q)W^2}{(2W - 1)^2} \]
\[ = \frac{W(1 - W)}{(2W - 1)^2} \quad (17) \]

**Remark 1.** It is important the MLE of \( y_i \) is independent of \( Q \) that is the force response from the card type II has no role in estimating \( y_i \).

**Theorem 2.** For SRSWR

(i) \( \hat{\pi} = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i = \frac{W - \hat{p}_{01}}{2W - 1} \) is an unbiased estimator of \( \pi \)

(ii) Variance of \( \hat{\pi} \) is
An alternative randomized response model using two deck of cards: a rejoinder

\[ V(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{W(1-W)}{n(2W-1)^2} \]

(iii) An unbiased estimator of is

\[ \hat{V}(\hat{\pi}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} (\hat{\gamma}_i - \hat{\pi})^2 \]

where \( \sum_{i=1}^{n} \) denote sum over units in \( s \) with repetition and \( \hat{p}_{01} = \sum_{i=1}^{n} x_{0i} / n \).

**PROOF.**

(i) \( E(\hat{\pi}) = E_p [E_R(\hat{\pi})] \)

(\( \text{where } E_p \text{ denotes expectation over sampling design } p \text{ (SRSWR)} \))

\[ = E_p \left[ \frac{1}{n} \sum_{i=ia} E_R(\hat{\gamma}_i) \right] \]

\[ = \frac{1}{N} \sum_{i=ia} \gamma_i \]

\[ = \pi \]

(ii) \( V(\hat{\pi}) = E[V_R(\hat{\pi})] + V_p [E_R(\hat{\pi})] \)

(\( \text{where } V_p \text{ is the variance with respect to the sampling design } p \) )

\[ = \frac{1}{n^2} \sum_{i=ia} V_R(\hat{\gamma}_i) + V_p \left( \frac{1}{n} \sum_{i=ia} \gamma_i \right) \]

\[ = \frac{\phi}{n} + \frac{1}{N} \sum_{i=1}^{N} (\gamma_i - \overline{Y})^2 \text{ where } \overline{Y} = Y / N \]

\[ = \frac{\pi(1-\pi)}{n} + \frac{W(1-W)}{n(2W-1)^2} \]

(iii) \( E[\hat{V}(\hat{\pi})] = V(\hat{\pi}) \)

The following Theorem 3 shows that the proposed estimator \( \hat{\pi} \) is always superior to the estimator \( \hat{\pi}_f \) proposed by Abdelfatah et al. (2011).
THEOREM 3.

\[ V(\hat{\pi}) \leq V(\hat{\pi}_f) \text{, equality holds if } Q = \frac{1}{2} \]

PROOF.

Consider the difference

\[
\frac{Q^3 + (1 - Q)^3}{[Q^2 + (1 - Q)^2]^2} - 1
\]

\[
= \frac{Q^3 + (1 - Q)^3 - [Q^2 + (1 - Q)^2]^2}{[Q^2 + (1 - Q)^2]^2}
\]

Now

\[
Q^3 + (1 - Q)^3 - [Q^2 + (1 - Q)^2]^2
\]

\[
= Q^3 + (1 - Q)^3 - \{Q^4 + 2Q^2(1 - Q)^2 + (1 - Q)^4\}
\]

\[
= Q(1 - Q)\{4Q^2 + 1 - 4Q\}
\]

\[
= Q(1 - Q)\{4Q^2 + 1 - 4Q\} \quad (18)
\]

\[
V(\hat{\pi}) = \frac{Q^3 + (1 - Q)^3}{4n(2W - 1)^2[Q^3 + (1 - Q)^3]^2} - \frac{(2\pi - 1)^2}{4n}
\]

\[
\geq \frac{1}{4n(2W - 1)^2} - \frac{(2\pi - 1)^2}{4n} \quad \text{(using (18))}
\]

\[
= \frac{1}{4n(2W - 1)^2} - \frac{1}{4n} + \frac{1}{4n} - \frac{(2\pi - 1)^2}{4n}
\]

\[
= \frac{W(1 - W)}{4n} + \frac{\pi(1 - \pi)}{4n}
\]

\[
= V(\hat{\pi})
\]

ACKNOWLEDGEMENT

The authors would like to thank the Editor, a referee and Giovanna Galatà Executive Editor for their valuable suggestions.
REFERENCES


SUMMARY

The Randomized response (RR) technique with two decks of cards proposed by Odumade and Singh (2009) can always be made more efficient than the RR techniques proposed by Warner (1965), Mangat and Singh (1990), and Mangat (1994) by adjusting the proportion of cards in the decks. Abdelfatah et al. (2011) modified Odumade and Singh (2009) RR technique and claimed that their method can be more efficient than the Warner (1965) model. In this paper it is shown that such claim is not valid and the RR technique proposed by Abdelfatah et al. (2011) is in fact less efficient than the Warner (1965) technique at equal protection of respondents. Such finding are recently shown by Giordano and Perri (2011).