Item sum techniques for quantitative sensitive estimation on successive occasions

Kumari Priyanka¹,a, Pidugu Trisandhya

“Department of Mathematics, Shivaji College, University of Delhi, India

Abstract

The problem of the estimation of quantitative sensitive variable using the item sum technique (IST) on successive occasions has been discussed. IST difference, IST regression, and IST general class of estimators have been proposed to estimate quantitative sensitive variable at the current occasion in two occasion successive sampling. The proposed new estimators have been elaborated under Trappmann et al. (Journal of Survey Statistics and Methodology, 2, 58–77, 2014) as well as Perri et al. (Biometrical Journal, 60, 155–173, 2018) allocation designs to allocate long list and short list samples of IST. The properties of all proposed estimators have been derived including optimum replacement policy. The proposed estimators have been mutually compared under the above mentioned allocation designs. The comparison has also been conducted with a direct method. Numerical applications through empirical as well as simplistic simulation has been used to show how the illustrated IST on successive occasions may venture in practical situations.

Keywords: sensitive variable, successive occasions, class of estimators, population mean, variance, bias, mean squared error, optimum matching fraction

1. Introduction

Data gathering on sensitive, incriminating, stigmatizing or too personal issues are an avowedly baffling task. The three possibilities: under reporting, over reporting and refusal to answer the sensitive questions are prevalent. Generally, socially undesirable characteristics such as drug addiction, tax evasion, plagiarism, criminal conviction, and unauthorized natural resource use are likely to be under reported whereas socially desirable characteristics such as energy conservation, reducing pollution, ecological and biological conservation, and participation in elections are likely to be over reported. However some may refuse to report because of social stigma or fear of privacy violations. All these phenomenon provoke non sampling errors that are difficult to deal with and can seriously damage the validity of the analysis. In order to overcome mis-reporting on sensitive issues, many data collection strategies have been developed to elicit a more honest response from respondents by increasing the obscurity of the survey process and ensuring privacy protection.

One important method of survey addressing sensitive issue is the indirect questioning method. The indirect questioning technique may be classified in three different categories: the randomized response technique (RRT), the item count technique (ICT), and the non RRT. The RRT was initiated by Warner (1965), the ICT was originally proposed by Miller (1984) for binary variables to estimate the prevalence of a stigmatizing behavior within the population. However the non RRT was initiated recently by Tian and Tang (2014).

¹Corresponding author: Department of Mathematics, Shivaji College, University of Delhi, New Delhi 110 027, India.
E-mail: priyanka.ism@gmail.com

Published 31 March 2019 / journal homepage: http://csam.or.kr
© 2019 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.
In this paper we focus on the generalization of second technique the ICT. The ICT is used in surveys that require the study of qualitative variable. However, Chaudhuri and Christofides (2013) proposed a generalization to ICT for estimating a quantitative sensitive variable. Trappmann et al. (2014) named this technique as item sum technique (IST). Perri et al. (2018) discussed the optimal sample size allocation in the IST. Further enhancement in IST literature can be seen in Hussain et al. (2015), Rueda et al. (2017).

In many fields of applied research it has been observed that data needs to be gathered for a sensitive variables which can change over time. The statistical tool recommended for such a situation is successive sampling. Jessen (1942) initiated the idea of sampling same population over time with partial replacement of units called successive sampling. However, an analysis of a sensitive variable in successive sampling has been initiated by Arnab and Singh (2013). They applied RRT on successive occasions. Additional literature addressing sensitive issue over successive occasions can be seen in Yu et al. (2015), Naeem and Shabbir (2016), Singh et al. (2017), Priyanka et al. (2018), and Priyanka and Trisandhya (2018). These researchers focused on a scrambled response technique or RRT to handle sensitive issues on successive occasions. However, the IST is now emerging as an alternative technique to deal with sensitive issues. Hence, an attempt has been made in the present work to use IST in successive sampling to estimate a sensitive population mean. To the best of our knowledge this is an initial attempt and will contribute another useful method in successive sampling literature.

Therefore, the IST difference, the IST regression and also the IST general class of estimators have been proposed in the present work to estimate a sensitive population mean in two occasion successive sampling. All these proposed estimators have been discussed under Trappmann et al. (2014) allocation and Perri et al. (2018) optimal allocation designs. Detail properties including optimum replacement strategies are elaborated. The proposed methods have been compared mutually as well as with the corresponding direct method. Numerical applications through empirical simulation and simplistic simulation show how the illustrated IST on successive occasions may develop in practical situations. Finally some concluding remarks have been forwarded.

2. The item sum technique

The well-known technique in sensitive characteristics estimation is the ICT; however, the ICT is generally applicable for dichotomous (qualitative) variables only. Hence, the ICT was generalized by Chaudhuri and Christofides (2013) that can be used to estimate the quantitative sensitive variable. Later Trappmann et al. (2014) named this generalized version of ICT as IST and used it to estimate some quantitative sensitive variable. The algorithm for the IST is as follows.

From a random sample \( (s) \), two random sub-samples \( (s_{ll} \) and \( s_{sl} ) \) are generated. The sub-sample \( s_{ll} \), is confronted with a long list (LL) of items containing the sensitive question and a number of innocuous/non-sensitive questions. However the respondents in sub-sample \( s_{ll} \) have been given a short list (SL) of items containing only the innocuous questions present in the LL sample. Respondents in each sample are asked to report the total score of all items given to them, without disclosing the individual scores for the items. The mean difference of the answers between the \( s_{ll} \) and \( s_{sl} \) is used as an unbiased estimator of the population mean of the sensitive variable. All sensitive and innocuous variables should be quantitative in nature and measured on the same scale as that of the sensitive variable in the IST.

The decisive point in the IST is how to split the total sample in to the LL sample and SL sample. Trappmann et al. (2014) allocated the same number of units to each sample irrespective of the vari-
Consider a finite population $P$ consisting of $N$ identifiable units for sampling over two successive occasions. Let $x$ denote the quantitative sensitive variable at the first occasion which changes to $y$ at second occasion. Similarly let $t_1$ be the non-sensitive/innocuous variable at the first occasion which changes to $t_2$ at the second occasion. Assume that $x_i, y_i, t_1, t_2$ denotes the value of $x, y, t_1,$ and $t_2$ respectively on the unit $i \in P$. To estimate the population mean of quantitative sensitive variable $\bar{y}$ at current occasion using the IST, the sampling design is proposed as:

At first occasion a sample of size $n$ is drawn using simple random sample without replacement (SRSWOR) which has been split to $s_{ll}$ and $s_{ul}$ samples called the LL-sample and SL-sample respectively. Now, at the second occasion considering the partial overlap case, two independent samples are considered, one is a matched sample of size $m = nll$ drawn as SRSWOR sub-sample from sample size $n$ at first occasion and second is a fresh sample of size $u = (n - m) = nul$, which is drawn afresh at current occasion. Further, the samples of sizes $m$ and $u$ are split into corresponding LL-sample and SL-samples as $s_{lll}, s_{ul}, s_{ll},$ and $s_{ull}$ respectively. The response obtained from the respondents on two occasions and the corresponding IST estimate based on different samples are presented in Table 1.

### 3. Proposed IST frame work in successive sampling design

**Table 1: Response received under item sum technique (IST)**

| Occasion | Sample size | Response received | IST estimate |
|----------|-------------|-------------------|--------------|
| I        | $n$         | $z_{1i} = x_i + t_{1i}$, if $i \in s_{ull}$, $t_{1i}$, if $i \in s_{ul}$, | $\hat{\mu} = \bar{z}_{1ll} - \bar{z}_{1ul}$ |
|          | $m$         | $z_{1i} = x_i + t_{1i}$, if $i \in s_{ull}$, $t_{1i}$, if $i \in s_{ul}$, | $\hat{\mu} = \bar{z}_{1ll} - \bar{z}_{1ul}$ |
| II       | $m$         | $z_{2i} = y_i + t_{2i}$, if $i \in s_{ull}$, $t_{2i}$, if $i \in s_{ul}$, | $\hat{\mu} = \bar{z}_{2ll} - \bar{z}_{2ul}$ |
|          | $u$         | $z_{2i} = y_i + t_{2i}$, if $i \in s_{ull}$, $t_{2i}$, if $i \in s_{ul}$, | $\hat{\mu} = \bar{z}_{2ll} - \bar{z}_{2ul}$ |

$z_{ji}, j = 1, 2$ denote the observed response at first and second occasion respectively on the $j^{th}$ observation.

$\bar{z}_{jll}, j = 1, 2; i \in \{n, m, u\}$ denote the mean of $z_j$ in the long list (LL) samples.

$\bar{z}_{jul}, j = 1, 2; i \in \{n, m, u\}$ denote the mean of $t_1$ and $t_2$ in the short list (SL) samples.
sample of size \( n \) is the IST estimator \( T_{1u} = \hat{y}_u \). Combining the two estimators as the convex linear combinations, the final estimator for sensitive population mean at current occasion is given by

\[
T_1 = \phi_1 T_u + (1 - \phi_1) T_{1m}, \tag{4.1}
\]

where \( T_u = \hat{y}_u \) and \( T_{1m} = \hat{y}_m + k(\hat{x}_n - \hat{x}_m) \); \( \phi_1 \in [0, 1] \) and \( k \) is a scalar quantity to be chosen suitably.

5. IST successive regression estimator

The regression estimator is another estimator in survey sampling theory. Hence, the estimator for the matched portion of the sample have been chosen as a regression type estimator given by

\[
T_{2m} = \hat{y}_m + \hat{b}(m_0)(\hat{x}_n - \hat{x}_m) \text{ with } \hat{b}(m_0) = \frac{s_x}{s_x^2(m_0)} \text{ and } \phi_2 \in [0, 1] \text{ is a scalar quantity to be chosen suitably.}
\]

6. IST successive general class of estimator

Many estimators such as ratio, product, exponential ratio, may be thought on similar lines for proposing an estimator based on matched sample of size \( m \). Therefore, in order to generalize the frame work, an IST general class of estimator has been proposed, so that the IST difference, IST regression and others may be viewed as members of the proposed class of estimator. Hence, the final estimator in this case is given as

\[
T_3 = \phi_3 T_u + (1 - \phi_3) T_{3m}, \tag{6.1}
\]

where \( T_{3m} = g(\hat{y}_m, \hat{x}_m, \hat{x}_n) \) is a function of \( \hat{y}_m, \hat{x}_m, \) and \( \hat{x}_n \). Following Srivastava and Jhajj (1980), Tracy et al. (1996), and Priyanka and Trisandhya (2018), the function \( g \) is assumed such that it satisfies following conditions:

(i) The point \( (\hat{y}_m, \hat{x}_m, \hat{x}_n) \) assumes the value in a closed convex subset \( \mathbb{R}^3 \) of three dimensional real space containing the point \( (\bar{Y}, \bar{X}, \bar{X}) \).

(ii) The function \( g(\hat{y}_m, \hat{x}_m, \hat{x}_n) \) is continuous and bounded in \( \mathbb{R}^3 \).

(iii) \( g(\bar{Y}, \bar{X}, \bar{X}) = \bar{Y} \) and \( g_1(\bar{Y}, \bar{X}, \bar{X}) = \partial g(\hat{y}_m, \hat{x}_m, \hat{x}_n)/\partial \hat{y}_m = 1 \), i.e., first order partial derivative of \( g \) with respect to \( \hat{y}_m \) at \( g(\bar{Y}, \bar{X}, \bar{X}) = \bar{Y} \Rightarrow g_1(K) = \partial g(\cdot)/\partial \hat{y}_m|_{K} = 1 \), where \( K = (\bar{Y}, \bar{X}, \bar{X}) \).

(iv) The first and second order partial derivatives of \( g(\hat{y}_m, \hat{x}_m, \hat{x}_n) \) exist and are continuous and bounded in \( \mathbb{R}^3 \).

7. Analysis of IST estimators on successive occasions

To elucidate the performances of proposed IST estimators, the bias, variance/mean squared error of the proposed estimators \( T_i \) \((i = 1, 2, 3)\) has been calculated as

\[
\mathbb{B}(T_i) = E \left( T_i - \bar{Y} \right) = E \left[ \phi_i (T_u - \bar{Y}) + (1 - \phi_i) (T_{1m} - \bar{Y}) \right] = \phi_i \mathbb{B}(T_u) + (1 - \phi_i) \mathbb{B}(T_{1m}), \quad i = 1, 2, 3.
\]
Since, $T_u$ is unbiased for $\bar{Y}$, so $\mathbb{E}(T_u) = 0$. Therefore, the bias of estimator $T_i$ is given as

$$\mathbb{E}(T_i) = (1 - \phi_i)\mathbb{E}(T_{im}), \quad i = 1, 2, 3.$$  \hspace{1cm} (7.1)

The variance of the estimator $T_i$ is computed as

$$\mathbb{V}(T_i) = E[(T_i - \bar{Y})^2] = E[\phi_i(T_u - \bar{Y})^2 + (1 - \phi_i)(T_{im} - \bar{Y})^2] = \phi_i \mathbb{V}(T_u) + (1 - \phi_i)^2 \mathbb{V}(T_{im}) + 2\phi_i(1 - \phi_i)\text{cov}(T_u, T_{im}), \quad i = 1, 2, 3.$$  \hspace{1cm} (7.2)

As $T_u$ and $T_{im}$ are based on two independent samples of sizes $u$ and $m$ respectively. So, $\text{cov}(T_u, T_{im}) = 0$. Therefore, the variance of estimator $T_i$ becomes

$$\mathbb{V}(T_i) = \phi_i \mathbb{V}(T_u) + (1 - \phi_i)^2 \mathbb{V}(T_{im}).$$  \hspace{1cm} (7.3)

It can be seen that, $\mathbb{V}(T_i)$ in equation (7.3) is a function of $\phi_i$. So, it has been optimized with respect to $\phi_i$ and optimum value of $\phi_i$ is obtained as:

$$\phi_{opt.} = \frac{\mathbb{V}(T_{im})}{\mathbb{V}(T_u) + \mathbb{V}(T_{im})}, \quad i = 1, 2, 3.$$  \hspace{1cm} (7.4)

Substituting the optimum value of $\phi_i$ from equation (7.4) in equation (7.3) the optimum variance of the proposed estimator $T_i$ is computed as

$$\mathbb{V}(T_{i, opt}) = \frac{\mathbb{V}(T_u) \times \mathbb{V}(T_{im})}{\mathbb{V}(T_u) + \mathbb{V}(T_{im})}, \quad i = 1, 2, 3.$$  \hspace{1cm} (7.5)

7.1. Bias and variance of $T_u$ and $T_{im}; i = 1, 2, 3$

The estimator $T_u$ is unbiased for $\bar{Y}$, hence, its variance is computed as

$$\mathbb{V}(T_u) = \left(\frac{S_i^2}{u_1}\right) + \left(\frac{S_i^2}{u_2}\right) - \left(\frac{S_i^2 + S_i^2}{N}\right).$$  \hspace{1cm} (7.6)

The estimator $T_{im}$ is also unbiased for $\bar{Y}$, so its variance is obtained as

$$\mathbb{V}(T_{im}) = \left(\frac{1}{m_1}\right) [k^2 S_i^2 - 2k \rho_{c_1} S_i S_{c_2} + S_{c_2}^2] + \left(\frac{1}{m_2}\right) [k^2 S_i^2 - 2k S_i S_{c_1} \rho_{c_1} + S_{c_1}^2]$$

$$+ \left(\frac{1}{m_3}\right) [2k \rho_{c_1} S_i S_{c_2} - k^2 S_{c_2}^2] + \left(\frac{1}{m_4}\right) [2k S_i S_{c_1} \rho_{c_1} - k^2 S_{c_1}^2] - \left[\frac{1}{N}\right] [S_i^2 + S_i^2].$$  \hspace{1cm} (7.7)

The above expression for variance of $T_{im}$ contain an unknown constant $k$, hence, it is optimized with respect to $k$ and optimum value of $k$ is obtained as

$$k_{opt.} = \frac{\left(\frac{1}{m_2} - \frac{1}{m_3}\right) S_i S_{c_2} \rho_{c_1} + \left(\frac{1}{m_3} - \frac{1}{m_1}\right) S_i S_{c_1} \rho_{c_2}}{\left(\frac{1}{m_3} - \frac{1}{m_2}\right) S_i^2 + \left(\frac{1}{m_2} - \frac{1}{m_1}\right) S_i^2}.$$
Now, as the estimator $\mathbb{T}_{2m}$ and $\mathbb{T}_{3m}$ are biased for $\hat{Y}$, hence the expression for their bias and mean squared error have been computed under the following transformations:

\[
\begin{align*}
\hat{z}_{2\text{all}} &= \hat{Z}_2(1 + e_1), \quad \hat{z}_{2\text{all}} = \hat{Z}_2(1 + e_1), \quad \hat{z}_{1\text{all}} = \hat{Z}_1(1 + e_1), \\
\hat{i}_{2\text{all}} &= T_2(1 + e_4), \quad \hat{i}_{2\text{all}} = T_2(1 + e_4), \quad \hat{i}_{1\text{all}} = T_1(1 + e_6), \quad \hat{i}_{1\text{all}} = T_1(1 + e_7), \\
x^2_2(m_{ll}) &= S^2_2(1 + e_8), \quad \hat{x}_{2\text{all}} = S^2_2(1 + e_9), \\
\hat{s}_{l\text{all}} &= \frac{\nu_{rl}}{(\nu_{2\theta})^2}, \quad \nu_{rl} = \frac{1}{n - 1} \sum(z_i - \bar{z})(z_i - \bar{z})'
\end{align*}
\]

such that, $E(e_i) = 0; |e_i| < 1$, where $i = 0, 1, 2, 3, 4, 5, 6, 7, 8$, and 9.

Under the above transformations, retaining the terms up to first order of approximation, we have for bias and mean squared error of $\mathbb{T}_{2m}$ as

\[
\mathbb{B}(\mathbb{T}_{2m}) = \frac{1}{m_{ll}} - \frac{1}{n_{ll}} \beta_{2\text{all}} \left[ S_2, n_{03} - S_2, n_{12} \right]
\]

and

\[
\mathbb{V}(\mathbb{T}_{2m}) = \frac{1}{m_{ll}} \left[ S^2_2, \beta^2_{2\text{all}} - 2S_2, \beta_{2\text{all}}, \rho_{2\text{all}} + S^2_2 \right] + \left( \frac{1}{n_{ll}} \right) \left[ S^2_2, \beta^2_{2\text{all}} - 2S_2, \beta_{2\text{all}}, \rho_{2\text{all}} + \left( \frac{1}{n_{ll}} \right) \left[ 2S_2, \beta_{2\text{all}}, \rho_{2\text{all}}, \rho_{2\text{all}} - S^2_2, \beta_{2\text{all}}, \rho_{2\text{all}} - S^2_2, \beta_{2\text{all}}, \rho_{2\text{all}} \right] \right] (7.8)
\]

Also, the bias and mean squared error of the estimator $\mathbb{T}_{3m}$ has been derived under above considered transformations as:

\[
\mathbb{T}_{3m} = g(\hat{y}_m, \hat{x}_m, \hat{x}_n).
\]

Expanding $g(\hat{y}_m, \hat{x}_m, \hat{x}_n)$ about the point $K = (\bar{Y}, \bar{X}, \bar{X})$ using Taylor series expansion, retaining terms up to first order of approximations, we have

\[
\mathbb{T}_{3m} = g(\bar{Y} + (\hat{y}_m - \bar{Y}), \bar{X} + (\hat{x}_m - \bar{X}), \bar{X} + (\hat{x}_n - \bar{X}))
\]

\[
= g(\hat{y}_m + (\hat{x}_m - \bar{X})G_2 + (\hat{x}_n - \bar{X})G_3 + (\hat{y}_m - \bar{Y})G_1 + (\hat{x}_n - \bar{X})G_{22} + (\hat{x}_n - \bar{X})G_{33} + (\hat{y}_m - \bar{Y})G_{12} + (\hat{y}_m - \bar{Y})(\hat{x}_n - \bar{X})G_{13} + (\hat{x}_n - \bar{X})(\hat{y}_m - \bar{Y})G_{23} + \cdots ) (7.10)
\]

where,

\[
G_1 = \frac{\partial g}{\partial \hat{y}_m}_{|K} = 1, \quad G_2 = \frac{\partial g}{\partial \hat{x}_m}_{|K}, \quad G_3 = \frac{\partial g}{\partial \hat{x}_n}_{|K}, \quad G_{11} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{y}_m^2}, \quad G_{22} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{x}_m^2}, \quad G_{33} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{x}_n^2}, \quad G_{12} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{y}_m \hat{x}_m}_{|K}, \quad G_{13} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{y}_m \hat{x}_n}_{|K}, \quad G_{23} = \frac{1}{2} \frac{\partial^2 g}{\partial \hat{x}_m \hat{x}_n}_{|K}.
\]

Bias and Mean squared error of the class of estimator $\mathbb{T}_{3m}$ to the first order approximations are ob-
tained as

\[
B(\mathbb{T}_{3m}) = \frac{1}{m_{ij}} \left[ S_i^2 G_{22} + \rho_{z_1 z_2} S_{z_1 z_2} G_{12} \right] + \frac{1}{m_{rl}} \left[ S_i^2 G_{22} + \rho_{r_1 r_2} S_{r_1 r_2} G_{12} \right] \\
+ \frac{1}{n_{ll}} \left[ S_i^2 G_{33} + \rho_{z_1 z_2} S_{z_1 z_2} G_{13} + S_i^2 G_{23} \right] + \frac{1}{n_{rl}} \left[ S_i^2 G_{33} + \rho_{r_1 r_2} S_{r_1 r_2} G_{13} + S_i^2 G_{23} \right] \\
- \frac{1}{N} \left[ S_i^2 (G_{22} + G_{33} + G_{23}) + S_i r_i \rho_{r_1 r_2} (G_{12} + G_{13}) + S_i^2 (G_{22} + G_{33} + G_{23}) \right] \\
+ S_i z_i \rho_{z_1 z_2} (G_{12} + G_{13}) \right],
\]

(7.11)

and

\[
V(\mathbb{T}_{3m}) = \left( \frac{1}{m_{ij}} \right) \left[ S_i^2 G_{22} + 2 \rho_{z_1 z_2} S_{z_1 z_2} G_{22} + S_i^2 \right] + \left( \frac{1}{m_{rl}} \right) \left[ S_i^2 G_{22} + 2 \rho_{r_1 r_2} S_{r_1 r_2} G_{22} + S_i^2 \right] \\
+ \left( \frac{1}{n_{ll}} \right) \left[ S_i^2 G_{33} + 2 \rho_{z_1 z_2} S_{z_1 z_2} G_{33} + 2 G_{23} S_i^2 \right] + \left( \frac{1}{n_{rl}} \right) \left[ S_i^2 G_{33} + 2 G_{23} S_i r_i \rho_{r_1 r_2} + 2 G_{23} S_i^2 \right] \\
- \frac{1}{N} \left[ S_i^2 G_{22} + 2 \rho_{z_1 z_2} S_{z_1 z_2} G_{22} + S_i^2 G_{33} + 2 S_i r_i \rho_{r_1 r_2} G_{22} + S_i^2 + S_i z_i G_{23} \right] \\
+ 2 \rho_{z_1 z_2} S_{z_1 z_2} G_{33} + 2 G_{23} S_i^2 + S_i^2 G_{23} + 2 S_i r_i \rho_{r_1 r_2} G_{33} + 2 G_{23} G_{33} S_i^2 \right].
\]

(7.12)

Clearly, we can see that equation (7.12) is a function of \( G_2 \) and \( G_3 \). So, after minimizing equation 15 by partially differentiating with respect to \( G_2 \) and \( G_3 \) respectively and equating to zero we get the optimized value of \( G_2 \) and \( G_3 \) as

\[
G_{2\text{opt.}} = \frac{\left( \frac{1}{m_{ij}} - \frac{1}{n_{ll}} \right) S_i z_i \rho_{z_1 z_2} + \left( \frac{1}{m_{rl}} - \frac{1}{n_{rl}} \right) S_i r_i \rho_{r_1 r_2}}{\left( \frac{1}{m_{ij}} - \frac{1}{n_{ll}} \right) S_i^2 + \left( \frac{1}{m_{rl}} - \frac{1}{n_{rl}} \right) S_i^2}.
\]

\[
G_{3\text{opt.}} = \frac{\left( \frac{1}{n_{ll}} - \frac{1}{N} \right) S_i z_i \rho_{z_1 z_2} + S_i^2 G_{23} + \left( \frac{1}{n_{rl}} - \frac{1}{N} \right) S_i r_i \rho_{r_1 r_2} + S_i^2 G_{23}}{\left( \frac{1}{m_{ij}} - \frac{1}{n_{ll}} \right) S_i^2 + \left( \frac{1}{m_{rl}} - \frac{1}{n_{rl}} \right) S_i^2}.
\]

7.2. Allocating LL & SL Sample using Trappmann et al. (2014) approach and Perri et al. (2018) approach

Since, in IST a sample is split in to LL sample and SL sample. Trappmann et al. (2014) considered equal number of units in both the samples irrespective of variability of the items in the two lists. Applying his approach on successive occasions we have the following allocations:

\[
m_{ll} = n_{ld} = \frac{n}{2}, \quad m_{ll} = m_{ld} = \frac{m}{2}, \quad \text{and} \quad u_{ll} = u_{ld} = \frac{u}{2}.
\]

However, Perri et al. (2018) concluded that the estimates may be affected due to high variability of items in LL sample and SL sample. Hence, they proposed optimal sample size allocation to LL and SL samples by minimizing the variance of IST estimates under a budget constraints. Hence, modifying these ideas to work for allocating LL sample and SL sample on various samples at first and second
### Table 2: Minimum variance

| $i$ | Est. $V_i(T_i)$ | Minimum Variance under Trappmann et al. (2014) approach |
|-----|-----------------|----------------------------------------------------------|
| 1   | $T_1$           | $V_{i}(T_{1})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |
| 2   | $T_2$           | $V_{i}(T_{2})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |
| 3   | $T_3$           | $V_{i}(T_{3})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |

### Minimum Variance under Perri et al. (2018) approach

| $i$ | Est. $V_i(T_i)$ | Minimum Variance under Perri et al. (2018) approach |
|-----|-----------------|----------------------------------------------------------|
| 1   | $T_1$           | $V_{i}(T_{1})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |
| 2   | $T_2$           | $V_{i}(T_{2})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |
| 3   | $T_3$           | $V_{i}(T_{3})_{\text{min}} = \frac{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}{\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4}$ with $\hat{\mu}_i = \min \left( \frac{-\hat{B}_2 + \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1}, \frac{-\hat{B}_2 - \sqrt{\hat{B}_2^2 + 4\hat{B}_1\hat{B}_3}}{2\hat{B}_1} \right) \in [0, 1]$ |

### occasion assuming same budget allocation for each LL and SL samples have:

$$
\begin{align*}
n_{ij} &= \frac{nS_{zi}}{S_{zi} + S_{ti}} = n\beta_i \text{ (say)}, \\
n_{ii} &= \frac{nS_{ti}}{S_{zi} + S_{ti}} = n\beta_i \text{ (say)}, \\
u_{ij} &= \frac{uS_{zi}}{S_{zi} + S_{ti}} = u\beta_i \text{ (say)}, \\
u_{ii} &= \frac{uS_{ti}}{S_{zi} + S_{ti}} = u\beta_i \text{ (say)}, \\
m_{ij} &= \frac{mS_{zi}}{S_{zi} + S_{ti}} = m\beta_i \text{ (say)}, \\
m_{ii} &= \frac{mS_{ti}}{S_{zi} + S_{ti}} = m\beta_i \text{ (say)}, \\
u_{ij} &= \frac{nuS_{zi}}{S_{zi} + S_{ti}} = nu\beta_i \text{ (say)}, \\
u_{ii} &= \frac{nuS_{ti}}{S_{zi} + S_{ti}} = nu\beta_i \text{ (say)}, \\
m_{ij} &= \frac{nuS_{zi}}{S_{zi} + S_{ti}} = nu\beta_i \text{ (say)}, \\
m_{ii} &= \frac{nuS_{ti}}{S_{zi} + S_{ti}} = nu\beta_i \text{ (say)},
\end{align*}
$$

Using the assumptions of both the approaches the minimum variance of proposed estimators $T_i$ has been obtained and are presented in Table 2. Where,

- $J_1 = K_0K_3f - K_0K_4j^2$, $J_2 = K_0K_3f - K_0K_2j - K_0K_3j + K_1K_4j$,
- $J_3 = K_0K_3j - f(K_0K_3j) - K_0K_3j - K_0K_3j - K_0K_3j - K_0K_3j - K_0K_3j$, $J_4 = K_0K_3j - K_0K_3j - K_0K_3j - K_0K_3j$,
- $I_1 = J_1J_2j + J_2J_3$, $I_2 = J_2K_1j - J_3K_3j$, $I_3 = J_2K_1j + J_3K_3j$,
- $K_0 = S_{zi}^2 + S_{ti}^2$, $\forall j = 1, 2, \ldots, 6$; $K_1 = 2S_{zi}^2 + 2S_{ti}^2$, $\forall j = 1, 3, 5$; $K_2 = S_{zi}^2 + S_{ti}^2$, $\forall j = 2, 4, 6$;
- $K_{12} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2 + 2k_2S_{zi}^2 + 2k_2S_{ti}^2$, $K_{13} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2 + 2k_2S_{zi}^2 + 2k_2S_{ti}^2$, $K_{14} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2$,
- $K_{22} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2$, $K_{23} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2$, $K_{24} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2$, $K_{32} = 2k_2S_{zi}^2 + 2k_2S_{ti}^2$. 

---

Kumari Priyanka, Pidugu Trisandhya
\( K_{33} = 2 \left( 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 + 2S_t^2 S_{\tau}^2 \rho_{\tau t} \right) - \left( S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 \right), \quad K_{34} = S_{\zeta t}^2 + S_{\tau t}^2. \)

\( K_{42} = \left( S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 + S_{\tau}^2 \right) \frac{1}{\beta_1} + \left( S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 + S_{\tau}^2 \right) \frac{1}{\beta_2}, \)

\( K_{43} = \left( 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 \right) \frac{1}{\beta_1} + \left( 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 S_{\zeta t}^2 - S_{\tau}^2 S_{\zeta t}^2 \right) \frac{1}{\beta_2}, \)

\( K_{44} = S_{\zeta t}^2 + S_{\tau t}^2, \quad K_{32} = 2 \left( S_{\tau}^2 S_{\zeta t}^2 + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 \right), \)

\( K_{53} = 2 \left( S_{\tau}^2 G_{3p} + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 \right), \)

\( K_{54} = \frac{S_{\tau}^2 G_{3p} + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2 + 2S_t^2 S_{\tau}^2 \rho_{\tau t} + S_{\tau}^2}{\beta_1} \frac{1}{\beta_2}, \quad G_{2t} = -S_{\zeta t} S_{\tau}^2 \rho_{\tau t} \frac{S_{\tau}^2 + S_{\tau t}^2}{S_{\tau}^2 + S_{\tau t}^2}, \quad G_{3t} = -S_{\zeta t} S_{\tau}^2 \rho_{\tau t} \frac{S_{\tau}^2 + S_{\tau t}^2}{S_{\tau}^2 + S_{\tau t}^2} \beta_1, \quad G_{3p} = -S_{\zeta t} S_{\tau}^2 \rho_{\tau t} \frac{S_{\tau}^2 + S_{\tau t}^2}{\beta_2} \beta_1, \quad \) and \( f = \frac{n}{N}. \)

8. Efficiency comparison

In order to compare various proposed IST estimators in successive sampling, the percent relative efficiencies have been computed for data considered in section 9 under Trappmann et al. (2014) as well as Perri et al. (2018) allocation designs as:

\[ E_1 = \frac{\mathcal{V}(T_1)_{\min}}{\mathcal{V}(T)_{\min}} \times 100, \quad E_2 = \frac{\mathcal{V}(T_2)_{\min}}{\mathcal{V}(T)_{\min}} \times 100, \quad \text{and} \quad E_3 = \frac{\mathcal{V}(T_3)_{\min}}{\mathcal{V}(T)_{\min}} \times 100. \]

9. Numerical demonstration

Population Source: [Free access to data by Statistical Abstracts of United States]

To evaluate the performance of proposed IST successive sampling estimators, numerical illustrations has been supplemented using natural population. The population consists of \( N = 51 \) states. Let the aim be to estimate rate of abortion in year 2004. Therefore, for IST successive sampling frame work
Table 3: Empirical results

| $\hat{\mu}_1$ | $\hat{\mu}_2$ | $\hat{\mu}_3$ | $\hat{\mu}_4$ | $\hat{\mu}_5$ | $E_1$ | $E_2$ | $E_3$ |
|--------------|--------------|--------------|--------------|--------------|------|------|------|
| 0.8424       | 0.8359       | 0.7617       | 0.3626       | *            |      |      | *    |
|              |              |              |              | *            | 114.9334 | 115.8558 | -    |

"*" = indicates that the optimum value of fraction of sample to be drawn afresh do not exist; '- ' denote corresponding percent relative efficiency cannot be computed.

we consider:

\[
\begin{align*}
    y &= \text{Rate of abortions in the year 2004} \\
    x &= \text{Rate of abortions in the year 2000} \\
    t_1 &= \text{Number of residents in the year 2000} \\
    t_2 &= \text{Number of residents in the year 2004}
\end{align*}
\]

Clearly the rate of abortion is sensitive; however, rate of residents is non-sensitive. Hence, the data is suitable to be applied for IST frame work. Since same study variable “rate of abortion”has been observed for two different years 2000 and 2004, therefore, the considered data is suitable for IST successive sampling frame work. The numerical calculations have been performed on the data with results represented in Table 3.

The value of $E_i$ ($i = 1, 2$) are observed to be more than 100, this indicates that optimum allocation design by Perri et al. (2018) is preferable over Trappmann et al. (2014) design. Therefore, the further numerical analysis has been done using Perri et al. (2018) allocation design.

10. Simulation study

An extensive simulation study has been done using Monte Carlo simulation for the data mentioned in Section 9. The 5,000 different Monte Carlo replications have been observed. The process is also repeated for a different combination of constants termed as sets. The variance/mean squared error of the proposed estimators $T_1$, $T_2$, and $T_3$ has been computed under Perri et al. (2018); in addition, the allocation design and are denoted by $V_p(T_1)$, $V_p(T_2)$, and $V_p(T_3)$ respectively. The percent of relative efficiencies for the IST successive difference and regression estimators with respect to IST successive general class of estimator have been computed as:

\[
E_{si} = \frac{V_p(T_i)}{V_p(T_s)} \times 100
\]

The procedure have been repeated for three different combinations of constants, termed as different sets given as:

\[
\begin{align*}
    \text{I} : & \quad n = 24, \quad u = 04, \quad m = 20 \\
    \text{II} : & \quad n = 24, \quad u = 08, \quad m = 16 \\
    \text{III} : & \quad n = 24, \quad u = 10, \quad m = 14.
\end{align*}
\]

Figure 1 and Figure 2 summarizes the outcomes of the simulation results.

11. Direct method

Estimators that use IST are less efficient than estimators obtained using direct questioning. Hence, in order to identify the amount of loss we compare the proposed class of estimator $T_3$ with respect to
ISTs for quantitative sensitive estimation on successive occasions

Figure 1: Simulated percent relative efficiency of the proposed IST general class of estimator with respect to proposed IST difference estimator for three different sets.

Figure 2: Simulated percent relative efficiency of the proposed IST general class of estimator with respect to proposed IST regression estimator for three different sets.

corresponding direct method. The estimator under direct questioning method is proposed as

$$T_D = \chi T_{uD} + (1 - \chi) T_{3mD}, \quad \chi \in [0, 1],$$

(11.1)

where

$$T_{uD} = \bar{y}_u,$$

(11.2)

$$T_{3mD} = d(\bar{y}_m, \bar{x}_m, \bar{x}_s).$$

(11.3)
where \( T_{3aD} \) follow similar regularity conditions as stated in Section 6. The, minimum mean squared error of the class of estimator \( T_{D} \) is obtained as

\[
\mathbb{V}(T_{D})_{\text{min}} = \frac{-\hat{\mu}_D J_{d1} + \hat{\mu}_D J_{d2} + J_{d3}}{\hat{\mu}_D J_{d4} - \hat{\mu}_D J_{d5} + K_{d1}}\tag{11.4}
\]

with

\[
\hat{\mu}_D = \min \left\{ \frac{I_{d2} + \sqrt{I_{d2}^2 - I_{d1} I_{d5}}}{I_{d1}}, \frac{I_{d2} - \sqrt{I_{d2}^2 - I_{d1} I_{d5}}}{I_{d1}} \right\} \quad \text{such that } \hat{\mu}_D \in [0, 1].\tag{11.5}
\]

where,

\[
J_{d1} = K_{d1} K_{d3} \hat{f}^2, \quad J_{d2} = f(K_{d1} K_{d3} - K_{d1} K_{d2} + f K_{d1} K_{d3}), \quad J_{d3} = K_{d1} K_{d2} - f K_{d1} K_{d3}, \quad J_{d4} = f(K_{d1} + K_{d3}), \quad J_{d5} = K_{d1} - K_{d2} + f(K_{d1} + K_{d3}), \quad I_{d1} = J_{d1} J_{d5} - J_{d2} J_{d4}, \quad I_{d2} = J_{d3} J_{d4} + J_{d1} K_{d1}, \quad I_{d3} = J_{d2} K_{d1} + J_{d3} J_{d5}, \quad K_{d1} = S_y^2, \quad K_{d2} = S_y^2 + S_x S_y D_2^2 + 2 S_x^2 D_2 \rho_{yx}, \quad K_{d3} = S_y^2 + S_x^2 D_2^2 + 2 S_x S_y \rho_{yx} D_2, \quad D_2 = \frac{-S_y S_x \rho_{yx}}{S_x^2},
\]

Further the simulated ratio of the mean squared error of \( T_{D} \) and \( T_{3} \) have been computed by considering 5,000 different samples using Monte Carlo simulation study for different sets and results are presented in Figure 3 and Figure 4 respectively

\[
\text{Ratio} = \frac{\mathbb{V}(T_{D})}{\mathbb{V}(T_{3})}\tag{11.6}
\]

**Remark 1.** The IST is used with an expectation to receive a true response as in IST directly; however, the response to sensitive questions are not being asked. Standard bias reduction methods like Jackknife, the Bootstrap and methods that use approximations of the bias function through asymptotic expansions of the bias that follow Kosmidis (2014) may be used if some false responses are still received that show bias in the estimator after explaining the IST efficiently to respondents.

### 12. Interpretations of results

The following interpretations can be drawn from empirical and simulation results:

1. It has been observed that IST is feasible in successive sampling to handle sensitive issues on successive occasions.

2. From Table 3, it is clear that \( E_1 \) and \( E_2 \) both are coming out to be greater than 100, this implies that optimum allocation design by Perri et al. (2018) is more efficient than allocation by Trappmann et al. (2014) design in two occasion successive sampling. It is to be noted that for the considered data the optimum fraction to be drawn afresh do not exist for IST successive general class of estimator, so corresponding efficiency cannot be computed. Hence, in order to check the validity of IST successive general class of estimator simulation has been carried out with several choices of parameters.
3. Simulation results in Figure 1 and Figure 2, justify that $E_{s1}$ and $E_{s2}$ are greater than 100 for all three considered sets. This indicates that IST successive general class of estimators is more efficient than IST successive regression and IST, a successive difference estimators. However, $E_{s1} > E_{s2}$ indicate that the IST successive difference estimator is better than the IST successive regression estimator. Also, the simulated percent relative efficiency increases as $\phi$ increases and is in accordance with the theory of successive sampling.

4. From Figure 3 and Figure 4, it is observed that as $\phi$ increases the simulated ratio of the mean squared error of the direct method and IST successive general class of estimator increases. The values indicate a loss in precision of the IST general class of estimator over the direct method. The issues under consideration are sensitive; therefore, a direct method is unsuitable because the privacy of respondents need to be considered; therefore, despite of loss in precision IST need to be preferred over direct method for sensitive issues on successive occasions.

13. Conclusion

The IST on successive occasions enables an estimation of the population mean of stigmatized quantitative variable using innocuous information, that reduce the social desirability response bias and provides privacy. Out of the three proposed IST successive estimators, the IST successive general class of estimator under both Trappmann et al. (2014) allocation designs as well as Perri et al. (2018) allocation designs have been proven to be more efficient than the other two. The optimum allocation of LL and SL samples by Perri et al. (2018) have been found to be more fruitful in successive sampling than allocation due to Trappmann et al. (2014). While comparing with direct method, certain
amount of loss in precision is observed but that is realistic as the survey issues are sensitive so there may be chances of complete refusal or partial refusal if we apply direct method. However, using IST on successive occasion atleast estimation of sensitive issues are possible. Therefore it can be concluded that the proposed IST successive estimators provide comfort and satisfaction to the respondents in terms of privacy protection as well as a methodological advancement in literature related to successive sampling dealing with sensitive issues. Hence, the IST successive class of estimators may be recommended for practical use by survey practitioners.

Acknowledgements

The authors are thankful to the reviewers and editors for their valuable suggestions that improved an earlier version of the paper. The authors are also thankful to SERB, New Delhi, India for providing the financial assistance to carry out the present work. The authors also sincerely acknowledge free access to data from the Statistical Abstracts of United States.

References

Arnab R and Singh S (2013). Estimation of mean of sensitive characteristics for successive sampling, Communications in Statistics - Theory and Methods, 42, 2499–2524.
Chaudhuri A and Christofides TC (2013). Indirect Questioning in Sample Surveys, Springer-Verlag, Berlin, Heidelberg, De.
Hussian Z, Shabbir N, and Shabbir J (2015). An alternative item sum technique for improved estimators of population mean in sensitive surveys, Hacettepe University Bulletin of Natural Sciences
ISTs for quantitative sensitive estimation on successive occasions

and Engineering Series B: Mathematics and Statistics, 46, 1–30.
Jessen RJ (1942). Statistical investigation of a sample survey for obtaining farm facts, Iowa Agriculture and Home Economics Experiment Station. Research Bulletin, 26, 1–104.

Kosmidis I (2014). Bias in parametric estimation: reduction and useful side-effects, WIREs Computational Statistics, 6, 185–196.

Miller JD (1984). A new survey technique for studying deviant behavior (Ph.D. thesis). The George Washington University, Washington DC.

Naeem N and Shabbir J (2016). Use of scrambled responses on two occasions successive sampling under non-response, Hacettepe University Bulletin of Natural Sciences and Engineering Series B: Mathematics and Statistics, 46. Available from: http://www.hjms.hacettepe.edu.tr/uploads/32c18b80-275f-4b5a-a28d-6a5890eeac3.pdf

Perri PF, Rueda Garcia M, and Cobo Rodriguez B (2018). Multiple sensitive estimation and optimal sample size allocation in the item sum technique, Biometrical Journal, 60, 155–173.

Priyanka K and Trisandhya P (2018). A composite class of estimators using scrambled response mechanism for sensitive population mean in successive sampling, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2017.1422762.

Priyanka K, Trisandhya P, and Mittal R (2018). Dealing sensitive characters on successive occasions through a general class of estimators using scrambled response techniques, Metron, 76, 203–230.

Rueda Garcia M, Perri PF, and Cobo Rodriguez B (2017). Advances in estimation by the item sum technique using Auxiliary information in complex surveys, Advances in Statistical Analysis, 102, 455–478.

Singh GN, Suman S, Khetan M, and Paul C (2017). Some estimation procedures of sensitive character using scrambled response techniques in successive sampling, Communications in Statistics - Theory and Methods, DOI:10.1080/03610926.2017.1327073.

Srivastava SK and Jhajj HS (1980). A class of estimators using auxiliary information for estimating finite population variance, Sankhya, C42, 87–96.

Tracy DS, Singh HP, and Singh R (1996). An alternative to the ratio-cum-product estimator in sample surveys, Journal of Statistical Planning and Inference, 53, 375–397.

Trappmann M, Krumpal I, Kirchner A, and Jann B (2014). Item sum: a new technique for asking quantitative sensitive questions, Journal of Survey Statistics and Methodology, 2, 58–77.

Tian GL and Tang ML (2014). Incomplete Categorical Data Design: Non-Randomized Response Techniques for Sensitive Questions in Surveys, Chapman & Hall/CRC, FL.

Warner SL (1965). Randomized response: a survey technique for eliminating evasive answer bias, Journal of the American Statistical Association, 60, 63–69.

Yu B, Jin Z, Tian J, and Gao G (2015). Estimation of sensitive proportion by randomized response data in successive sampling, Computational and Mathematical Methods in Medicine, 2015, 1–6.

Received October 4, 2018; Revised January 24, 2019; Accepted January 31, 2019