



Received: 21 October 2016
Accepted: 11 April 2017
First Published: 18 April 2017

*Corresponding author: Olusegun S. Ewemooje, Department of Statistics, Federal University of Technology, Akure, Nigeria
E-mails: osewemooje@futa.edu.ng, sunsege@yahoo.com

Reviewing editor:
Guohua Zou, Chinese Academy of Sciences, China

Additional information is available at the end of the article

STATISTICS | RESEARCH ARTICLE

Estimating two sensitive characters with equal probabilities of protection

Olusegun S. Ewemooje^{1*}

Abstract: This work seeks to improve the procedure for estimating two sensitive characteristics at a time by introducing equal probabilities of protection on the randomized devices. Equal protection estimators have been proposed. It was discovered that equal probabilities of protection increases the efficiency. The proposed model and estimators have been shown to be more robust and efficient than the conventional ones.

Subjects: Science; Mathematics & Statistics; Social Science; Behavioral Sciences

Keywords: equal protection model; percentage relative efficiency; robust estimator; sensitive characteristics

1. Introduction

Sample Survey research is one of the most essential areas of measurement in applied social sciences research. This covers any measurement techniques that involve asking questions from respondents. Beam (2012) in his book *The Problem with Survey Research* argues that all survey research instruments produce unreliable and potentially inaccurate results. Also, non-response or false response in human survey involving sensitive characters contributes great bias to obtaining correct estimate of the population parameter. For various reasons respondents in a sample survey may prefer not to disclose to the interviewer the correct answers to some questions. Such questions may be sensitive or highly personal, this situation may likely lead to refusal of answers or ambiguous answer. The resulting ambiguous answer bias is difficult to measure. This makes Randomized Response Technique (RRT) appropriate for such survey. RRT is a research method used in structured survey interview; it was first introduced by Warner (1965) and modified by Greenberg, Abul-Ela, Simmons, and Horvitz (1969), Mangat and Singh (1990), Mangat (1994), Odumade and Singh (2009), Adebola and Johnson (2015) and Adebola, Adediran, and Ewemooje (2017). The successes of the methods depend on the



Olusegun S. Ewemooje

ABOUT THE AUTHOR

Olusegun S. Ewemooje obtained Professional Diploma in Statistics, BSc and MSc in Statistics from University of Ibadan, Nigeria. He is currently on his PhD in Statistics at the same University. He is also a Graduate Statistician of the Royal Statistical Society. He is a lecturer at the Department of Statistics, Federal University of Technology, Akure, Nigeria. His research interest includes Sample Survey Designs and Applications, Education Statistics, Modelling and Environmental Statistics.

PUBLIC INTEREST STATEMENT

The proposed model is to improve the procedure for estimating two sensitive characteristics at a time by introducing equal probabilities of protection on the randomized devices. The proposed model and estimators have been shown to be more robust and efficient than the conventional ones. Therefore, with the use of proposed model social science researchers and those in behavioural science have little or no problem in surveying sensitive issues such as induced abortion, drug abuse, tax evasion, illegal possession of arms and all embarrassing or illegal activities.

respondent being persuaded that by participating, his or her personal status with regards to the sensitive issue will not be exposed.

However, a few work has been done on estimating two sensitive characteristics at the same time, these include: Lee, Sedory, and Singh (2013) which extended Warner’s (1965) procedure to capture two characteristics, Ewemooje and Amahia (2015, 2016) used Mangat’s (1994) procedure while Batool and Shabbir (2015) used Mangat and Singh (1990) procedure.

2. Lee et al. (2013) procedures

2.1. The simple model

A simple random and with replacement sample of n respondents from the given population was selected. Two shuffled decks of cards are presented to each respondent in the sample. The two decks are marked: Deck I and Deck II and each deck is comprised of two types of cards that indicate whether or not the respondent possesses a certain sensitive characteristic. Each respondent is requested to draw one card from each deck of cards, match his or her status with the statement on the card drawn from Deck I and Deck II, and then response in terms of “Yes or No” without revealing the statement written on the card to the interviewer. The probabilities of (Yes, Yes), (Yes, No), (No, Yes), and (No, No) are denoted by θ_{11} , θ_{10} , θ_{01} and θ_{00} respectively as follows:

$$\theta_{11} = (2P - 1)(2T - 1)\pi_{AB} + (2P - 1)(1 - T)\pi_A + (1 - P)(2T - 1)\pi_B + (1 - P)(1 - T) \quad (1)$$

$$\theta_{10} = -(2P - 1)(2T - 1)\pi_{AB} + (2P - 1)T\pi_A - (1 - P)(2T - 1)\pi_B + (1 - P)T \quad (2)$$

$$\theta_{01} = -(2P - 1)(2T - 1)\pi_{AB} - (2P - 1)(1 - T)\pi_A + P(2T - 1)\pi_B + P(1 - T) \quad (3)$$

$$\theta_{00} = (2P - 1)(2T - 1)\pi_{AB} - (2P - 1)T\pi_A - P(2T - 1)\pi_B + PT \quad (4)$$

where π_A , π_B and π_{AB} are the population proportions of respondents belonging to character A, B and AB, respectively.

Simple model which is a special case of Christofides (2005) has unbiased estimators of the proportion π_A , π_B and π_{AB} as:

$$\hat{\pi}_A^* = \frac{\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2(2P - 1)} \quad (5)$$

$$\hat{\pi}_B^* = \frac{\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2T - 1)}{2(2T - 1)} \quad (6)$$

$$\hat{\pi}_{AB}^* = \frac{(P + T)\hat{\theta}_{11} + (T - P)\hat{\theta}_{10} + (P - T)\hat{\theta}_{01} + (2 - P - T)\hat{\theta}_{00} - T(1 - P) - P(1 - T)}{2(2P - 1)(2T - 1)} \quad (7)$$

where P and T are the probabilities of belonging to sensitive characteristic A in Deck I and sensitive characteristic B in Deck II, respectively, for $T \neq 0.5$ and $P \neq 0.5$. Also, $\hat{\theta}_{11} = N_{11}/n$, $\hat{\theta}_{10} = N_{10}/n$, $\hat{\theta}_{01} = N_{01}/n$ and $\hat{\theta}_{00} = N_{00}/n$.

The variances of the estimators are given as:

$$V(\hat{\pi}_A^*) = \frac{\pi_A(1 - \pi_A)}{n} + \frac{P(1 - P)}{n(2P - 1)^2} \quad (8)$$

$$V(\hat{\pi}_B^*) = \frac{\pi_B(1 - \pi_B)}{n} + \frac{T(1 - T)}{n(2T - 1)^2} \tag{9}$$

$$V(\hat{\pi}_{AB}^*) = \frac{\pi_{AB}(1 - \pi_{AB})}{n} + \frac{(2P - 1)^2 T(1 - T)\pi_A + P(1 - P)(2T - 1)^2 \pi_B + PT(1 - P)(1 - T)}{n(2P - 1)^2 (2T - 1)^2} \tag{10}$$

for $T = P \neq 0.5$.

2.2. The crossed model

The procedure and all assumptions are the same as in simple model. The method also uses two decks of cards but with the statements on the cards differ for the Crossed Model (see Lee et al., 2013 for details). Using crossed model, each response, (Yes, Yes), (Yes, No), (No, Yes), and (No, No) for individuals in population may also occur in four different ways; the probabilities of getting each response were given as:

$$\theta_{11}^* = \pi_{AB}\{PT + (1 - P)(1 - T)\} - \pi_A(1 - P)(1 - T) - \pi_B(1 - P)(1 - T) + (1 - P)(1 - T) \tag{11}$$

$$\theta_{10}^* = -\pi_{AB}\{PT + (1 - P)(1 - T)\} - \pi_A\{(1 - P)T - 1\} - \pi_B(1 - P)T + (1 - P)T \tag{12}$$

$$\theta_{01}^* = -\pi_{AB}\{PT + (1 - P)(1 - T)\} - \pi_A P(1 - T) - \pi_B\{P(1 - T) - 1\} + P(1 - T) \tag{13}$$

$$\theta_{00}^* = \pi_{AB}\{PT + (1 - P)(1 - T)\} - \pi_A PT - \pi_B PT + PT \tag{14}$$

The proposed unbiased estimator of the population proportions π_A , π_B and π_{AB} for the crossed model is given by:

$$\hat{\pi}'_A = \frac{1}{2} + \frac{(T - P + 1)(\hat{\theta}_{11}^* - \hat{\theta}_{00}^*) + (P + T - 1)(\hat{\theta}_{10}^* - \hat{\theta}_{01}^*)}{2(P + T - 1)} \tag{15}$$

$$\hat{\pi}'_B = \frac{1}{2} + \frac{(T - P + 1)(\hat{\theta}_{11}^* - \hat{\theta}_{00}^*) + (P + T - 1)(\hat{\theta}_{01}^* - \hat{\theta}_{10}^*)}{2(P + T - 1)} \tag{16}$$

$$\hat{\pi}'_{AB} = \frac{PT\hat{\theta}_{11}^* - (1 - P)(1 - T)\hat{\theta}_{00}^*}{\{PT + (1 - P)(1 - T)\}(P + T - 1)} \tag{17}$$

for $P + T \neq 1$.

The variances of these estimators are given as:

$$V(\hat{\pi}'_A) = \frac{\pi_A(1 - \pi_A)}{n} + \frac{(1 - P)[T\{PT + (1 - P)(1 - T)\}(1 - \pi_A - \pi_B + 2\pi_{AB})]}{n(P + T - 1)^2} \tag{18}$$

$$V(\hat{\pi}'_B) = \frac{\pi_B(1 - \pi_B)}{n} + \frac{(1 - T)[P\{PT + (1 - P)(1 - T)\}(1 - \pi_A - \pi_B + 2\pi_{AB})]}{n(P + T - 1)^2} \tag{19}$$

$$V(\hat{\pi}'_{AB}) = \frac{\pi_{AB}(1 - \pi_{AB})}{n} + \frac{1}{n\{PT + (1 - P)(1 - T)\}(P + T - 1)^2} \left[\pi_{AB}\{P^2T^2 + (1 - P)^2(1 - T)^2 - \{PT + (1 - P)(1 - T)\}(P + T - 1)^2\} + PT(1 - P)(1 - T)(1 - \pi_A - \pi_B) \right] \tag{20}$$

for $P + T \neq 1$.

On the probability of protection for the respondents in each sensitive character, Lee et al. (2013) decided to keep the choice of $P = 0.7$ and $T = 0.7$ in both the simple and crossed model, Ewemooje and Amahia (2015) chose $P = 0.6$ and $\lambda = 0.7$ while Batool and Shabbir (2015) used various

probabilities but inferred that their proposed model perform better when there is equal probabilities of protection on the sensitive characteristics i.e. $W_1 = Q_1 = P_1 = T_1 = P = T = 0.7$. Ewemooje and Amahia (2016) also looked at the efficiency of Lee et al.'s (2013) simple and crossed models in the face of equal and unequal probabilities of protection. It was discovered that equal probabilities of protection increases the efficiency. Hence, this work aimed at improving the procedure by introducing equal probabilities of protection on the randomized devices.

3. Equal protection model

The model allows for equal probabilities of protection on the randomized devices for two sensitive characteristics where respondent is required to truthfully answer “yes” or “no” to the sensitive question; if “no”, the respondent is to go through two randomized devices (deck of cards). He or she is required to draw a card from deck I of cards containing two statements:

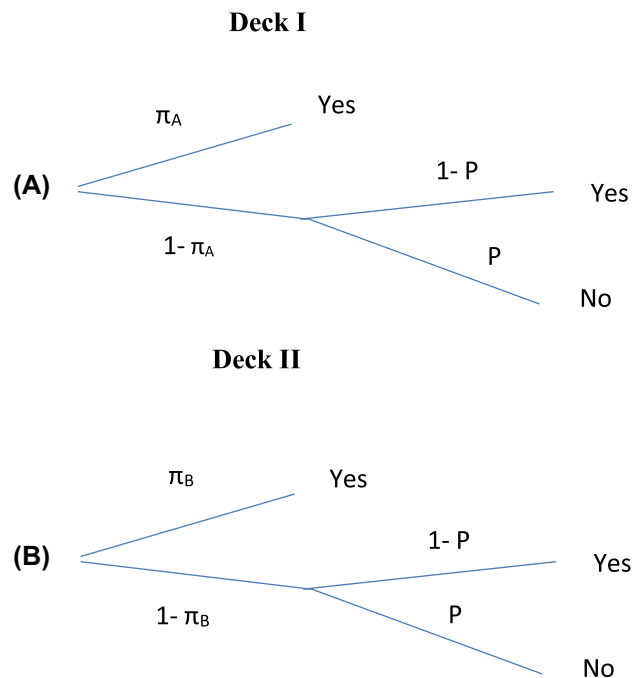
- I belong to character A with probability P .
- I do not belong to character A with probability $1 - P$.

And respond “yes” or “no” accordingly without reporting the statement on the card to the interviewer. Also, the respondent proceeds to next stage by following the same procedure of drawing another card from deck II of cards containing either of the two statements:

- I belong to character B with probability P .
- I do not belong to character B with probability $1 - P$.

And answer “yes” or “no” accordingly without reporting the statement on the card to the interviewer. The procedure can be duly represented by Figure 1.

Figure 1. Tree diagram showing proposed model.



The responses obtained can be grouped into four different places: n_{11} , number of respondents that answered “yes” to character A and “yes” to character B; n_{10} , number of respondents that answer “yes” to character A and “no” to character B; n_{01} , number of respondents that answered “no” to character A and “yes” to character B and n_{00} , number of respondents that answer “no” to character A and “no” to character B. Note that $n_{11} + n_{10} + n_{01} + n_{00} = n$ i.e. $\sum_{i=0}^1 \sum_{j=0}^1 n_{ij} = n$. Consequently, the probabilities of (yes, yes), (yes, no), (no, yes), and (no, no) can be represented with as θ_{11} , θ_{10} , θ_{01} and θ_{00} respectively where $\sum_{i=0}^1 \sum_{j=0}^1 \theta_{ij} = 1$.

Using the proposed model, we have the expression for the probabilities as:

$$\theta_{11} = P^2 \pi_{AB} + P(1 - P)\pi_A + P(1 - P)\pi_B + (1 - P)^2 \tag{21}$$

$$\theta_{10} = -P^2 \pi_{AB} + P^2 \pi_A - P(1 - P)\pi_B + P(1 - P) \tag{22}$$

$$\theta_{01} = -P^2 \pi_{AB} - P(1 - P)\pi_A + P^2 \pi_B + P(1 - P) \tag{23}$$

$$\theta_{00} = P^2 \pi_{AB} - P^2 \pi_A - P^2 \pi_B + P^2 \tag{24}$$

The distance between the observed probabilities, $\hat{\theta}_{ij}$, and the true probabilities, θ_{ij} is minimized using the expression:

$$D = \frac{1}{2} \sum_{i=0}^1 \sum_{j=0}^1 (\theta_{ij} - \hat{\theta}_{ij})^2 \tag{25}$$

We further differentiate D with respect to π_A , π_B and π_{AB} and equate to zero i.e. $\frac{dD}{d\pi_A} = 0$, $\frac{dD}{d\pi_B} = 0$ and $\frac{dD}{d\pi_{AB}} = 0$. Then solve simultaneously in order to obtain unbiased estimators of π_A , π_B and π_{AB} .

The proposed equal protection unbiased estimators of π_A , π_B and π_{AB} are given by:

$$\hat{\pi}_A = \frac{\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P} \tag{26}$$

$$\hat{\pi}_B = \frac{\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P} \tag{27}$$

$$\hat{\pi}_{AB} = \frac{(4P - 1)\hat{\theta}_{11} - \hat{\theta}_{10} - \hat{\theta}_{01} - (4P - 3)\hat{\theta}_{00} + (2P - 1)^2}{4P^2} \tag{28}$$

for $P > 0$.

3.1. Tests for unbiasedness of the estimators

In testing for unbiasedness of $\hat{\pi}_A$, we have:

$$\begin{aligned} E(\hat{\pi}_A) &= E \left[\frac{\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P} \right] \\ &= \frac{E(\hat{\theta}_{11}) + E(\hat{\theta}_{10}) - E(\hat{\theta}_{01}) - E(\hat{\theta}_{00}) + (2P - 1)}{2P} \\ &= \frac{P^2 \pi_{AB} + P(1 - P)\pi_A + P(1 - P)\pi_B + (1 - P)^2 - P^2 \pi_{AB} + P^2 \pi_A - P(1 - P)\pi_B + P(1 - P)}{2P} \\ &\quad + \frac{P^2 \pi_{AB} + P(1 - P)\pi_A - P^2 \pi_B - P(1 - P) - P^2 \pi_{AB} + P^2 \pi_A + P^2 \pi_B - P^2 + (2P - 1)}{2P} \\ &= \frac{2\pi_A [P^2 + P(1 - P) + (1 - P)^2 - P^2 + (2P - 1)]}{2P} \end{aligned}$$

$$E(\hat{\pi}_A) = \frac{2P\pi_A}{2P} = \pi_A$$

Hence, $\hat{\pi}_A$ is unbiased.

In testing for unbiasedness of $\hat{\pi}_B$, we have:

$$\begin{aligned} E(\hat{\pi}_B) &= E\left[\frac{\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P}\right] \\ &= \frac{E(\hat{\theta}_{11}) - E(\hat{\theta}_{10}) + E(\hat{\theta}_{01}) - E(\hat{\theta}_{00}) + (2P - 1)}{2P} \\ &= \frac{P^2\pi_{AB} + P(1 - P)\pi_A + P(1 - P)\pi_B + (1 - P)^2 + P^2\pi_{AB} - P^2\pi_A + P(1 - P)\pi_B - P(1 - P)}{2P} \\ &\quad + \frac{-P^2\pi_{AB} - P(1 - P)\pi_A + P^2\pi_B + P(1 - P) - P^2\pi_{AB} + P^2\pi_A + P^2\pi_B - P^2 + (2P - 1)}{2P} \\ &= \frac{2\pi_B[P^2 + P(1 - P) + (1 - P)^2 - P^2 + (2P - 1)]}{2P} \end{aligned}$$

$$E(\hat{\pi}_B) = \frac{2P\pi_B}{2P} = \pi_B$$

Hence, $\hat{\pi}_B$ is unbiased.

In testing for unbiasedness of $\hat{\pi}_{AB}$, we have:

$$\begin{aligned} E(\hat{\pi}_{AB}) &= E\left[\frac{(4P - 1)\hat{\theta}_{11} - \hat{\theta}_{10} - \hat{\theta}_{01} - (4P - 3)\hat{\theta}_{00} + (2P - 1)^2}{4P^2}\right] \\ &= \frac{(4P - 1)E(\hat{\theta}_{11}) - E(\hat{\theta}_{10}) - E(\hat{\theta}_{01}) - (4P - 3)E(\hat{\theta}_{00}) + (2P - 1)^2}{4P^2} \\ &= \frac{(4P - 1)[P^2\pi_{AB} + P(1 - P)\pi_A + P(1 - P)\pi_B + (1 - P)^2] + P^2\pi_{AB} - P^2\pi_A + P(1 - P)\pi_B}{4P^2} \\ &\quad + \frac{-P(1 - P) + P^2\pi_{AB} + P(1 - P)\pi_A - P^2\pi_B - P(1 - P)}{4P^2} \\ &\quad + \frac{-(4P - 3)[P^2\pi_{AB} - P^2\pi_A - P^2\pi_B + P^2] + (2P - 1)^2}{4P^2} \\ &= \frac{\pi_{AB}[P^2(4P - 1) + 2P^2 - P^2(4P - 3)] + \pi_A[P(1 - P)(4P - 1) - P^2 + P(1 - P)]}{4P^2} \\ &\quad + \frac{P^2(4P - 3) + \pi_B[P(1 - P)(4P - 1) - P^2 + P(1 - P) + P^2(4P - 3)]}{4P^2} \\ &\quad + \frac{(1 - P)^2(4P - 1) - 2P(1 - P) - P^2(4P - 3) + (2P - 1)^2}{4P^2} \end{aligned}$$

$$E(\hat{\pi}_{AB}) = \frac{4P^2\pi_{AB}}{4P^2} = \pi_{AB}$$

Hence, $\hat{\pi}_{AB}$ is also unbiased.

3.2. Variance estimation

To obtain the variance of the proposed estimators $\hat{\pi}_A$, $\hat{\pi}_B$ and $\hat{\pi}_{AB}$ presented in Equations (26)–(28), respectively, we have:

$$V(\hat{\pi}_A) = V \left\{ \frac{\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P} \right\}$$

$$V(\hat{\pi}_B) = V \left\{ \frac{\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)}{2P} \right\}$$

$$V(\hat{\pi}_{AB}) = V \left\{ \frac{(4P - 1)\hat{\theta}_{11} - \hat{\theta}_{10} - \hat{\theta}_{01} - (4P - 3)\hat{\theta}_{00} + (2P - 1)^2}{4P^2} \right\}$$

In solving these variances, the following variance and covariance operators were used:

$$\begin{aligned} V(x_{11}) &= \theta_{11}(1 - \theta_{11}), V(x_{10}) = \theta_{10}(1 - \theta_{10}), V(x_{01}) = \theta_{01}(1 - \theta_{01}), V(x_{00}) \\ &= \theta_{00}(1 - \theta_{00}), C(x_{11}, x_{10}) = \theta_{11}\theta_{10}, C(x_{11}, x_{01}) = \theta_{11}\theta_{01}, C(x_{11}, x_{00}) = \theta_{11}\theta_{00}, C(x_{10}, x_{01}) \\ &= \theta_{10}\theta_{01}, C(x_{10}, x_{00}) = \theta_{10}\theta_{00}, C(x_{01}, x_{00}) = \theta_{01}\theta_{00} \end{aligned}$$

where V and C are the operators of variance and covariance over the randomized response device, respectively. x_{11} is “yes” response for character A and “yes” response for character B, x_{10} is “yes” response for character A and “no” response for character B, x_{01} is “no” response for character A and “yes” response for character B while x_{00} is “no” response for character A and “no” response for character B.

Note that; $V(x_{11} + x_{10} - x_{01} - x_{00}) = V(x_{11}) + V(x_{10}) + V(x_{01}) + V(x_{00})$
 $- 2[C(x_{11}, x_{10}) - C(x_{11}, x_{01}) - C(x_{11}, x_{00}) - C(x_{10}, x_{01}) - C(x_{10}, x_{00}) + C(x_{01}, x_{00})]$

Solving this rigorously, the variances for the unbiased estimators $\hat{\pi}_A$, $\hat{\pi}_B$ and $\hat{\pi}_{AB}$ are given by:

$$V(\hat{\pi}_A) = \frac{\pi_A [(2P - 1) - P\pi_A] + (1 - P)}{nP} \tag{29a}$$

$$V(\hat{\pi}_B) = \frac{\pi_B [(2P - 1) - P\pi_B] + (1 - P)}{nP} \tag{30a}$$

$$V(\hat{\pi}_{AB}) = \frac{\pi_{AB} [(2P - 1)^2 - P^2\pi_{AB}] + (2P - 1)(1 - P)\pi_A + (1 - P)(2P - 1)\pi_B + (1 - P)^2}{nP^2} \tag{31a}$$

for $P > 0$.

The sample estimators of the variances $V(\hat{\pi}_A)$, $V(\hat{\pi}_B)$ and $V(\hat{\pi}_{AB})$ are given as follows:

$$\hat{V}(\hat{\pi}_A) = \frac{\hat{\pi}_A [(2P - 1) - P\hat{\pi}_A] + (1 - P)}{nP} \tag{29b}$$

$$\hat{V}(\hat{\pi}_B) = \frac{\hat{\pi}_B [(2P - 1) - P\hat{\pi}_B] + (1 - P)}{nP} \tag{30b}$$

$$\hat{V}(\hat{\pi}_{AB}) = \frac{\hat{\pi}_{AB} [(2P - 1)^2 - P^2\hat{\pi}_{AB}] + (2P - 1)(1 - P)\hat{\pi}_A + (1 - P)(2P - 1)\hat{\pi}_B + (1 - P)^2}{nP^2} \tag{31b}$$

for $P > 0$.

3.2.1. Remarks

An unbiased estimator of the proportion of people belonging to at least one characteristic A or B is given as:

$$\hat{\pi}_{A \cup B} = \hat{\pi}_A + \hat{\pi}_B - \hat{\pi}_{AB} \quad (32)$$

An unbiased estimator of the proportion of people belonging to precisely characteristic A and not B is given as:

$$\hat{\pi}_{A-B} = \hat{\pi}_A - \hat{\pi}_{AB} \quad (33)$$

An unbiased estimator of the proportion of people belonging to precisely characteristic B and not A is given as:

$$\hat{\pi}_{B-A} = \hat{\pi}_B - \hat{\pi}_{AB} \quad (34)$$

An unbiased estimator of the population total of people belonging to the two sensitive characteristics in a target population, N, is given as:

$$\hat{Y}_A = N\hat{\pi}_A$$

$$\hat{Y}_A = \frac{N[\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)]}{2P} \quad (35)$$

$$\hat{Y}_B = N\hat{\pi}_B$$

$$\hat{Y}_B = \frac{N[\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2P - 1)]}{2P} \quad (36)$$

$$\hat{Y}_{AB} = N\hat{\pi}_{AB}$$

$$\hat{Y}_{AB} = \frac{N[(4P - 1)\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - (4P - 3)\hat{\theta}_{00} + (2P - 1)^2]}{4P^2} \quad (37)$$

It then follows that the variances of the estimate of the population parameters are given by:

$$V(\hat{Y}_J) = N^2V(\hat{\pi}_J)$$

$$V(\hat{Y}_J) = \frac{N^2[\pi_J[(2P - 1) - P\pi_J] + (1 - P)]}{nP}$$

$$V(\hat{Y}_J) = \frac{NK[\pi_J[(2P - 1) - P\pi_J] + (1 - P)]}{P}, \quad J = A, B \quad (38)$$

$$V(\hat{Y}_{AB}) = N^2V(\hat{\pi}_{AB})$$

$$V(\hat{Y}_{AB}) = \frac{N^2[\pi_{AB}[(2P - 1)^2 - P^2\pi_{AB}] + (2P - 1)(1 - P)\pi_A + (1 - P)(2P - 1)\pi_B + (1 - P)^2]}{nP^2}$$

$$V(\hat{Y}_{AB}) = \frac{NK[\pi_{AB}[(2P - 1)^2 - P^2\pi_{AB}] + (2P - 1)(1 - P)\pi_A + (1 - P)(2P - 1)\pi_B + (1 - P)^2]}{P^2} \quad (39)$$

where $K = \frac{N}{n}$ and $P > 0$.

4. Efficiency comparison

It has been shown by several authors that to ensure reliable result in practical survey and increase confidentiality of the respondent, the probabilities of protection should be greater than 0.5 and not too high. Hence, we set $P = 0.7$ and $T = 0.7$ for the purpose of comparing the proposed model with the Lee et al. (2013).

4.1. Efficiency comparison with simple model

Here, we compare the proposed model with the simple model by obtaining the percentage relative efficiency of the proposed estimators with respect to the simple model estimators using Equations (40)–(42):

$$\text{PRE}(\hat{\pi}_A) = \frac{V(\hat{\pi}_A^*)}{V(\hat{\pi}_A)} \times 100 \tag{40}$$

$$\text{PRE}(\hat{\pi}_B) = \frac{V(\hat{\pi}_B^*)}{V(\hat{\pi}_B)} \times 100 \tag{41}$$

$$\text{PRE}(\hat{\pi}_{AB}) = \frac{V(\hat{\pi}_{AB}^*)}{V(\hat{\pi}_{AB})} \times 100 \tag{42}$$

4.2. Efficiency comparison with crossed model

Comparing the proposed model with the crossed model using the percentage relative efficiency of the proposed estimators with respect to the crossed model estimators with the use of Equations (43)–(45), we obtained the results shown in Tables 3 and 4.

$$\text{PRE}(\hat{\pi}_A) = \frac{V(\hat{\pi}_A')}{V(\hat{\pi}_A)} \times 100 \tag{43}$$

$$\text{PRE}(\hat{\pi}_B) = \frac{V(\hat{\pi}_B')}{V(\hat{\pi}_B)} \times 100 \tag{44}$$

$$\text{PRE}(\hat{\pi}_{AB}) = \frac{V(\hat{\pi}_{AB}')}{V(\hat{\pi}_{AB})} \times 100 \tag{45}$$

5. Results and discussion

The proposed model is more efficient than the simple model of Lee et al. (2013) at all levels. The percentage relative efficiency (PRE) of $\hat{\pi}_A$ and $\hat{\pi}_B$ ranges between 597.61 and 1,223.01 each while the PRE of $\hat{\pi}_{AB}$ ranges between 706.23 and 824.68. Also, the PRE of $\hat{\pi}_A$ and $\hat{\pi}_B$ increases with increase in π_A and π_B while PRE of $\hat{\pi}_{AB}$ decreases (see Table 1).

Table 2 shows that at all levels of π_{AB} , the percentage relative efficiency (PRE) decreases with increase in the values of π_A and π_B . When $\pi_{AB} = 0.05$; the percentage relative efficiency (PRE) ranges between 706.23 and 824.68, when $\pi_{AB} = 0.1$; PRE ranges between 701.58 and 812.81 while when $\pi_{AB} = 0.2$; PRE ranges between 713.55 and 831.59. However, as the value of π_{AB} increases there is increase in the PRE.

Table 1. Percentage relative efficiency (PRE) of proposed over Lee et al. (2013)

π_A	π_B	PRE (\hat{x}_A)	PRE (\hat{x}_B)	PRE (\hat{x}_{AB})
0.1	0.1	601.67	601.67	824.68
	0.2	601.67	597.61	798.58
	0.3	601.67	609.24	776.81
	0.4	601.67	637.32	758.37
	0.5	601.67	686.81	742.56
	0.6	601.67	770.09	728.84
	0.7	601.67	917.72	716.83
	0.8	601.67	1,223.01	706.23
0.2	0.1	597.61	601.67	798.58
	0.2	597.61	597.61	776.81
	0.3	597.61	609.24	758.37
	0.4	597.61	637.32	742.56
	0.5	597.61	686.81	728.84
	0.6	597.61	770.09	716.83
	0.7	597.61	917.72	706.23
0.3	0.1	609.24	601.67	776.81
	0.2	609.24	597.61	758.37
	0.3	609.24	609.24	742.56
	0.4	609.24	637.32	728.84
	0.5	609.24	686.81	716.83
	0.6	609.24	770.09	706.23
0.4	0.1	637.32	601.67	758.37
	0.2	637.32	597.61	742.56
	0.3	637.32	609.24	728.84
	0.4	637.32	637.32	716.83
	0.5	637.32	686.81	706.23
0.5	0.1	686.81	601.67	742.56
	0.2	686.81	597.61	728.84
	0.3	686.81	609.24	716.83
	0.4	686.81	637.32	706.23
0.6	0.1	770.09	601.67	728.84
	0.2	770.09	597.61	716.83
	0.3	770.09	609.24	706.23
0.7	0.1	917.72	601.67	716.83
	0.2	917.72	597.61	706.23
0.8	0.1	1,223.01	601.67	706.23

Note: Simple model for $\pi_{AB} = 0.05$, $P = 0.7$ and $T = 0.7$.

Table 2. Percentage relative efficiency (PRE) of proposed over Lee et al. (2013)

π_A	π_B	PRE ($\hat{\pi}_{AB}$)		
		0.05	0.1	0.2
0.1	0.1	824.68	812.81	831.59
	0.2	798.58	788.58	805.95
	0.3	776.81	768.24	784.41
	0.4	758.37	750.94	766.05
	0.5	742.56	736.03	750.22
	0.6	728.84	723.06	736.42
	0.7	716.83	711.67	724.30
	0.8	706.23	701.58	713.55
0.2	0.1	798.58	788.58	805.95
	0.2	776.81	768.24	784.41
	0.3	758.37	750.94	766.05
	0.4	742.56	736.03	750.22
	0.5	728.84	723.06	736.42
	0.6	716.83	711.67	724.30
	0.7	706.23	701.58	713.55
0.3	0.1	776.81	768.24	784.41
	0.2	758.37	750.94	766.05
	0.3	742.56	736.03	750.22
	0.4	728.84	723.06	736.42
	0.5	716.83	711.67	724.30
	0.6	706.23	701.58	713.55
0.4	0.1	758.37	750.94	766.05
	0.2	742.56	736.03	750.22
	0.3	728.84	723.06	736.42
	0.4	716.83	711.67	724.30
	0.5	706.23	701.58	713.55
0.5	0.1	742.56	736.03	750.22
	0.2	728.84	723.06	736.42
	0.3	716.83	711.67	724.30
	0.4	706.23	701.58	713.55
0.6	0.1	728.84	723.06	736.42
	0.2	716.83	711.67	724.30
	0.3	706.23	701.58	713.55
0.7	0.1	716.83	711.67	724.30
	0.2	706.23	701.58	713.55
0.8	0.1	706.23	701.58	713.55

Note: Simple model for different levels of π_{AB} when $P = 0.7$ and $T = 0.7$.

Table 3. Percentage relative efficiency (PRE) of proposed over Lee et al. (2013)

π_A	π_B	PRE ($\hat{\pi}_A$)	PRE ($\hat{\pi}_B$)	PRE ($\hat{\pi}_{AB}$)
0.1	0.1	332.53	332.53	207.48
	0.2	299.87	312.09	171.19
	0.3	267.21	297.27	140.92
	0.4	234.56	286.02	115.28
	0.5	201.90	277.20	93.29
	0.6	169.24	270.09	74.21
	0.7	136.58	264.24	57.51
	0.8	103.93	259.34	42.77
0.2	0.1	312.09	299.87	171.19
	0.2	281.20	281.20	140.92
	0.3	250.30	266.81	115.28
	0.4	219.41	254.77	93.29
	0.5	188.51	243.74	74.21
	0.6	157.62	232.33	57.51
	0.7	126.72	218.35	42.77
0.3	0.1	297.27	267.21	140.92
	0.2	266.81	250.30	115.28
	0.3	236.34	236.34	93.29
	0.4	205.88	223.52	74.21
	0.5	175.42	210.27	57.51
	0.6	144.96	194.57	42.77
0.4	0.1	541.67	316.60	183.60
	0.2	468.75	284.09	142.84
	0.3	395.83	256.90	108.78
	0.4	322.92	231.53	79.88
	0.5	250.00	204.84	55.06
0.5	0.1	277.20	201.90	93.29
	0.2	243.74	188.51	74.21
	0.3	210.27	175.42	57.51
	0.4	176.81	161.02	42.77
0.6	0.1	270.09	169.24	74.21
	0.2	232.33	157.62	57.51
	0.3	194.57	144.96	42.77
0.7	0.1	264.24	136.58	57.51
	0.2	218.35	126.72	42.77
0.8	0.1	259.34	103.93	42.77

Note: Crossed model for $\pi_{AB} = 0.05$, $P = 0.7$ and $T = 0.7$.

The proposed model is more efficient than the crossed model of Lee et al. (2013) when considering the two sensitive characteristics independently while the interaction show that the proposed is only efficient when $\pi_A + \pi_B \leq 0.7$. The percentage relative efficiency (PRE) of $\hat{\pi}_A$ and $\hat{\pi}_B$ ranges between 103.93 and 332.53 each while the PRE of $\hat{\pi}_{AB}$ ranges between 108.78 and 207.48 as seen in Table 3. However, the PRE of $\hat{\pi}_A$ and $\hat{\pi}_B$ decreases with increase in π_A and π_B while PRE of $\hat{\pi}_{AB}$ also decreases.

In Table 4; when $\pi_{AB} = 0.05$ and $\pi_{AB} = 0.1$, the proposed model becomes inefficient against the crossed model as $\pi_A + \pi_B > 0.7$ while when $\pi_{AB} = 0.2$, the proposed model becomes more efficient

Table 4. Percentage relative efficiency (PRE) of proposed over Lee et al. (2013)

π_A	π_B	PRE ($\hat{\pi}_{AB}$)		
		0.05	0.1	0.2
0.1	0.1	207.48	249.75	339.23
	0.2	171.19	210.91	293.00
	0.3	140.92	178.31	254.14
	0.4	115.28	150.58	221.02
	0.5	93.29	126.69	192.46
	0.6	74.21	105.89	167.58
	0.7	57.51	87.63	145.71
	0.8	42.77	71.47	126.33
0.2	0.1	171.19	210.91	293.00
	0.2	140.92	178.31	254.14
	0.3	115.28	150.58	221.02
	0.4	93.29	126.69	192.46
	0.5	74.21	105.89	167.58
	0.6	57.51	87.63	145.71
	0.7	42.77	71.47	126.33
0.3	0.1	140.92	178.31	254.14
	0.2	115.28	150.58	221.02
	0.3	93.29	126.69	192.46
	0.4	74.21	105.89	167.58
	0.5	57.51	87.63	145.71
	0.6	42.77	71.47	126.33
0.4	0.1	183.60	150.58	221.02
	0.2	142.84	126.69	192.46
	0.3	108.78	105.89	167.58
	0.4	79.88	87.63	145.71
	0.5	55.06	71.47	126.33
0.5	0.1	93.29	126.69	192.46
	0.2	74.21	105.89	167.58
	0.3	57.51	87.63	145.71
	0.4	42.77	71.47	126.33
0.6	0.1	74.21	105.89	167.58
	0.2	57.51	87.63	145.71
	0.3	42.77	71.47	126.33
0.7	0.1	57.51	87.63	145.71
	0.2	42.77	71.47	126.33
0.8	0.1	42.77	71.47	126.33

Note: Crossed model for different levels of π_{AB} when $P = 0.7$ and $T = 0.7$.

than the crossed model at all levels of π_A and π_B . Therefore, as the values of π_{AB} increases there is observed increase in the PRE.

6. Conclusion

It has been shown that the proposed model is more efficient than the simple model of Lee et al. (2013) at all levels. The proposed model is also more efficient than the crossed model of Lee et al.

(2013) when considering the two sensitive characteristics independently while the interaction show that the proposed is efficient when $\pi_A + \pi_B \leq 0.7$. This is because as the sum increases to one less interaction occurs but as π_{AB} increases to 0.2, it becomes more efficient at all levels. It was also observed that there is system breakdown as $P = 0.5$ or $T = 0.5$ and $P + T = 1$, in simple and crossed models, respectively, while the proposed model is estimable at all point greater than zero. Therefore, it is concluded that the proposed model and estimators are more robust and efficient than the conventional estimators.

Acknowledgement

The author acknowledges with thanks immense suggestions and contributions of the reviewers toward improving this work.

Funding

The author received no direct funding for this research.

Author details

Olusegun S. Ewemooje¹

E-mails: osewemooje@futa.edu.ng, sunsege@yahoo.com

ORCID ID: <http://orcid.org/0000-0003-3236-6018>

¹ Department of Statistics, Federal University of Technology, Akure, Nigeria.

Citation information

Cite this article as: Estimating two sensitive characters with equal probabilities of protection, Olusegun S. Ewemooje, *Cogent Mathematics* (2017), 4: 1319607.

References

- Adebola, F. B., Adediran, A. A., & Ewemooje, O. S. (2017). Hybrid tripartite randomized response technique. *Communications in Statistics-Theory and Methods*, doi:10.1080/03610926.2016.1277760
- Adebola, F. B., & Johnson, O. O. (2015). An improved Warner's randomized response model. *International Journal of Statistics and Applications*, 5, 263–267. doi:10.5923/j.statistics.20150506.01
- Batool, F., & Shabbir, J. (2015). A two stage design for multi variate estimation of proportions. *Communications in Statistics-Theory and Methods*, 45, 5412–5426. doi:10.1080/03610926.2014.942435
- Beam, G. (2012). *The problem with survey research*. New Brunswick, NJ: Transaction. ISBN: 978-1-4128-4603-5.
- Christofides, T. C. (2005). Randomized response technique for two sensitive characteristics at the same time. *Metrika*, 62, 53–63. <https://doi.org/10.1007/s001840400355>
- Ewemooje, O. S., & Amahia, G. N. (2015). Improved randomized response technique for two sensitive attributes. *Afrika Statistika*, 10, 839–852. doi:10.16929/as/2015.639.78
- Ewemooje, O. S., & Amahia, G. N. (2016). Improving the efficiency of randomized response technique for two sensitive characters. *FUTA Journal of Research in Sciences*, 12, 65–72.
- Greenberg, B. G., Abul-El, Jr., A.-L. A., Simmons, W. R., & Horvitz, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, 64, 520–539. <https://doi.org/10.1080/01621459.1969.10500991>
- Lee, C., Sedory, S. A., & Singh, S. (2013). Estimating at least seven measures of qualitative variables from a single sample using randomized response technique. *Statistics & Probability Letters*, 83, 399–409. <https://doi.org/10.1016/j.spl.2012.10.004>
- Mangat, N. S. (1994). An improved randomized response strategy. *Journal of the Royal Statistical Society, Series B*, 56, 93–95.
- Mangat, N. S., & Singh, R. (1990). An alternative randomized response procedure. *Biometrika*, 77, 439–442.
- Odumade, O., & Singh, S. (2009). Efficient use of two decks of cards in randomized response sampling. *Communications in Statistics-Theory and Methods*, 38, 439–446. <https://doi.org/10.1080/03610920802225065>
- Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63–69. <https://doi.org/10.1080/01621459.1965.10480775>



© 2017 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

