Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 73(1) 2019 41–52

Scrambled Response Techniques in Two Wave Rotation Sampling for Estimating Population Mean of Sensitive Characteristics with Case Study

Kumari Priyanka[1], Pidugu Trisandhya[1] and Richa Mittal[2]
[1]*Shivaji College, University of Delhi, New Delhi*
[2]*NIT Calicut*

Received 11 August 2017; Revised 01 August 2018; Accepted 03 January 2019

SUMMARY

Present work is an attempt to use non-sensitive auxiliary variable and scrambled response techniques (SRT) to estimate population mean of a sensitive variable. A class of estimator is proposed to estimate the population mean of a sensitive variable in sampling over two successive waves. Various members of the proposed class of estimators has been discussed. The proposed class of estimator has been analysed theoretically as well as empirically. It has been compared with modified general successive sampling estimator and with some of members of its own class. Simulation study has also been discussed. A case study of drug usage by students in a college on two successive waves has also been carried out.

Keywords: Randomized response technique, Scrambled response technique, Sensitive variable, Scrambling variable, Exponential ratio type estimators, Successive sampling, Optimum rotation rate, Non-sensitive auxiliary variable, Mean squared error, Bias.

Mathematics Subject Classification: 62D05

1. INTRODUCTION

Analysis of sensitive issues like negligence of governmental rules, number of abortion before marriage, bribing for some entrance exam, sexual indulgence during teenage, status of extramarital relationship, employing child labourers, child-sexual abuse, voluntary prostitution, commencement of crime, honour killing, drug intake etc., usually lead to over or under reporting of the true facts due to social or moral inclinations and stigma. Thus a significant deviation occurs in the results owing to socially desirable answers which do not comply to real scenario subsisting in the society.

There are two approaches to estimate population proportion or population mean of a quantitative sensitive variable. First approach is to reduce the stigma involved in answering such sensitive questions by providing certain privacy through a randomized response device following certain randomized response rule (Randomized Response Model). Warner (1965) was the first to provide such a randomizing model and later on extensive literature have been added by Horvitz *et al.* (1967), Greenberg *et al.* (1971), Christofides (2003, 2005), Kim and Elam (2007), Wu *et al.* (2008),Yan *et al.* (2009), Arnab (2011), Diana and Perri (2011), Arnab *et al.* (2012), Singh and Sedory (2012) and Sihm and Gupta (2015) etc.

All these authors have focused on estimation of population mean or proportion of sensitive characters using some randomised response models.

This approach becomes practically next to impossible when it comes to observe a very large sample since lifestyle has drastically changed and people are living a very fast life with certain time constraints so complete refusal to response is also encountered due to time consuming procedure involved in randomised response model. In such situations, second approach known as Scrambled Response Technique (SRT) which was introduced by Warner (1971) but was left for exploration and very first attempts were made

Corresponding author: Kumari Priyanka
E-mail address: priyanka.ism@gmail.com

by Pollock and Bek (1976) and Eichhorn and Hayre (1983), works as saviour. This technique reduces the impossibility of conducting a survey having large sample size with a sensitive issue to be addressed. In this technique to estimate the population mean of sensitive variable the respondent is asked to answer freely about the stigmatizing character by adding or multiplying a corrective scrambling factor to his/her response hiding real response from the interviewer. In this line a rich literature is available from Saha (2007), Koyuncu *et al.* (2014) and Hussain and Al-Zhrani (2016) etc. Mukherjee (2016) highlighted important issues related to sensitive estimation theory.

Moreover, these above said issues have been addressed through a single time survey in the literature available on sensitive variable analysis; instead these issues are required to be monitored continuously over time, since doing so will reflect the change of social scenario related to the sensitive issues as well as changed level of sensitivity of issue with respect to time. For example, any government of a county may be interested to record the mean number of rape cases in the country at starting of their ruling period. After recording them one time the government may be interested to decrease these for ensuring the better society. For this government can make strict laws against the rapist, more awareness of such laws can be spread amongst the females, it may also increase the level of security for females at work place and so on. After such precautious measures government may wish to see the changed level of the society by recoding the mean number of rape cases at the end of their ruling tenure. In order to monitor such a variable more than once, statistical tool generally recommended in literature is successive or rotation sampling. Jessen (1942) started the theory of rotation sampling by utilising all the information collected from previous wave. His pioneer work in this line has been followed by Patterson (1950),Sen(1973), Feng and Zou (1997), Singh and Priyanka (2008), Bandyopadhyay and Singh (2014), Priyanka and Mittal (2014, 2015a, 2015b), Priyanka *et al.* (2015) and many others.

None of the above works in successive sampling analyses sensitive issues which change over time. Very few attempts namely Arnab and Singh (2013) and Yu *et al.* (2014) are found which dealt with sensitive issues on successive waves while using randomized response technique. Hence, motivated with this scope of study, the present article endeavours to propose a class of estimator to estimate population mean of a sensitive variable and properties of proposed estimators including the optimum rotation rates have been derived up to first order of approximations. Discussion has been made regarding the distribution of scrambling variable. Also an empirical study has been worked out for the proposed class of estimator on two successive waves by means of a case study of drug usage by undergraduate students in a college. Simulation studies are rationalized to show the feasibility of proposed estimators.

2. SURVEY DESIGN AND ANALYSIS

2.1 Formulation

Let $U_1, U_2, \ldots, U_N$ be the finite population of N units, which has been sampled over two successive waves. It is assumed that size of the population remains unchanged but values of units change over two successive waves. The sensitive variable under study be denoted by $x(y)$ on the first (second) waves respectively. It is assumed that information on non-sensitive auxiliary variable z, stable in nature over the successive waves with completely known population mean $\bar{Z}$, is readily available on both the successive waves and positively correlated to x and y respectively. Simple random sample (without replacement) of n units is taken at the first wave. A random sub-sample of $m = n\lambda$ units is retained for use at the second wave. Now at the current wave, a simple random sample (without replacement) of $u = (n-m) = n\mu$ units are drawn afresh from the remaining $(N-n)$ units of the population so that the sample size on the second wave remains the same. Let μ and $\lambda(\mu+\lambda=1)$ are the fractions of fresh and matched samples respectively at the second (current) successive wave. Let S_1 and S_2 be two scrambling variables and the scrambled response for the sensitive variable $x(y)$ are perturbed to $g(h)$ respectively on first(second) waves. The following notations are considered here after:

$\bar{X}, \bar{Y}, \bar{G}, \bar{H}, \bar{Z}, \bar{S}_1, \bar{S}_2$	:Population means of the variables x, y, g, h, z, S_1, S_2 respectively.
$\bar{h}_u, \bar{h}_m, \bar{g}_m, \bar{g}_n$	:Sample mean of the variables based on sample sizes shown in suffices.
$\bar{z}_u, \bar{z}_m, \bar{z}_n$	:Sample mean of the non-sensitive auxiliary variable based on sample sizes shown in suffice.

$\rho_{yx},\ \rho_{xz},\ \rho_{yz},\ \rho_{hg},\ \rho_{hz},\ \rho_{xz}$	:Correlation coefficient between the variables shown in suffices.
$C_x,\ C_y,\ C_z$	: Coefficient of variation of variables shown in suffices.
$S_x^2,\ S_y^2,\ S_z^2$	:Population mean squared of $x,\ y$ and z respectively.
$\sigma_x^2,\ \sigma_y^2,\ \sigma_z^2,\ \sigma_{S_1}^2,\ \sigma_{S_2}^2$	:Population variances $x,\ y,\ z,\ S_1$ and S_2 respectively

2.2 Scrambled Response Model

In accordance with the sensitive issues, the original sensitive variables are perturbed using scrambling variables which are termed as scrambled response technique. Motivated by the literature which are available in this line proposed by eminent researchers, we intended to discuss a model where the sensitive variable $x(y)$ are perturbed to $g(h)$ respectively on first(second) waves as:

$$g = x\left[S_1 + \frac{S_2}{x}\right] \text{ on first wave} \tag{1}$$

and

$$h = y\left[S_1 + \frac{S_2}{y}\right] \text{ on second wave} \tag{2}$$

The scrambling variable $S_1,\ S_2$ may follow any distribution. The values of relevant parameters under the above scrambled response model are computed as:

$$\bar{X} = \frac{\bar{G} - \bar{S}_2}{\bar{S}_1} \text{ on first wave} \tag{3}$$

and

$$\bar{Y} = \frac{\bar{H} - \bar{S}_2}{\bar{S}_1} \text{ on second wave} \tag{4}$$

such that:

$$\rho_{hg} = \frac{\rho_{yx}\sigma_y\sigma_x\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{X}\bar{Y}\sigma_{S_1}^2 + \sigma_{S_2}^2}{\sqrt{\sigma_y^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \sigma_{S_1}^2\bar{Y}^2 + \sigma_{S_2}^2}\sqrt{\sigma_x^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \sigma_{S_1}^2\bar{X}^2 + \sigma_{S_2}^2}},$$

$$\rho_{hz} = \frac{\rho_{yz}\sigma_y\bar{S}_1}{\sqrt{\sigma_y^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \sigma_{S_1}^2\bar{Y}^2 + \sigma_{S_2}^2}},$$

$$\rho_{gz} = \frac{\rho_{xz}\sigma_x\bar{S}_1}{\sqrt{\sigma_x^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \sigma_{S_1}^2\bar{X}^2 + \sigma_{S_2}^2}}.$$

The renowned scrambled response model by Pollock and Bek (1976) and Eichhorn and Hayre (1983) becomes special cases of the above model for different choices of scrambling variables. These particular cases are discussed below:

Case 1: Taking $S_1 = 1$ in equation (1) and equation (2), the additive scrambled response model proposed by Pollock and Bek (1976) is obtained as $g_A = x + S_2$ and $h_A = y + S_2$. Various parameters related to additive model are given as

$$\left(\bar{X}\right)_A = \left(\bar{G}\right)_A - \bar{S}_2 \tag{5}$$

and

$$\left(\bar{Y}\right)_A = \left(\bar{H}\right)_A - \bar{S}_2 \tag{6}$$

such that

$$\rho_{hg} = \frac{\rho_{yx}\sigma_y\sigma_x + \sigma_{S_2}^2}{\sqrt{\sigma_y^2 + \sigma_{S_2}^2}\sqrt{\sigma_x^2 + \sigma_{S_2}^2}},\ \rho_{hz} = \frac{\rho_{yz}\sigma_y}{\sqrt{\sigma_y^2 + \sigma_{S_2}^2}},$$

$$\rho_{gz} = \frac{\rho_{xz}\sigma_x}{\sqrt{\sigma_x^2 + \sigma_{S_2}^2}},\ C_h^2 = \frac{\sigma_y^2 + \sigma_{S_2}^2}{\left[\bar{Y} + \bar{S}_2\right]^2} \text{ and}$$

$$C_g^2 = \frac{\sigma_x^2 + \sigma_{S_2}^2}{\left[\bar{X} + \bar{S}_2\right]^2}.$$

Case 2: Taking $S_2 = 0$ in equation (1) and equation (2), the multiplicative model proposed by Pollock and Bek (1976) and elaborated in detail by Eichhorn and Hayre (1983) is obtained as $g_M = x\,S_1$ and $h_M = y\,S_1$ with

$$\left(\bar{X}\right)_M = \frac{\left(\bar{G}\right)_M}{\bar{S}_1} \tag{7}$$

and

$$\left(\bar{Y}\right)_M = \frac{\left(\bar{H}\right)_M}{\bar{S}_1} \tag{8}$$

such that

$$\rho_{hg} = \frac{\sigma_{S_1}^2 + \rho_{yx}C_yC_x\left(\sigma_{S_1}^2 + \bar{S}_1^2\right)}{\sqrt{\sigma_y^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{Y}^2\sigma_{S_1}^2}\sqrt{\sigma_x^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{X}^2\sigma_{S_1}^2}},$$

$$\rho_{hz} = \frac{\rho_{yz}\bar{S}_1C_y}{\sqrt{\sigma_y^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{Y}^2\sigma_{S_1}^2}},$$

$$\rho_{gz} = \frac{\rho_{xz}\bar{S}_1C_x}{\sqrt{\sigma_x^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{X}^2\sigma_{S_1}^2}},$$

$$C_h^2 = \frac{\sqrt{\sigma_y^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{Y}^2\sigma_{S_1}^2}}{\left[\bar{Y}\ \bar{S}_1\right]^2} \text{ and}$$

$$C_g^2 = \frac{\sqrt{\sigma_x^2\left(\sigma_{S_1}^2 + \bar{S}_1^2\right) + \bar{X}^2\sigma_{S_1}^2}}{\left[\bar{X}\ \bar{S}_1\right]^2}$$

2.1 Remark. The scrambling variables S_1 and S_2 are such that $E(S_1)=\bar{S}_1$, $E(S_2)=\bar{S}_2$, $V(S_1)=\sigma^2_{S_1}$, $V(S_2)=\sigma^2_{S_2}$.

2.2 Remark. $(\bar{Y})_A$ and $(\bar{Y})_M$ denote population mean of sensitive variable y under additive model and multiplicative model respectively.

2.3 Remark. In order to acquire suitable estimator of sensitive population mean at current wave in two wave successive sampling, an appropriate estimator of population mean of coded response variable $\bar{H}$ need to be investigated and substituted in equation (4). Hence, next section is devoted to investigation of suitable estimator for $\bar{H}$.

2.3 The Proposed Class of estimator on Successive waves

In this section, a class of estimator have been proposed as a convex linear combination of two different classes of estimators based on two independent sample of sizes u and m respectively on second(current) wave.

2.3.1 Class of Estimator based on fresh sample on the second wave

For defining the estimator based on fresh sample at second wave, the following estimators based on sample of size u can be used to estimate population mean of coded response variable:

$$T_{u1} = \bar{h}_u \exp\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right),$$

$$T_{u2} = \bar{h}_u\left(\frac{\bar{Z}}{\bar{z}_u}\right),$$

$$T_{u3} = \bar{h}_u\left(\frac{\bar{Z}}{\bar{z}_u}\right)\exp\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right),$$

$$T_{u4} = \bar{h}_u \exp\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right)\exp\left(\frac{\bar{Z}}{\bar{z}_u}\right),$$

$$T_{u5} = \bar{h}_u \exp\left(\frac{\bar{z}_u}{\bar{Z}}\right),$$

$$T_{u6} = \bar{h}_u \exp\left(\frac{\bar{z}_u}{\bar{Z}}\right)\exp\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right),$$

$$T_{u7} = \bar{h}_u\left(\frac{\bar{z}_u}{\bar{Z}}\right),$$

$$T_{u8} = \bar{h}_u \exp\left(\frac{\bar{Z}}{\bar{z}_u}\right), \text{ etc.,}$$

Following the estimation procedure acquired by Srivastava (1971), a class of estimator for population mean of coded response variable based on fresh sample is defined as:

$$T_u = \bar{h}_u F_u(a)\ ;\ \text{where, } a = \left(\frac{\bar{z}_u}{\bar{Z}}\right) \tag{9}$$

where, $F_u(a)$ is a parametric function, such that $F_u(1)=1$ and the first and second order partial derivatives of F with respect to $'a'$ exists and are known constants at a given point $a = 1$.

2.3.2 Class of Estimators Based on the matched sample at current wave

Based on sample of size $'m'$ at current(second) wave, one may think of ratio type, exponential type and ramification of these estimators as:

$$T_{m1} = \bar{h}^*_m \exp\left(\frac{\bar{g}^*_n}{\bar{g}^*_m}\right)$$

where, $\bar{h}^*_m = \bar{h}_m \exp\left(\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_m}\right)$,

$\bar{g}^*_m = \bar{g}_m \exp\left(\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_m}\right)$ and

$$\bar{g}^*_n = \bar{g}_n \exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m2} = \bar{h}_m\left(\frac{\bar{g}_n}{\bar{g}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_m}\right),$$

$$T_{m3} = \bar{h}_m\left(\frac{\bar{g}_n}{\bar{g}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m4} = \bar{h}_m\left(\frac{\bar{g}_n}{\bar{g}_m}\right)\exp\left(\frac{\bar{z}_n-\bar{z}_m}{\bar{z}_n+\bar{z}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m5} = \bar{h}_m \exp\left(\frac{\bar{g}_n-\bar{g}_m}{\bar{g}_n+\bar{g}_m}\right)\exp\left(\frac{\bar{z}_n-\bar{z}_m}{\bar{z}_n+\bar{z}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m6} = \bar{h}_m \exp\left(\frac{\bar{g}_n-\bar{g}_m}{\bar{g}_n+\bar{g}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m7} = \bar{h}_m\left(\frac{\bar{g}_n}{\bar{g}_m}\right)\exp\left(\frac{\bar{z}_n-\bar{z}_m}{\bar{z}_n+\bar{z}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right),$$

$$T_{m8} = \bar{h}_m\left(\frac{\bar{g}_n}{\bar{g}_m}\right)\exp\left(\frac{\bar{z}_n}{\bar{z}_m}\right)\exp\left(\frac{\bar{z}_n}{\bar{Z}}\right),$$

etc.,

Motivated by the above estimation techniques and following Srivastava (1971), a class of estimators has been proposed for population mean of a coded response variable on current (second) wave based on the matched sample of size m as

$$T_m = \bar{h}_m \; F_m(b, c, d) \tag{10}$$

where $b = \dfrac{\bar{g}_m}{\bar{g}_n},\ c = \dfrac{\bar{z}_m}{\bar{z}_n},\ d = \dfrac{\bar{z}_n}{\bar{Z}}$. Where $F_m(.)$ is a parametric function, such that it follows similar conditions as considered for $F_u(a)$ given in section 2.3.1.

2.3.3 Combined Class of Estimator

By considering convex linear combination of the estimators based on sample size u and m, the final estimator of the population mean of coded response variable $\bar{H}$ is obtained as

$$T = \phi \; T_u + (1 - \phi) \; T_m \tag{11}$$

where $\phi \in [0, 1]$ is a scalar quantity to be chosen suitably.

2.4 Remark. From section 2.3.1 and section 2.3.2, the estimators can be combined to be a member of the final class of estimator T defined in equation (11) as

$$T_k = \phi_k \; T_{u\,i} + (1 - \phi_k) \; T_{m\,j};$$

where $\{i, j, k\} \in \{1, 2, 3, ...\}$.

3. ANALYSIS OF THE PROPOSED CLASS OF ESTIMATORS

3.1 Bias and Mean Squared Error

The properties of the proposed class of estimators are derived under the following large sample approximations :

$$\bar{h}_u = \bar{H}(1 + e_0),\ \bar{h}_m = \bar{H}(1 + e_1),\ \bar{g}_m = \bar{G}(1 + e_2),$$
$$\bar{g}_n = \bar{G}(1 + e_3),\ \bar{z}_u = \bar{Z}(1 + e_4),$$
$$\bar{z}_m = \bar{Z}(1 + e_5),\ \bar{z}_n = \bar{Z}(1 + e_6),$$

such that

$$E(e_j) = 0\ ;\ |e_j| < 1,$$

where, $j = 1, 2, 3, 4, 5$ and 6.

3.1.1 Bias and Mean Squared Error of T_u

The expressions of bias and mean squared error of the class of estimators T_u are derived as

$$T_u = \bar{h}_u \; F_u(a)$$

Using $F_u(1)=1$ and expanding $F_u(a)$ about the point $a = 1$ in second order Taylor series, we have

$$T_u = \bar{h}_u\left[1 + (a - 1)P_1 + (a - 1)^2 P_2 + \ldots\right]$$

$$= \bar{H}\left[1 + e_0 + e_4 P_1 + e_4^2 P_2 + e_0 e_4 P_1 + \ldots\right] \tag{12}$$

$$P_1 = \left.\frac{\partial F_u}{\partial a}\right|_a \text{ and } P_2 = \left.\frac{1}{2}\frac{\partial^2 F_u}{\partial a^2}\right|_a.$$

3.1 Theorem. Bias of the class of estimator T_u to the first order approximations for large N, are obtained as

$$B(T_u) \;=\; \frac{1}{u}\left[\rho_{hz}\; P_1 \;+\; P_2\right] S_h C_h \tag{13}$$

Proof. Retaining terms in equation (12) up to first order of approximations, we have,

$$(T_u - \bar{H}) = \bar{H}\left[\; e_0 + e_4 P_1 + e_4^2 P_2 + e_0 e_4 P_1\right]$$

Taking Expectations on both sides in the above equation and assuming $N \to \infty$ we get bias of T_u up to first order approximation as

$$B(T_u) \;=\; \frac{1}{u}\left[S_h C_h \rho_{hz}\; P_1 \;+\; \bar{H} C_z^2\; P_2\right]$$

3.2 Theorem. Mean squared error of the class of estimator T_u to the first order approximations and for large N, are obtained as

$$M(T_u)_{opt.} = \frac{1}{u}\left[1 - \rho_{h\,z}^2\right] S_h^2 \tag{14}$$

Proof. Retaining terms in equation (12) up to first order of approximations, we have,

$$(T_u \;-\; \bar{H})^2 \;=\; \left(\bar{H}\;[\; e_0 + e_4\; P_1]\right)^2$$

Taking Expectations on both sides in the above equation assuming $N \to \infty$ and optimizing with respect to P_1, we get the optimum mean squared error of T_u up to first order approximation as in equation (14).

3.1.2 Bias and Mean Squared Error of T_m

$$T_m = \bar{h}_m \; F_m(b, c, d)$$

Expanding $F_m(b, c, d)$ about the point $Q = (1, 1, 1)$ in first order Taylor's series and using $F_m(1, 1, 1) = 1$, we get

$$T_m = \bar{h}_m \Big[1 + (b-1)\,Q_1 + (c-1)\,Q_2 + (d-1)\,Q_3 + (b-1)^2 Q_{11} + (c-1)^2 Q_{22} + (d-1)^2 Q_{33} + (b-1)(c-1)\,Q_{12} + (b-1)(d-1)Q_{13} + (c-1)(d-1)Q_{23} + \ldots\Big]$$

$$= \bar{H}\Big[1 + (e_2 - e_3 - e_2e_3 + e_3^2)Q_1 + (e_5 - e_6 - e_5e_6 + e_6^2)Q_2 + (e_6)Q_3 + (e_2 - e_3 - e_2e_3 + e_3^2)^2 Q_{11} + (e_5 - e_6 - e_5e_6 + e_6^2)^2 Q_{22} + (e_6)^2 Q_{33} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_5 - e_6 - e_5e_6 + e_6^2)Q_{12} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_6)Q_{13} + (e_5 - e_6 - e_5e_6 + e_6^2)(e_6)Q_{23} + \ldots\Big] \tag{15}$$

where,

$$Q \equiv (1, 1, 1),\; Q_1 = \left.\frac{\partial F_m}{\partial b}\right|_Q,\; Q_2 = \left.\frac{\partial F_m}{\partial c}\right|_Q,\; Q_3 = \left.\frac{\partial F_m}{\partial d}\right|_Q,$$

$$Q_{11} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial b^2}\right|_Q,\; Q_{22} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial c^2}\right|_Q,\; Q_{33} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial d^2}\right|_Q,$$

$$Q_{12} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial b \partial c}\right|_Q,\; Q_{13} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial b \partial d}\right|_Q,\; Q_{23} = \left.\frac{1}{2}\frac{\partial^2 F_m}{\partial c \partial d}\right|_Q$$

3.3 Theorem. Bias of the class of estimator T_m to the first order approximations are derived as

$$B(T_m) = \left[\left(\frac{1}{m} - \frac{1}{n}\right)\left(\rho_{gz} Q_{12} + \rho_{hg} Q_1 + \rho_{hz} Q_2\right) + \left(\frac{1}{n}\right)\rho_{hz} Q_3\right] S_h C_h \tag{16}$$

Proof.

$$(T_m - \bar{H}) = \bar{H}\Big[(e_2 - e_3 - e_2e_3 + e_3^2)Q_1 + (e_5 - e_6 - e_5e_6 + e_6^2)Q_2 + (e_6)Q_3 + (e_2 - e_3 - e_2e_3 + e_3^2)^2 Q_{11} + (e_5 - e_6 - e_5e_6 + e_6^2)^2 Q_{22} + (e_6)^2 Q_{33} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_5 - e_6 - e_5e_6 + e_6^2)Q_{12} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_6)Q_{13} + (e_5 - e_6 - e_5e_6 + e_6^2)(e_6)\,Q_{23} + \ldots\Big]$$

Taking expectations on both sides in the above equation and assuming $N \to \infty$ we get bias of T_m up to first order approximation as

$$B(T_m) = \bar{H}\left[\left(\frac{1}{m} - \frac{1}{n}\right)\left(\rho_{gz} C_g C_z Q_{12} + \rho_{hg} C_g C_h Q_1 + \rho_{hz} C_h C_z Q_2\right) + \left(\frac{1}{n}\right)\rho_{hz} C_h C_z Q_3\right]$$

3.4 Theorem. Mean squared error of the class of estimators T_m to the first order of approximations is obtained as

$$M(T_m)_{opt.} = \left[\left(\frac{1}{m} - \frac{1}{n}\right)\left((Q_1^*)^2 + (Q_2^*)^2 + 2(Q_1^*)(Q_2^*)\rho_{gz} + 2(Q_1^*)\rho_{hg} + 2(Q_2^*)\rho_{hz}\right) + \frac{1}{m} + \frac{1}{n}\left((Q_3^*)^2 + 2(Q_3^*)\rho_{hz}\right)\right] S_h^2 \tag{17}$$

Proof.

$$(T_m - \bar{H})^2 = \Big(\bar{H}\Big[(e_2 - e_3 - e_2e_3 + e_3^2)Q_1 + (e_5 - e_6 - e_5e_6 + e_6^2)Q_2 + (e_6)Q_3 + (e_2 - e_3 - e_2e_3 + e_3^2)^2 Q_{11} + (e_5 - e_6 - e_5e_6 + e_6^2)^2 Q_{22} + (e_6)^2 Q_{33} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_5 - e_6 - e_5e_6 + e_6^2)Q_{12} + (e_2 - e_3 - e_2e_3 + e_3^2)(e_6)Q_{13} + (e_5 - e_6 - e_5e_6 + e_6^2)(e_6)Q_{23} + \ldots\Big]\Big)^2$$

Taking expectations on both sides in the above equation and assuming $N \to \infty$ we get mean squared error of T_m up to first order approximation as

$$M(T_m)_{opt.} = \left[\left(\frac{1}{m} - \frac{1}{n}\right)\left(Q_1^2 + Q_2^2 + 2Q_1Q_2\rho_{gz} + 2Q_1\rho_{hg} + 2Q_2\rho_{hz}\right) + \frac{1}{m} + \frac{1}{n}\left(Q_3^2 + 2Q_3\rho_{hz}\right)\right] S_h^2$$

which is minimized for

$$Q_1 = \frac{\rho_{hz}\rho_{gz} - \rho_{hg}}{1 - \rho_{gz}^2}\left[Say\left(Q_1^*\right)\right],\quad Q_2 = \frac{\rho_{hz}\rho_{gz} - \rho_{hz}}{1 - \rho_{gz}^2}\left[Say\left(Q_2^*\right)\right]$$

and $Q_3 = -\rho_{hz}\left[Say\left(Q_3^*\right)\right]$

Substituting the optimum values of Q_1^*, Q_2^* and Q_3^* in the above equation, the optimum mean squared error is obtained as in equation (17).

3.1.3 Bias and Mean Squared Error of T

3.5 Theorem. Bias of the class of estimators T to the first order of approximations are obtained as

$$B(T) = \phi B(T_u) + (1 - \phi)\, B(T_m) \tag{18}$$

where $B(T_u)$ and $B(T_m)$ are given in equations (13) and (16) respectively.

Proof. The bias of the class of estimators T is given by

$$B(T) = E\left[T - \bar{H}\right]$$
$$= E\left[\phi\left(T_u - \bar{H}\right) + (1 - \phi)\left(T_m - \bar{H}\right)\right]$$
$$= \phi B\left(T_u\right) + (1 - \phi)\, B\left(T_m\right)$$

Substituting the values of $B\left(T_u\right)$ and $B\left(T_m\right)$ from the equations (13) and (16) in the above equation, we have the expression for the bias of the class of estimators T given in equation (18).

3.6 Theorem. Mean squared error of the class of estimators T to first order of approximations are obtained as

$$M\left(T\right) = \phi^2 M\left(T_u\right)_{opt.} + (1 - \phi)^2\, M\left(T_m\right)_{opt.} \qquad (19)$$

when $M\left(T_u\right)_{opt.}$ and $M\left(T_m\right)_{opt.}$ are given in equations (14) and (17) respectively.

Proof. The Mean squared error of the class of estimators T is given by

$$M(T) = E\left[T - \bar{H}\right]^2$$
$$= E\left[\phi\left(T_u - \bar{H}\right) + (1 - \phi)\left(T_m - \bar{H}\right)\right]^2$$
$$= \phi^2 M\left(T_u\right) + (1-\phi)^2 M\left(T_m\right) + 2\phi(1-\phi)\operatorname{cov}\left(T_u, T_m\right)$$

As T_u and T_m are based on two non-overlapping samples of sizes u and m respectively. So $\operatorname{cov}\left(T_u, T_m\right) = 0.$ By substituting the optimum values of T_u and T_m from the equations (14) and (17) in the above equation, we obtain the expression for the mean squared error of the class of estimators T as in equation (19).

3.2 Minimum Mean Squared Error of the Proposed class of Estimator T

It can be seen that the mean squared error of the class of estimators T is a function of unknown constant ϕ therefore, it is minimized with respect to ϕ and subsequently the optimum value of ϕ is obtained as

$$\phi_{opt.} = \frac{M\left(T_m\right)_{opt.}}{M\left(T_u\right)_{opt.} + M\left(T_m\right)_{opt.}} \qquad (20)$$

Substituting the value of $\phi_{opt.}$ from equation (20) in equation (19), we get the optimum mean squared error of the class of estimator T as

$$M\left(T\right)_{opt.} = \frac{M\left(T_u\right)_{opt.} \times M\left(T_m\right)_{opt.}}{M\left(T_u\right)_{opt.} + M\left(T_m\right)_{opt.}} \qquad (21)$$

Further, substituting the values $M\left(T_u\right)_{opt.}$ and $M\left(T_m\right)_{opt.}$ from equations (14) and equation (17) in equation(21), the simplified values of $M\left(T\right)_{opt.}$ is derived as

$$M\left(T\right)_{opt.} = \frac{J_1\mu - J_2}{K_3\mu^2 - \mu J_3 - K_1}\left(\frac{S_h^2}{n}\right) \qquad (22)$$

where,

$K_1 = 1 - \rho_{hz}^2,\ K_2 = 1 + Q_1^2 + Q_2^2 + 2Q_1Q_2\rho_{gz} + 2Q_1\rho_{hg} + 2Q_2\rho_{hz},$
$K_3 = Q_3^2 + 2Q_3\rho_{hz} - \left(K_2 - 1\right),\ J_1 = K_1K_3,\ J_2 = K_1\left(K_2 + K_3\right)$ *and*
$J_3 = K_2 + K_3 - K_1.$

3.3 Optimum rotation rate for the proposed class estimator

Since the mean squared error of the proposed class of estimator $M\left(T\right)_{opt.}$ is the function of μ, which is the rotation rates or the fractions of sample to be drawn afresh at current wave. To estimate population mean with maximum precision and minimum cost, the mean squared error of the estimator T is derived in equation (22) have been optimized with respect to μ. Hence, optimum fraction of sample to be drawn afresh say $\hat{\mu}_f$ have been obtained for the estimator $M\left(T\right)_{opt.}$ and are given as

$$\hat{\mu}_f = \min\left\{\frac{I_2 + \sqrt{I_2^2 - I_1I_3}}{I_1},\ \frac{I_2 - \sqrt{I_2^2 - I_1I_3}}{I_1}\right\}$$

such that $\hat{\mu}_f \in [0,\ 1]$ (23)

$I_1 = J_1K_3,\ I_2 = J_2K_3$ and $I_3 = K_1J_1 + J_3J_2.$

Substituting the optimum value of $\hat{\mu}_f$ in equation (22), we have the optimum value of the mean squared error of the class of estimators T with respect to ϕ as well as μ as,

$$M\left(T\right)_{opt.^*} = \frac{J_1\hat{\mu}_f - J_2}{K_3\hat{\mu}_f^2 - \hat{\mu}_f J_3 - K_1}\left(\frac{S_h^2}{n}\right). \qquad (24)$$

4. SOME ESTIMATORS

The general successive sampling estimator proposed by Jessen (1942) when modified to work for sensitive population mean estimation given by T_J when no additional non-sensitive auxiliary variable is used at any occasion and a member of its own class, T_1 when non-sensitive auxiliary information is used have been listed with their properties to be used for comparison purpose of final estimator. They are given as

$$T_J = \phi_J \bar{h}_u + (1-\phi_J)\left[\bar{h}_m + K^*(\bar{g}_n - \bar{g}_m)\right]$$

and

$$T_1 = \phi_1 T_{u1} + (1 - \phi_1) T_{m1} ;$$

where,

$$T_{u1} = \bar{h}_u \exp\left(\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u}\right),\ T_{m1} = \bar{h}_m^* \left(\frac{\bar{g}_n^*}{\bar{g}_m^*}\right)$$

where,

$$\bar{h}_m^* = \bar{h}_m \exp\left(\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\right),\ \bar{g}_m^* = \bar{g}_m \exp\left(\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_m}\right),\ \text{and}$$

$\bar{g}_n^* = \bar{g}_n \exp\left(\frac{\bar{Z} - \bar{z}_n}{\bar{Z} + \bar{z}_n}\right)$ $\{\phi_J, \phi_1\} \in [0, 1]$ are unknown constants and K^* is suitably chosen.

The minimum mean squared error of T_J and T_1 are respectively given as

$$M(T_J)_{opt.^*} = \left(\frac{1-\hat{\mu}_1 \rho_{hg}^2}{1-\hat{\mu}_1^2 \rho_{hg}^2}\right)\left(\frac{S_h^2}{n}\right) \quad (25)$$

and

$$M(T_1)_{opt.^*} = \frac{J_{11}\hat{\mu}_2 - J_{12}}{K_{13}\hat{\mu}_2^2 - \hat{\mu}_2 J_{13} - K_{11}}\left(\frac{S_h^2}{n}\right). \quad (26)$$

where,

$$\hat{\mu}_1 = \min\left\{\frac{1 + \sqrt{1 - \rho_{hg}^2}}{\rho_{hg}^2},\ \frac{1 - \sqrt{1 - \rho_{hg}^2}}{\rho_{hg}^2}\right\}$$

$$\text{such that } \hat{\mu}_1 \in [0, 1] \quad (27)$$

and

$$\hat{\mu}_2 = \min\left\{\frac{I_{12} + \sqrt{I_{12}^2 - I_{11}I_{13}}}{I_{11}},\ \frac{I_{12} - \sqrt{I_{12}^2 - I_{11}I_{13}}}{I_{11}}\right\}$$

$$\text{such that } \hat{\mu}_2 \in [0, 1] \quad (28)$$

where,

$K_{11} = \frac{5}{4} - \rho_{hz}$, $K_{12} = 2 - 2\rho_{hg}$, $K_{13} = \frac{5}{4} - 2\rho_{hg} - 2 - \rho_{hz}$, $I_{11} = J_{11}K_{13}$, $I_{12} = J_{12}K_{13}$, $I_{13} = K_{11}J_{11} + J_{13}J_{12}$, $J_{11} = K_1 K_{13}$, $J_{12} = K_{11}(K_{12} + K_{13})$ and $J_{13} = K_{12} + K_{13} - K_{11}$.

5. ESTIMATORS FOR SENSITIVE POPULATION MEAN AT CURRENT WAVE

Substituting the population mean of coded response variable $\bar{H}$ in equation (4) by its estimators T_J, T_1 and T respectively, the corresponding estimators for sensitive population mean at current wave $\hat{\bar{Y}}_J$, $\hat{\bar{Y}}_1$ and $\hat{\bar{Y}}$ have been obtained and presented in Table 1. The estimators $\hat{\bar{Y}}_J$, $\hat{\bar{Y}}_1$ and $\hat{\bar{Y}}$ are biased, so the mean squared error of sensitive population mean estimators $\hat{\bar{Y}}_J$, $\hat{\bar{Y}}_1$ and $\hat{\bar{Y}}$ have been computed and are presented in Table 1.

6. EFFICIENCY COMPARISON

In order to compare the proposed class of estimators $\hat{\bar{Y}}$, with respect to the estimators $\hat{\bar{Y}}_J$ and $\hat{\bar{Y}}_1$ respectively, their percent relative efficiencies have been computed and are presented below

$$E_1 = \frac{M\left[\left(\hat{\bar{Y}}_J\right)\right]}{M\left[\left(\hat{\bar{Y}}\right)\right]} \times 100 \text{ and } E_2 = \frac{M\left[\left(\hat{\bar{Y}}_1\right)\right]}{M\left[\left(\hat{\bar{Y}}\right)\right]} \times 100 \quad (29)$$

6.1 Remark. The two scrambling variables S_1 and S_2 used to perturb the true response through the scrambled response model may follow any distribution. In this paper for the considered model on two wave successive sampling, following Pollock and Bek (1976) and Eichhorn and Hayre (1983), we consider scrambling variable S_1 and S_2 to follow normal distribution such that $S_1 \sim N(1, 0.6)$ and $S_2 \sim N(0, 1)$

7. NUMERICAL DEMONSTRATION

To judge the performance of the proposed class of estimators, the following numerical illustrations has

Table 1. Sensitive population mean estimator and their Mean squared error (MSE)

S.No.	Estimators	MSE
1	$\left(\hat{\bar{Y}}_J\right) = \frac{T_J - \bar{S}_2}{\bar{S}_1}$	$M\left[\left(\hat{\bar{Y}}_J\right)\right] = \frac{M(T_J)_{opt.^*}}{\bar{S}_1^2}$
2	$\left(\hat{\bar{Y}}_1\right) = \frac{T_1 - \bar{S}_2}{\bar{S}_1}$	$M\left[\left(\hat{\bar{Y}}_1\right)\right] = \frac{M(T_1)_{opt.^*}}{\bar{S}_1^2}$
3	$\left(\hat{\bar{Y}}\right) = \frac{T - \bar{S}_2}{\bar{S}_1}$	$M\left[\left(\hat{\bar{Y}}\right)\right] = \frac{M(T)_{opt.^*}}{\bar{S}_1^2}$

been worked out for a completely known population with population parameters as follows:

Population - I

$N = 51,\ n = 20,\ S_x^2 = 4.3451\times10^6,\ S_y^2 = 4.1604\times10^6,$
$S_z^2 = 4.2152\times10^6,\ \bar{X} = 1923.3,\ \bar{Y} = 1947.8,\ \bar{Z} = 1923.3,$
$\rho_{yx} = 0.7,\ \rho_{xz} = 0.7,\ \rho_{yz} = 0.7.$

The optimum values of $\hat{\mu}$'s for $\hat{\hat{Y}}$, $\hat{\hat{Y}}_J$, $\hat{\hat{Y}}_1$ and percent relative efficiencies E_1 and E_2 have been computed for the above data and are presented in Table 2.

Table 2. Percent relative efficiency of $\hat{\hat{Y}}$ with respect to $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$ at their respective optimum conditions for Population-I

$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_f$		E_2
0.6072	0.5536	0.5751	129.7108	103.9351

7.1 Case Study: Usage of Drugs (Cigarette, Alcohol, Gutkha, Paan Masala etc.)

Population-II: For practicing the literal feasibility of the proposed estimators $\hat{\hat{Y}}$, $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$, a case study has been designed for two waves and data have been collected from 315 under graduate students of a College (University of Delhi), India through a survey conducted on two successive waves. For convenience 315 random numbers S_1 and S_2 have been generated assuming $S_1 \sim N(1, 0.6)$ and $S_2 \sim N(0,1)$. Following sensitive and non-sensitive variables of the interest have been considered:

x : Average monthly expenditure on drug usage in July, 2015, by the i[th] student.

y : Average monthly expenditure on drug usage in April, 2016, by the i[th] student.

z : Average monthly pocket money from all sources in July, 2015 of the i[th] student.

And hence the scrambled response was collected from the respondents in the form of $g = x\left(S_1 + \frac{S_2}{x}\right)$ and $h = y\left(S_1 + \frac{S_2}{y}\right)$.

Therefore, the optimum rotation rate, percent relative efficiencies of the proposed class of estimator $\hat{\hat{Y}}$ with respect to $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$ are obtained and are listed in Table 3.

Table 3. Percent relative efficiency of $\hat{\hat{Y}}$ with respect to $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$ at their respective optimum conditions for Population-II

$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_f$	E_1	E_2
0.6935	0.6432	0.6562	142.9261	103.0541

8. MONTE CARLO SIMULATION STUDY

Using Monte carlo simulation for the above said data (Population-II), the simulation study has been carried out by considering 5000 different samples. The simulated percent relative efficiencies E_{S_1} and E_{S_2} for the proposed class of estimator when compared with respect to the estimator $\hat{\hat{Y}}_J$ and the estimator $\hat{\hat{Y}}_1$ respectively have been computed for

Table 4. Simulation results for Population-II

ϕ	Percent Relative Efficiency	Set I	Set II	Set III
0.1	E_{S_1}	146.1175	144.2562	143.4881
	E_{S_2}	108.2009	109.8896	109.6529
0.2	E_{S_1}	146.1322	144.2804	143.5617
	E_{S_2}	108.1823	109.8596	109.5713
0.3	E_{S_1}	146.1353	144.2709	143.5538
	E_{S_2}	108.1783	109.8714	109.5800
0.4	E_{S_1}	146.1338	143.2810	143.5605
	E_{S_2}	108.1803	109.8588	109.5726
0.5	E_{S_1}	146.1361	144.2714	143.5736
	E_{S_2}	108.1774	109.8707	109.5581
0.6	E_{S_1}	146.1338	144.2607	143.5579
	E_{S_2}	108.1803	109.8840	109.5755
0.7	E_{S_1}	146.1273	144.2708	143.5326
	E_{S_2}	108.1885	109.8715	109.6036
0.8	E_{S_1}	146.262	144.2672	143.5371
	E_{S_2}	108.1898	109.8760	109.5985
0.9	E_{S_1}	146.1213	144.2673	143.5531
	E_{S_2}	108.1961	109.8564	109.5808

many combinations of constants for varying ϕ which are termed as different sets and are given below. The simulation results are presented in Table 4.

Set I : n = 45, u = 12, m = 33,

Set II : n = 45, u = 18, m = 27 and

Set III : n = 45, u = 27, m = 18.

9. DIRECT METHOD

It is quite evident that for assuring the privacy of respondents, some cost has to be paid in terms of loss in efficiency as compared to direct method of questioning.

Hence, in this section, the proposed class of estimators $\hat{\hat{\mathrm{Y}}}$ have been compared with its respective direct estimator $\overline{\mathrm{Y}}_{\mathrm{D}}$ i.e., when no scrambling technique is used, which is given as

$$\overline{Y}_D = \chi T_{ud} + (1-\chi) T_{md}; \ \chi \in [0,1] \tag{30}$$

where,

$$\mathrm{T}_{\mathrm{ud}} = \overline{\mathrm{y}}_{\mathrm{u}} \mathrm{F}_{\mathrm{u}}(\mathrm{a}) \ ; \text{ where, } \mathrm{a} = \left(\frac{\overline{\mathrm{z}}_{\mathrm{u}}}{\overline{\mathrm{Z}}}\right) \tag{31}$$

$$\mathrm{T}_{\mathrm{md}} = \overline{\mathrm{y}}_{\mathrm{m}} \ \mathrm{F}_{\mathrm{md}}\left(\mathrm{b}^*, \ \mathrm{c}, \ \mathrm{d}\right) \tag{32}$$

where,

$\mathrm{b}^* = \frac{\overline{\mathrm{x}}_{\mathrm{m}}}{\overline{\mathrm{x}}_{\mathrm{n}}}$, $\mathrm{c} = \frac{\overline{\mathrm{z}}_{\mathrm{m}}}{\overline{\mathrm{z}}_{\mathrm{n}}}$, $\mathrm{d} = \frac{\overline{\mathrm{z}}_{\mathrm{n}}}{\overline{\mathrm{Z}}}$ and $\mathrm{F}_{\mathrm{md}}\left(\mathrm{b}^*, \mathrm{c}, \mathrm{d}\right)$ is a parametric function of $\left(\mathrm{b}^*, \mathrm{c}, \mathrm{d}\right)$, such that $\mathrm{F}_{\mathrm{md}}(1, 1, 1) = 1$ and $\mathrm{F}_{\mathrm{md}}\left(\mathrm{b}^*, \mathrm{c}, \mathrm{d}\right)$ satisfies conditions similar to those given for $\mathrm{F}_{\mathrm{u}}(\mathrm{a})$.

The minimum mean squared error of the class of estimator $\overline{\mathrm{Y}}_{\mathrm{D}}$ to the first order approximations is given as

$$M\left(\overline{Y}_D\right)_{opt.^*} = \frac{J_{d1}\hat{\mu}_d - J_{d2}}{K_{d3}\hat{\mu}_d^2 - \hat{\mu}_d J_{d3} - K_{d1}} \left(\frac{S_y^2}{n}\right) \tag{33}$$

where,

$$K_{d1} = 1 - \rho_{yz}^2,$$

$$K_{d2} = 1 + Q_{d1}^2 + Q_{d2}^2 + 2Q_{d1}Q_{d2}\rho_{xz} + 2Q_{d1}\rho_{yx} + 2Q_{d2}\rho_{yz},$$

$$K_{d3} = Q_{d3}^2 + 2Q_{d3}\rho_{yz} - \left(K_{d2} - 1\right), \ Q_{d1} = \frac{\rho_{yz}\rho_{xz} - \rho_{yx}}{1 - \rho_{xz}^2},$$

$$Q_{d2} = \frac{\rho_{yx}\rho_{xz} - \rho_{yz}}{1 - \rho_{xz}^2}, \ Q_{d3} = -\rho_{yz}, \ J_{d1} = K_{d1}K_{d3},$$

$$J_{d2} = K_{d1}\left(K_{d2} + K_{d3}\right) \text{ and } J_{d3} = K_{d2} + K_{d3} - K_{d1}.$$

with,

$$\hat{\mu}_{\mathrm{d}} = \min\left\{\frac{\mathrm{I}_{\mathrm{d2}} + \sqrt{\mathrm{I}_{\mathrm{d2}}^2 - \mathrm{I}_{\mathrm{d1}}\mathrm{I}_{\mathrm{d3}}}}{\mathrm{I}_{\mathrm{d1}}}, \ \frac{\mathrm{I}_{\mathrm{d2}} - \sqrt{\mathrm{I}_{\mathrm{d2}}^2 - \mathrm{I}_{\mathrm{d1}}\mathrm{I}_{\mathrm{d3}}}}{\mathrm{I}_{\mathrm{d1}}}\right\}$$

$$\text{such that } \hat{\mu}_{\mathrm{d}} \in [0, 1] \tag{34}$$

where,

$$\mathrm{I}_{\mathrm{d1}} = \mathrm{J}_{\mathrm{d1}}\mathrm{K}_{\mathrm{d3}}, \ \mathrm{I}_{\mathrm{d2}} = \mathrm{J}_{\mathrm{d2}}\mathrm{K}_{\mathrm{d3}} \text{ and } \mathrm{I}_{\mathrm{d3}} = \mathrm{K}_{\mathrm{d1}}\mathrm{J}_{\mathrm{d1}} + \mathrm{J}_{\mathrm{d3}}\mathrm{J}_{\mathrm{d2}}.$$

In order to judge the scrambling model effect, the percent relative efficiency of the proposed class of estimator of sensitive population mean $\hat{\hat{\mathrm{Y}}}$ with respect to its direct method have been obtained as

$$\mathrm{E}_{\mathrm{D}} = \frac{\mathrm{M}\left[\overline{\mathrm{Y}}_{\mathrm{D}}\right]_{\mathrm{opt.}^*}}{\mathrm{M}\left[\left(\hat{\hat{\mathrm{Y}}}\right)\right]} \times 100 \tag{35}$$

The above percent relative efficiency have been checked out for population-I as well as Population-II and are presented in Table 5.

Table 5. Percent relative efficiency of $\overline{\mathrm{Y}}_{\mathrm{D}}$ with respect to $\hat{\hat{\mathrm{Y}}}$ at their respective optimum conditions for Population-Iand Population-II

Population	$\hat{\mu}_{\mathrm{D}}$	E_{D}
I	0.5232	102.7517
II	0.6296	78.7496

10. RESPONDENTS PRIVACY PROTECTION

While dealing with the sensitive issues on successive waves, there are certain hesitations which respondents feel while responding to the interviewer and hence they tend to delusive. Therefore to minimize this personal bias and falsification, scrambling is applied to ensure respondent's privacy. So following Diana and Perri (2011), normalized measure of privacy protection on first and second wave are computed under the considered model and are given as

Privacy veil at previous wave under considered model for Population-II:

$$N_0(x) = 1-\rho^2_{x.\,g\,z} = 1-\left(\frac{\rho^2_{x\,z}+\rho^2_{x\,g}-2\rho_{xz}\ \rho_{xg}\ \rho_{zg}}{1\ -\ \rho^2_{zg}}\right) = 0.1839 \tag{36}$$

Privacy veil at current wave under considered model for Population-II:

$$N_0(y) = 1-\rho^2_{y.\,h\,z} = 1-\left(\frac{\rho^2_{yz}+\rho^2_{yh}-2\rho_{yz}\ \rho_{yh}\ \rho_{zh}}{1-\rho^2_{zh}}\right) = 0.1601 \tag{37}$$

Where $\rho_{x.\,g\,z}$ and $\rho_{y.\,h\,z}$ are the multiple correlation coefficients of regression line X on G and Z and Y on H and Z respectively. It is studied that value of $N_0(x)\ [N_0(y)]$ closer to 1, reflects that maximum veil of privacy is present for the respondent. On the other hand, as the value of $N_0(x)\ [N_0(y)]$ approaches to 0, this implies that the privacy declines and least veil will be provided to the respondent in terms of privacy.

11. DISCUSSION OF RESULTS

1. Following observations can be drawn from Table 1 and Table 2:
   (a) It can be seen that the proposed class of estimators is performing better than $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$ for considered Population-I as well as Population-II**.**
   (b) The optimum fraction of sample to be drawn afresh exists for the proposed estimators.
   (c) It is further observed that $E_1 > E_2$, this implies that the proposed class of estimators $\hat{\hat{Y}}$ proves more efficient when compared with the estimator $\hat{\hat{Y}}_J$ than with the estimator $\hat{\hat{Y}}_1$ in terms of percent relative efficiency for considered Population-I and Population-II.
2. From Table 3, it can be observed that, when the value of ϕ increases the simulated percent relative efficiency E_{S_1} also increases for the considered Set-I,Set-II and Set-III respectively, whereas the simulated percent relative efficiency E_{S_2} decreases for the considered Set-I, Set-II and Set-III.
3. From Table 4, it is clear that the proposed class of estimator when compared with the direct method, the percent relative efficiency is coming out to be less than or equal to 100, this shows that cost has to be paid in terms of loss in efficiency for scrambling the data.
4. Normalized measure of privacy has been calculated under the considered model at both waves. Although, $N_0(x)\ [N_0(y)]$ are not statistically strong value for justifying the privacy provided by considered scrambled response technique but are built as a simple tool which may provide some veil of privacy to the respondent.

12. CONCLUSION

The estimation of sensitive population mean in two wave successive sampling is possible. The considered scrambled response model together with proposed class of estimators for estimating sensitive population mean at current wave is fruitful in terms of precision when compared to $\hat{\hat{Y}}_J$ and $\hat{\hat{Y}}_1$. It is further vindicated that more efficiency is obtained when proposed class of estimator is compared with $\hat{\hat{Y}}_J$ than with $\hat{\hat{Y}}_1$. This shows that the use of non-sensitive auxiliary variable together with considered scrambled response model is effective in estimating sensitive population mean. Therefore, the proposed class of estimators for estimating sensitive population mean may be recommended for future use in practise.

ACKNOWLEDGEMENTS

Authors are thankful to anonymous reviewer for their careful reading and constructive suggestions which led to improvement over an earlier version of the paper. Authors are also thankful to SERB, New Delhi, India for providing the financial assistance to carry outthe present work.

REFERENCES

Arnab, R. (2011), Alternative estimators for randomized response techniques in multi-character surveys, Comm. Statist.-Theo. Meth., **40**, 1839-1848.

Arnab R, Singh S. and North, D. (2012), Use of two decks of cards in randomized response techniques for complex survey designs, Comm. Statist.- Theo. Meth., **41**, 3198-3210.

Arnab R, and Singh S. (2013), Estimation of mean of sensitive characteristics for successive sampling, Comm. Statist.-Theo. Meth., **42**, 2499-2524.

Christofides, T. C. (2003), A generalized randomized response technique, Metrika, **57(2)**, 195-200.

Christofides, T. C. (2005), Randomized response in stratified sampling, J. statist. Plann. Infer. **128(1)**, 303-310.

Cochran, W. C. (1977), Sampling Techniques, John Wiley & Sons, New York, NY, USA, 3rd edition.

Dianna, G. and Perri, P. F. (2011), A class of estimators for quantitative sensitive data, Stat. Pap. **52**, 633-650.

Eichhorn, B.H. and Hayre, L.S. (1983), Scrambled randomized response method for obtaining sensitive quantitative data. J. Statist. Plann. Infer. **7**, 307-316.

Feng, S. and Zou, G. (1997), Sample rotation method with auxiliary variable, Comm. Stat.- Theo. Meth. **26**, 1497-1509.

Greenberg, B.G., Kubler, R.R., Horvitz, D.G. (1971), Application of RR technique in obtaining quantitative data, J. Amer. Statist. Asso., **66**, 243-250.

Horvitz, D.G., Shah, B.V. and Simmons, W.R. (1967), The unrelated question randomized response model, Proc. Soci. Statist., Amer. Statist. Asso., 65-72.

Hussain, Z. and Al-Zahrani, B. (2016): Mean and sensitivity estimation of a sensitive variable through additive scrambling. Comm. Stat. – Theo. Meth., **45**, 182-193.

Jessen, R. J. (1942), Statistical investigation of a sample survey for obtaining farm facts, Iowa Agri. Exp. Stat. Road Bull. **304**, 1-104.

Kim, J.-M. and Elam, M. E. (2007), A stratified unrelated question randomized response model, Statis. Pap., **48(2)**, 215-233.

Koyuncu, N., Gupta, S. and Sousa, R., (2014), Exponential-type estimators of the mean of a sensitive variable in the presence of a non-sensitive auxiliary information. Comm. Stat.-Sim. Comp., **43**, 1583-1594.

Mukherjee,R. (2016), Optimal Randomized Response Revisited, Chapter-20, Hand book of statistics. Vol.34, Elsevier.

Patterson, H.D. (1950), Sampling on successive occasions with partial replacement of units, J. Royal Statis. Soci. **12**, 241-255.

Pollock, K. H. and Bek, Y. (1976), A comparison of three randomized response models for quantitative data. J. Amm. Sta. Asso., **71**, 884-886.

Priyanka K. and Mittal, R. (2014), Effective rotation patterns for median estimation in successive sampling. Statis. Trans. **15**, 197-220.

Priyanka, K., Mittal R. and Min-Kim, J. (2015), Multivariate rotation design for population mean in sampling on successive occasions. Comm. Statis. Appli. Meth. **22**, 445-462.

Priyanka K. and Mittal, R. (2015a), Estimation of population median in two-occasion Rotation Sampling. J. Stat. App. Prob. Lett. **2**, 205-219.

Priyanka, K. and Mittal, R. (2015b), A class of estimators for population median in two occasion rotation sampling. HJMS, **44**, 189-202.

Srivastava,S.K., (1971), A generalized estimator for the mean of a finite population using multi-auxiliary information, Journal of American Statistical Association, **Vol.66, issue-334**, 404-407.

Sen, A. R. (1973), Theory and application of sampling on repeated occasions with several auxiliary variables. Biometrics, **29**, 381-385.

Saha, A. (2007), A simple randomized response technique in complex surveys. Metron, LXV, 59-66.

Singh, G. N. and Priyanka, K. (2008), Search of good rotation patterns to improve the precision of estimates at current occasion, Comm. Stat. – Theo. Meth. **37**, 337 - 348.

Singh, S. and Sedory, S. A. (2012), A true simulation study of three estimators at equal protection of respondents in randomized response sampling, StatisticaNeer., **75(4)**, 442-451.

Sihm, J.-S. and Gupta, S. (2015), A two-stage binary optional randomized response model, Comm. Statist.-Sim. Comp., **44**, 2278-2296.

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

Warner, S. L. (1971), The linear randomized response model, J. Amer. Statist. Asso. **66**, 884-888.

Wu, J.-W., Tian, G.-L.and Tang, M.-L. (2008), Two new models for survey sampling with sensitive characteristics: Design and Analysis, Metrika, 67, 251-263.

Yan , Z., Wang, J. and Lai, J. (2009), An efficiency and protection degree-based comparison among the quantitative randomized response strategies, Comm. Statist.- Theo. Meth., **38**, 400-408.

Yu, B., Jin, Z., Tian, J. and Gao, G. (2015), Estimation of sensitive proportion by randomized response data in successive sampling, Compu. Mathemat. Meth. Med. DOI: 10.1155/2015/172918.