Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Trends in Bioinformatics
ISSN 1994-7941
DOI: 10.3923/tb.2018.44.55

Research Article
Missing Data in Clinical Trials: Stratified Singh and Grewal's Randomized Response Model Using Geometric Distribution

[1]Tanveer Ahmad Tarray and [2]Housila P. Singh

[1]Department of Mathematical Sciences, Islamic University of Science and Technology, Awantipora, 192122, Kashmir, India
[2]School of Studies in Statistics, Vikram University, 456010, Ujjain, India

Abstract
Background and Objective: Singh and Grewal's model estimates the population proportion on stratified sampling schemes. The crux of this paper was to develop a new stratified Singh and Grewal's model using Geometric Distribution. **Materials and Methods:** The equilibrium point of the model was investigated and a new stratified sampling and stratified double sampling randomized response model based on Singh and Grewal's model using geometric distribution was proposed. **Results:** This study showed that the proposed method is more efficient than the one recently envisaged by Singh and Grewal's model. Numerical illustrations and graphical representations are also given in support of the present study. **Conclusion:** A new dexterous stratified randomized response model has been proposed and shown theoretically as well as numerically that the proposed model is more efficient than Singh and Grewal's randomized response model.

Key words: Randomized response sampling, estimation of proportion, respondents protection, sensitive characteristic

Citation: Tanveer Ahmad Tarray and Housila P. Singh, 2018. Missing data in clinical trials: Stratified singh and Grewal's randomized response model using geometric distribution. Trends Bioinform., 11: 44-55.

Corresponding Author: Tanveer Ahmad Tarray, Department of Mathematical Sciences, Islamic University of Science and Technology, Awantipora, 192122 Kashmir, India Tel: +9596560866

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Warner[1] introduced a randomized response model to estimate a population proportion for sensitive attribute including homosexuality, drug addiction or abortion. Kuk[2] remarked that an undesirable feature of randomized response technique is that under the existing scheme, some of the respondents will have to answer "Yes" or "No" to the sensitive question and there respondents are likely to become doubtful and uncooperative. Researchers agree that it is difficult to convince the respondents that they are protected through randomization. Warner[1] mentioned that interviewees often act as if they believe the random device simply determines whether they are required to reveal or not to reveal a secret. Further a minute problem arises when a respondent without the sensitive attribute becomes too eager to declare his innocence. To overcome such difficulties Kuk[2] introduced an ingenious randomized response model in which if a respondent belong to a sensitive group A, then he/she is instructed to use a deck of cards having θ_1 proportion of cards with the statement, "I belong to group A" and if the respondent belong to non-sensitive group A^c then the respondent is requested to use a different deck of cards having θ_2 proportion of cards with the statement, "I do not belong to group A". Let π be the true proportion of persons belonging to the sensitive group A. Thus, the probability of "Yes" answer in the Kuk[2] model is given by:

$$\theta = \theta_1\pi+(1-\pi)\,\theta_2 \tag{1}$$

An unbiased estimator of the population proportion π is given by:

$$\hat{\pi}_{kuk} = \frac{\{(n_1/n)-\theta_2\}}{(\theta_1-\theta_2)},\ \theta_1 \neq \theta_2 \tag{2}$$

with variance:

$$V(\hat{\pi}_{kuk}) = \frac{\theta(1-\theta)}{n(\theta_1-\theta_2)^2} \tag{3}$$

Recently Singh and Grewal[3] have suggested improvement in the Kuk[2] model geometric distribution as a randomization device. They have claimed that their method is more protective and efficient than the Kuk[2] model while doing surveys in practice.

Land *et al.*[4] proposed the estimators for the mean number of persons who have the rare sensitive attribute using unrelated question randomized response model by utilizing a Poisson distribution. Lee *et al.*[5] applied stratified sampling schemes (stratified sampling and stratified double sampling) to Land *et al.*[4] randomized response model using the Poisson distribution[6-17].

MODEL DESCRIPTION AND ANALYSIS

Singh and Grewal[3] suggested an unbiased estimator of population proportion π possessing sensitive attribute using geometric distribution in Kuk[2] model. This procedure is unique and very useful method among different randomized response procedures so far. This study extended Singh and Grewal[3] study to stratified sampling. It assumed of that the sizes of stratified population are known.

Let a population with size N be divided into disjoint L groups (strata) with size N_h (h = 1, 2,..., L). In the proposed procedure, an individual respondent in the sample of stratum h is provided with two decks of cards in the same way as in Kuk[2] model. In the first deck of cards, let θ^*_{i1} be the proportion of cards with the statement "I belong to a sensitive group A" and $(1-\theta^*_{i1})$ be the proportion of cards with the statement, "I do not belong to a sensitive group A". In the second deck of cards, let θ^*_{i2} be the proportion of cards with the statement, "I do not belong to group A" and $(1-\theta^*_{i1})$ be the proportion of cards with the statement, "I belong to a sensitive group A". Up to here, it is same as the Kuk[2] randomized response model. Now in the suggested model, if a respondent belongs to a sensitive group A, he/she is instructed to draw cards, one-by-one with replacement, from the first deck of cards until he/she gets the first card bearing the statement of his/her own status and requested to report the total number of cards, say X_h drawn by him/her to obtain the first card of his/ her own status. If a respondent belongs to group A^c, he/she is instructed to draw cards, one-by-one using with replacement, from the second deck of cards until he/she gets the first card bearing the statement of his/her own status and requested to report the total number of cards, say Y_h, drawn by him/her to obtain the first card of his/her own status. Thus, $X_h \sim G(\theta^*_{h1})$ and $Y_h \sim G(\theta^*_{h2})$ that is X_h and Y_h follows geometric distributions with parameters θ^*_{h1} and θ^*_{h2}, respectively, because cards are drawn based on with replacement sampling[3].

By this randomization device, the probability of a "Yes" in stratum h is given by:

$$\theta_h = \theta^*_{h2}\pi_{Sh} + \theta^*_{h1}(1-\pi_{Sh}) \tag{4}$$

where, π_{Sh} is the population proportion possessing sensitive attribute A in the h^{th} stratum.

Let Z_{hj} be the number of cards reported by the j^{th} respondent in the h^{th} respondent, then an estimator of the population proportion π_{Sh} is given by:

$$\hat{\pi}_{Sh} = \frac{\left\{\dfrac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})},\ \theta^*_{h1} \neq \theta^*_{h2} \tag{5}$$

Since the selections in different strata are made independently, the estimators for individual strata can be added together to obtain an estimator for the entire population. Thus an estimator of the population proportion $\pi_S = \sum_{h=1}^{L} w_h\pi_{Sh}$ with sensitive attribute, is defined by:

$$\hat{\pi}_S = \sum_{h=1}^{L} w_h\hat{\pi}_{Sh} = \sum_{h=1}^{L} w_h \frac{\left\{\dfrac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})} \tag{6}$$

where, $w_h = N_h/N$, so that, $\sum_{h=1}^{L} w_h = 1$, N is the number of units in the whole population and N_h is the total number of units in stratum h (h = 1, 2,..., L).

Theorem 2.1: The proposed estimator $\hat{\pi}_S$ is an unbiased estimate for the population proportion π_S.

Proof: Given $X_h \sim G(\theta^*_{h1})$ and $Y_h \sim G(\theta^*_{h2})$, therefore, we have:

$$E(Z_{hj}) = \frac{\pi_{Sh}}{\theta^*_{h1}} + \frac{(1-\pi_{Sh})}{\theta^*_{h2}} \tag{7}$$

By taking the expected value on both sides of Eq. 6 and using Eq. 7 we have:

$$E(\hat{\pi}_S) = \sum_{h=1}^{L} w_h E(\hat{\pi}_{Sh})$$

$$= \sum_{h=1}^{L} w_h\ E\left[\frac{\left\{\dfrac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}\right]$$

$$= \sum_{h=1}^{L} w_h \left[\frac{\left\{\dfrac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} E(Z_{hj}) - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}\right]$$

$$= \sum_{h=1}^{L} w_h \left[\frac{\left\{\dfrac{\left(\theta^*_{h1}\theta^*_{h2}/n_h\right)n_h\left\{\theta^*_{h2}\pi_{Sh} + (1-\pi_{Sh})\theta^*_{h1}\right\}}{(\theta^*_{h1}\theta^*_{h2})} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}\right]$$

$$= \sum_{h=1}^{L} w_h \frac{\left\{\theta^*_{h2}\pi_{Sh} + (1-\pi_{Sh})\theta^*_{h1} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}$$

$$= \sum_{h=1}^{L} w_h \frac{\pi_{Sh}\left\{\theta^*_{h2}-\theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}$$

$$= \sum_{h=1}^{L} w_h\pi_{Sh} = \pi_S$$

which proves the theorem.

Theorem 2.2: The variance of the estimator $\hat{\pi}_{Sh}$ is:

$$V(\hat{\pi}_S) = \sum_{h=1}^{L}\frac{w_h^2}{n_h}\times \left[\pi_{Sh}(1-\pi_{Sh}) + \frac{\left\{\theta^{*\,2}_{h2}(1-\theta^*_{h1})\pi_{Sh} + \theta^{*\,2}_{h1}(1-\theta^*_{h2})(1-\pi_{Sh})\right\}}{(\theta^*_{h2}-\theta^*_{h1})^2}\right] \tag{8}$$

Proof: Since the responses are independent and each unbiased estimator $\hat{\pi}_{Sh}$ has its own variance, the variance of $\hat{\pi}_S$ is:

$$V(\hat{\pi}_S) = V\left(\sum_{h=1}^{k} w_h\hat{\pi}_{Sh}\right) = \sum_{h=1}^{k} w_h^2 V(\hat{\pi}_{Sh})$$

$$= \sum_{h=1}^{k} w_h^2 V\left[\frac{\left\{\dfrac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right\}}{(\theta^*_{h2}-\theta^*_{h1})}\right]$$

$$= \sum_{h=1}^{L} w_h^2 \left[\frac{\left\{\left(\theta^*_{h1}\theta^*_{h2}\right)^2 \sum_{j=1}^{n_h} V(Z_{hj})\right\}}{n_h^2(\theta^*_{h2}-\theta^*_{h1})^2}\right] \tag{9}$$

$$= \sum_{h=1}^{L} w_h^2 \left[\frac{\left(\theta^*_{h1}\theta^*_{h2}\right)^2 \sum_{j=1}^{n_h}\left\{E(Z^2_{hj}) - \left((E(Z_{hj})\right)^2\right\}}{n_h^2(\theta^*_{h2}-\theta^*_{h1})^2}\right]$$

Since $X_h \sim G(\theta^*_{h1})$ and $Y_h \sim G(\theta^*_{h2})$, therefore, we have:

$$E(Z^2_{hj}) = \pi_{Sh}\left(\frac{2-\theta^*_{h1}}{\theta^{*\,2}_{h1}}\right) + (1-\pi_{Sh})\left(\frac{2-\theta^*_{h2}}{\theta^{*\,2}_{h12}}\right) \tag{10}$$

Using Eq. 7, 10 in Eq. 9 we get Eq. 8. Thus the theorem is proved.

Now we consider the problem of allocation of the sample to the different strata using (i) Proportional allocation, (ii) Optimum allocation and (iii) Neyman allocation methods. The variances of the proposed estimator $\hat{\pi}_S$ are also obtained under three allocation methods.

Proportional allocation: Proportional allocation is a technique to define sample size. Since the sample size in stratum h is defined as $n_h = n\, w_h$, where $w_h = N_h/N$, the variance of the estimator $\hat{\pi}_S$ is given in the following theorem.

Theorem 2.3: Under the Proportional allocation $n_h = n\, w_h$ (h = 1, 2,..., L), the variance $\hat{\pi}_S$ is given by:

$$V(\hat{\pi}_S)_P = \frac{1}{n}\sum_{h=1}^{L} w_h V_{h0} \tag{11}$$

Where:

$$V_{h0} = \left[\frac{\pi_{Sh}(1-\pi_{Sh}) + \left\{\theta^*_{h2}(1-\theta^*_{h1})\pi_{Sh} + \theta^*_{h1}(1-\theta^*_{h2})(1-\pi_{Sh})\right\}}{(\theta^*_{h2}-\theta^*_{h1})^2}\right] \tag{12}$$

Proof: Proof is simple so omitted.

Optimum allocation: In optimum allocation, the sample sizes are defined to minimize variance with a given cost. In stratified sampling, the cost function is defined as:

$$C = c_0 + \sum_{h=1}^{L} n_h c_h$$

where, c_0 is a fixed cost and c_h is a cost per sample unit in stratum h. For fixed cost, by the Cauchy-schwarz inequality, the sample size n_h to minimize $V(\hat{\pi}_S)$ is given by:

$$n_h = n \times \frac{W_h\sqrt{V_{h0}/c_h}}{\sum_{h=1}^{L} W_h\sqrt{V_h/c_h}} \tag{13}$$

Theorem 2.4: The variance of the proposed estimator $\hat{\pi}_S$ under the optimum allocation (Eq. 13) is given by:

$$V(\hat{\pi}_S)_o = \frac{1}{n}\left[\sum_{h=1}^{L} W_h\sqrt{V_{h0}c_h}\right] \times \left[\sum_{h=1}^{L} W_h\sqrt{\frac{V_{h0}}{c_h}}\right] \tag{14}$$

Proof: Proof is simple so omitted.

Neyman allocation: Information on $\hat{\pi}_{Sh}$ is usually unavailable. But if prior information on $\hat{\pi}_{Sh}$ is available from the past experience then we may derive the following Neyman allocation formula.

Theorem 2.5: The Neyman allocation n to $n_1, n_2,..., n_{L-1}$ and n_L to derive the minimum variance of $\hat{\pi}_S$ subject $n = \sum_{h=1}^{L} n_h$ is approximately given by:

$$\frac{n_h}{n} = \frac{w_h\sqrt{V_{ho}}}{\sum_{h=1}^{L} w_h\sqrt{V_{ho}}} \tag{15}$$

Proof: The minimum variance of $\hat{\pi}_S$ is given by[18]:

$$V(\hat{\pi}_S)_N = \frac{1}{n}\left[\sum_{h=1}^{L} w_h\sqrt{V_{h0}}\right]^2 \tag{16}$$

By substituting (n_h-1) for n_h in Eq. 7, the unbiased estimator of the minimal variance of $\hat{\pi}_S$ can be obtained.

ESTIMATION OF A POPULATION PROPORTION OF A SENSITIVE ATTRIBUTE IN STRATIFIED DOUBLE SAMPLING

When a population is stratified and the sizes of strata are unknown, the stratified double sampling method and the allocation problems are used, which are studied in this section. In the stratified double sampling method, it could have some information of the strata from a large sample randomly selected and re-sampled based on the estimated sizes for each stratum[5].

Since the sizes of the strata are not given, the respondents are asked a question to be stratified in the first phase. From a population with size N and L strata, n' samples are randomly taken and asked 'Do you belong to stratum h?'.

By this question, the sample is stratified into L strata. Let the number of respondents in stratum h be n'_h, then w_h and w'_h defined as:

$$w_h = N_h/N \text{ and } w'_h = n'_h/n'$$

where, w_h is a population proportion and w'_h is a sample proportion in stratum h. Also w'_h is an unbiased estimator for w_h (i.e., $E(w'_h) = w_h$).

In the second phase, the n_h respondents are randomly sampled with replacement from the n'_h samples in the previous phase and the randomized device is utilized. The estimator $\hat{\pi}'_S$ for $\hat{\pi}_S$ is given by:

$$\begin{aligned}\hat{\pi}'_S &= \sum_{h=1}^{L} w'_h \hat{\pi}_{Sh} \\ &= \sum_{h=1}^{L} w'_h \frac{\left\{\frac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right\}}{(\theta^*_{h2} - \theta^*_{h1})}\end{aligned} \quad (17)$$

Theorem 3.1: The estimator $\hat{\pi}'_S$ is the unbiased estimator of $\hat{\pi}_S$.

Proof: Given $X_h \sim \text{iid } G(\theta^*_{h1})$ and $Y_h \sim \text{iid } G(\theta^*_{h2})$, therefore, we have:

$$\begin{aligned}E(\hat{\pi}'_S) &= E_1\left[E_2\left\{\sum_{h=1}^{L}\frac{w'_h}{(\theta^*_{h2}-\theta^*_{h1})}\left(\frac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} Z_{hj} - \theta^*_{h1}\right) \mid w'_h\right\}\right] \\ &= E_1\left[\sum_{h=1}^{L}\frac{w'_h}{(\theta^*_{h2}-\theta^*_{h1})}\left(\frac{\theta^*_{h1}\theta^*_{h2}}{n_h}\sum_{j=1}^{n_h} E_2(Z_{hj}) - \theta^*_{h1}\right)\right] \\ &= E_1\left[\sum_{h=1}^{L}\frac{w'_h}{(\theta^*_{h2}-\theta^*_{h1})}\left(\left(\theta^*_{h1}\theta^*_{h2}/n_h\right) n_h \frac{\left\{\theta^*_{h2}\pi_{Sh} + (1-\pi_{Sh})\theta^*_{h1}\right\}}{\theta^*_{h1}\theta^*_{h2}} - \theta^*_{h1}\right)\right] \\ &= E_1\left[\sum_{h=1}^{L}\frac{w'_h}{(\theta^*_{h2}-\theta^*_{h1})} \times \left(\theta^*_{h2}\pi_{Sh} + \theta^*_{h1}(1-\pi_{Sh}) - \theta^*_{h1}\right)\right] \\ &= E_1\left[\sum_{h=1}^{L}\frac{w'_h}{(\theta^*_{h2}-\theta^*_{h1})}\left(\theta^*_{h2}-\theta^*_{h1}\right)\pi_{Sh}\right] \\ &= E_1\left[\sum_{h=1}^{L} w'_h \pi_{Sh}\right] = \sum_{h=1}^{L} E_1(w'_h)\pi_{Sh} \\ &= \sum_{h=1}^{L} w_h \pi_{Sh} = \pi_S\end{aligned}$$

which proves the theorem.

Theorem 3.2: The variance of $\hat{\pi}'_S$ is given by:

$$\begin{aligned}V(\hat{\pi}'_S) &= \frac{1}{n'}\left[\sum_{h=1}^{L} w_h V_{h0} + \sum_{h=1}^{L} w_h(\pi_{Sh} - \pi_S)^2\right] \\ &+ \left[\sum_{h=1}^{L}\frac{w_h}{n'}\left(\frac{1}{\nu_h} - 1\right)V_{h0}\right]\end{aligned} \quad (18)$$

where, $0 \le \nu_h = (n_h/n'_h) \le 1$ is fixed.

Proof: Let $\hat{\pi}'_{Sh}$ be the estimator in stratum h for the population proportion of the sensitive attribute in the stratified double sampling with sample n_h and $\hat{\pi}'_{Sh}$ be the estimator for the sensitive attribute in the first phase with sample size n'_h we have:

$$\begin{aligned}\hat{\pi}'_S &= \sum_{h=1}^{L} w'_h \hat{\pi}_{Sh} \\ &= \sum_{h=1}^{L} w'_h \hat{\pi}'_{Sh} + \sum_{h=1}^{L} w'_h(\hat{\pi}_{Sh} - \hat{\pi}'_{Sh})\end{aligned} \quad (19)$$

The variance of the first terms is:

$$\begin{aligned}V\left(\sum_{h=1}^{L} w'_h \hat{\pi}'_{Sh}\right) \\ = \frac{1}{n'}\left(\sum_{h=1}^{L} w_h V_{h0} + \sum_{h=1}^{L} w_h(\pi_{Sh} - \pi_S)^2\right)\end{aligned} \quad (20)$$

and using $n_h = \nu_h n'_h = \nu_h w'_h n'$, the second term is:

$$\begin{aligned}E_1\left[V_2\left\{\sum_{h=1}^{L} w'_h(\hat{\pi}_{Sh} - \hat{\pi}'_{Sh})\right\}\right] \\ = E_1\left[\sum_{h=1}^{L}\left(\frac{1}{n_h} - \frac{1}{n'_h}\right) w'^2_h V_{h0}\right] \\ = E_1\left[\sum_{h=1}^{L}\frac{w'_h}{n'}\left(\frac{1}{\nu_h} - 1\right)V_{h0}\right] \\ = \sum_{h=1}^{L}\frac{w_h}{n'}\left(\frac{1}{\nu_h} - 1\right)V_{h0}\end{aligned} \quad (21)$$

Equation 19 is shown by adding Eq. 20 and 21.

In proportional allocation, the sample size n is allocated by the size of the strata, n'_h. Instead of N and N_h using the samples n_h and n'_h in the first phase and the sample size $n_h = n \times n'_h/n'$ the variance of $\hat{\pi}_S$ is given by:

$$V(\hat{\pi}'_S)_P = \frac{1}{n'}\sum_{h=1}^{L} W_h(\pi_{Sh}-\pi_S)^2 + \frac{1}{n}\sum_{h=1}^{L} w_h V_{h0} \tag{22}$$

In the optimum allocation, the cost function is defined as:

$$C = c'n' + \sum_{h=1}^{L} c_h n_h \tag{23}$$

where, c' is a cost for allocation and c_h is a cost per sample unit in stratum h. Since n_h is a random variable, take the expectation of the cost in Eq. 23 to optimize n' and v_h:

$$E(C) = C^* = c'n' + \sum_{h=1}^{L} c_h E(n_h) = c'n' + n'\sum_{h=1}^{L} c_h v_h w_h \tag{24}$$

By the Cauchy-Schwarz inequality, the optimum value for v_h to minimize the product of the variance in Eq. 18 and the expected cost in Eq. 24 is given by:

$$v_h = \sqrt{\frac{c'}{c_h}\times\frac{V_{h0}}{\sum_{h=1}^{L} w_h(\pi_{Sh}-\pi_S)^2}}. \tag{25}$$

Inserting v_h into Eq. 24, the optimum sample size n' is given by:

$$n' = \frac{C^*}{c' + \sum_{h=1}^{L} c_h w_h v_h} \tag{26}$$

Putting Eq. 24 and 26 in Eq. 19 we get the minimum variance of $\hat{\pi}'_S$ as:

$$V(\hat{\pi}'_S)_o = \frac{1}{C^*}\left[\sqrt{c'\sum_{h=1}^{L} w_h(\pi_{Sh}-\pi_S)^2} + \sum_{h=1}^{L} w_h\sqrt{c_h V_{h0}}\right]^2 \tag{27}$$

STRATIFIED SAMPLING AND STRATIFIED DOUBLE SAMPLING-PROPORTION ALLOCATION

From Eq. 25 and 22 that the variance of the proposed estimator $\hat{\pi}'_S$ in stratified double sampling can be expressed as:

$$V(\hat{\pi}'_S)_P = V(\hat{\pi}_S)_P \frac{1}{n'}\sum_{h=1}^{L} w_h(\pi_{Sh}-\pi_S)^2$$

or:

$$\left[V(\hat{\pi}'_S)_P - V(\hat{\pi}_S)_P\right] = \frac{1}{n'}\sum_{h=1}^{L} w_h(\pi_{Sh}-\pi_S)^2 \tag{28}$$

This is the increase variance in proportional allocation from stratified double sampling when the sizes of the stratified populations are unknown.

RELATIVE EFFICIENCY

Singh and Grewal[3] obtained the variance of the estimator $\hat{\pi}_{SG}$ (Singh and Grewal[3]) for the population proportion π_S possessing sensitive attribute based on a Geometric distribution:

$$V(\hat{\pi}_{SG}) = \frac{1}{n}\left[\pi_S(1-\pi_S) + \frac{\left\{\theta_2^*(1-\theta_1^*)\pi_S + \theta_1^*(1-\theta_2^*)(1-\pi_S)\right\}}{(\theta_2^*-\theta_1^*)^2}\right] \tag{29}$$

where, θ_1^* and θ_2^* are known the proportion of cards with the statement "I belong to sensitive group A" and "I do not belong to sensitive group A", respectively.

It assume that there are two strata (i.e., L = 2) in the population, $n = n_1 + n_2$, $\theta_{11}^* = \theta_{21}^* = \theta_1^*$, $\theta_{12}^* = \theta_{22}^* = \theta_2^*$, $\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$ and $\hat{\pi}_S = w_1\hat{\pi}_{S1} + w_2\hat{\pi}_{S2}$. under this supposition, the percent relative efficiency of the proposed estimator π_S with respect to Singh and Grewal[3] estimator $\hat{\pi}_{SG}$ is given by:

$$PRE(\hat{\pi}_S, \hat{\pi}_{SG}) = \frac{V(\hat{\pi}_{SG})}{V(\hat{\pi}_S)}\times 100 = \frac{(1/n)\left[zz_1 + B\right]}{\left[(w_1^2/n_1)\{zz_2 + A_{S1}\} + (w_2^2/n_2)\{zz_3 + A_{S2}\}\right]}\times 100 \tag{30}$$

Where:

$$zz_1 = (w_1\pi_{S1} + w_2\pi_{S2})(1 - w_1\pi_{S1} - w_2\pi_{S2})$$

$$zz_1 = \pi_{S1}(1-\pi_{S1}), zz_2 = \pi_{S2}(1-\pi_{S2})$$

$$B = \frac{\left\{\theta_2^{*2}(1-\theta_1^*)\pi_S + \theta_1^{*2}(1-\theta_2^*)(1-\pi_S)\right\}}{(\theta_2^*-\theta_1^*)^2}$$

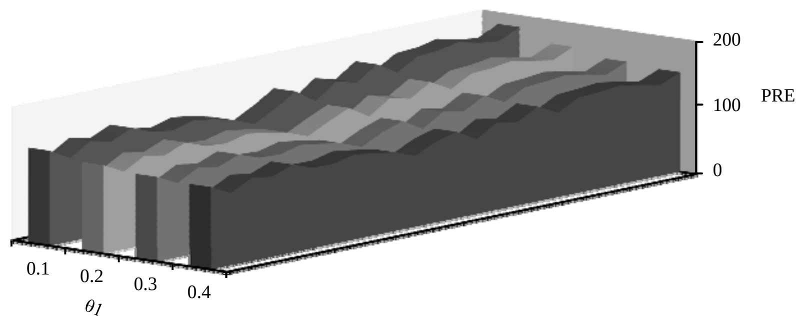

Fig. 1: Relative efficiency (%) of $\hat{\pi}_{SG}$ with respect to $\hat{\pi}_S$

Table 1: Relative efficiency (%) of $\hat{\pi}_{SG}$ with respect to $\hat{\pi}_S$ when n =1000

π_{S1}	π_{S2}	w_1	w_2	n_1	n_2	$\theta_1^* = 0.1$, $\theta_2^* = 0.5$	$\theta_1^* = 0.2$, $\theta_2^* = 0.6$	$\theta_1^* = 0.3$, $\theta_2^* = 0.7$	$\theta_1^* = 0.4$, $\theta_2^* = 0.8$
0.10	0.90	0.70	0.30	500	500	182.02	163.95	153.41	149.27
0.11	0.89	0.70	0.30	400	600	167.20	150.27	140.28	136.65
0.12	0.88	0.70	0.30	500	500	171.00	155.75	146.72	143.11
0.13	0.87	0.60	0.40	400	600	178.63	162.16	152.35	148.11
0.14	0.86	0.60	0.40	500	500	171.78	157.40	148.88	144.85
0.15	0.85	0.60	0.40	400	600	169.96	155.47	146.76	142.94
0.16	0.84	0.40	0.60	500	500	151.75	138.60	130.89	126.61
0.17	0.83	0.40	0.60	400	600	170.30	155.54	146.81	142.13
0.18	0.82	0.40	0.60	500	500	149.27	136.83	129.51	125.44
0.19	0.81	0.30	0.70	400	600	157.72	143.88	135.69	131.13
0.20	0.80	0.30	0.70	500	500	134.42	122.91	116.11	112.26
0.21	0.79	0.30	0.70	400	600	155.18	142.02	134.21	129.86
0.22	0.78	0.70	0.30	500	500	135.73	128.08	123.33	121.28
0.23	0.77	0.70	0.30	400	600	118.85	112.26	108.12	106.42
0.24	0.76	0.70	0.30	500	500	131.10	124.27	120.01	118.14
0.25	0.75	0.60	0.40	400	600	139.06	130.68	125.50	123.10
0.26	0.74	0.60	0.40	500	500	143.21	135.01	129.99	127.50
0.27	0.73	0.60	0.40	400	600	134.59	126.97	122.24	120.03
0.28	0.72	0.40	0.60	500	500	138.88	129.36	123.66	120.49
0.29	0.71	0.40	0.60	400	600	151.32	141.08	134.88	131.54
0.30	0.70	0.40	0.60	500	500	137.16	128.13	122.70	119.67
0.31	0.69	0.30	0.70	400	600	144.21	133.94	127.75	124.31
0.32	0.68	0.30	0.70	500	500	125.15	116.42	111.17	108.21
0.33	0.67	0.30	0.70	400	600	142.34	132.56	126.65	123.37

or:

$$B = \frac{\left\{\theta_2^{*2}(1-\theta_1^*)zz_1 + \theta_1^{*2}(1-\theta_2^*)zz_1\right\}}{(\theta_2^* - \theta_1^*)^2}$$

$$A_{S1} = \frac{\left\{\theta_2^{*2}(1-\theta_1^*)\pi_{S1} + \theta_1^{*2}(1-\theta_2^*)(1-\pi_{S1})\right\}}{(\theta_2^* - \theta_1^*)^2}$$

$$A_{S2} = \frac{\left\{\theta_2^{*2}(1-\theta_1^*)\pi_{S2} + \theta_1^{*2}(1-\theta_2^*)(1-\pi_{S2})\right\}}{(\theta_2^* - \theta_1^*)^2}$$

This study computed the PRE ($\hat{\pi}_S$, $\hat{\pi}_{SG}$) for n = 1000 with two strata having θ_2^* = 0.5, 0.6, 0.7, 0.8 and θ_1^* = 0.1, 0.2, 0.3, 0.4 and different values of n_1, n_2; w_1, w_2 and $\hat{\pi}_{S1}$, $\hat{\pi}_{S2}$. Findings are displaced in Table 1. It is observed from Table 1 that the percent relative efficiencies are greater than 100 which means that the proposed estimator $\hat{\pi}_S$ is more efficient than the Singh and Grewal[3] estimator $\hat{\pi}_{SG}$. This fact is also presented in Fig. 1. However this is limited empirical study so this conclusion should not be extrapolated.

Further we will do an efficiency comparison of the proposed stratified randomized response technique with that of Singh and Grewal[3] by comparing the variances.

Theorem 5.1: Assume that there are two strata (i.e., L = 2) in the population, $n = n_1 + n_2$, $\theta_{11}^* = \theta_{21}^* = \theta_1^*$, $\theta_{12}^* = \theta_{22}^* = \theta_2^*$,

$\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$, $\hat{\pi}_S = w_1\hat{\pi}_{S1} + w_2\hat{\pi}_{S2}$ and $\pi_{S1} \neq \pi_{S2}$. The proposed estimator $\hat{\pi}_S$ under proportional allocation is always better than Singh and Grewal[3] estimator $\hat{\pi}_{SG}$.

Proof: From Singh and Grewal[3], the variance of $\hat{\pi}_{SG}$ for two strata (i.e., L = 2) in the population, is given by:

$$V(\hat{\pi}_{SG}) = \left[\frac{\pi_S(1-\pi_S)}{n} + B\right] \tag{31}$$

where, $\hat{\pi}_{SG}$ stands for Singh and Grewal[3] estimator, $\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$ we have:

$$\begin{aligned}
& n\left[V(\hat{\pi}_{SG}) - V(\hat{\pi}_S)_P\right] = \\
& [\pi_S(1-\pi_S) - w_1\pi_{S1}(1-\pi_{S1}) \\
& -w_2\pi_{S2}(1-\pi_{S2}) + B - w_1A_{S1} - w_2A_{S2}] \\
& = [(w_1\pi_{S1} + w_2\pi_{S2})(1 - w_1\pi_{S1} - w_2\pi_{S2}) \\
& -w_1\pi_{S1}(1-\pi_{S1}) - w_2\pi_{S2}(1-\pi_{S2}) + B \\
& -w_1A_{S1} - w_2A_{S2}] \\
& = [w_1w_2(\pi_{S1} - \pi_{S2})^2 \\
& +(B - w_1A_{S1} - w_2A_{S2}] \\
& = w_1w_2(\pi_{S1} - \pi_{S2})^2 + 0 \\
& = w_1w_2(\pi_{S1} - \pi_{S2})^2
\end{aligned}$$

which is always positive. Hence, the theorem.

Theorem 5.2: Suppose that there are two strata (i.e., L = 2) in the population, $n = n_1 + n_2$, $\theta_{11}^* = \theta_{21}^* = \theta_1^*$, $\theta_{12}^* = \theta_{22}^* = \theta_2^*$, $\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$, $\hat{\pi}_S = w_1\hat{\pi}_{S1} + w_2\hat{\pi}_{S2}$ and $\pi_{S1} \neq \pi_{S2}$. The proposed estimator $\hat{\pi}_S$ under Neyman allocation is always better than Singh and Grewal[3] estimator $\hat{\pi}_{SG}$.

Proof: From Singh and Grewal[3], the variance of $\hat{\pi}_{SG}$ for two strata (i.e., L = 2) in the population, is given by:

$$V(\hat{\pi}_{SG}) = \left[\frac{\pi_S(1-\pi_S)}{n} + B\right] \tag{32}$$

where, $\hat{\pi}_{SG}$ stands for Singh and Grewal[3] estimator and $\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$.

For L = 2 and the assumptions stated in the theorem 5.2, we write the variance of the proposed estimator $\hat{\pi}_S$ under Neyman allocation as:

$$V(\hat{\pi}_S)_N = \frac{1}{n}\left[\begin{array}{l} w_1\{\pi_{S1}(1-\pi_{S1}) + A_{S1}\}^{1/2} \\ +w_2\{\pi_{S2}(1-\pi_{S2}) + A_{S2}\}^{1/2} \end{array}\right]^2 \tag{33}$$

From above equations, we have:

$$\begin{aligned}
& n\left[V(\hat{\pi}_{SG}) - V(\hat{\pi}_S)_N\right] \\
& = \Big[(w_1\pi_{S1} + w_2\pi_{S2})\{1 - w_1\pi_{S1} - w_2\pi_{S2}\} \\
& \quad + \frac{\left\{\theta_2^{*2}(1-\theta_1^*)zz_1 + \theta_1^{*2}(1-\theta_2^*)zz_2\right\}}{(\theta_2^* - \theta_1^*)^2} \\
& \quad -w_1^2\pi_{S1}(1-\pi_{S1}) - w_2^2\pi_{S2}(1-\pi_{S2}) \\
& \quad -w_1^2A_{S1} - w_1^2A_{S2} - 2w_1w_2\sqrt{A_{S1}^*A_{S2}^*}\Big]
\end{aligned} \tag{34}$$

where, $A_{S1}^* = \left[\pi_{S1}(1-\pi_{S1}) + A_{S1}\right]$ and $A_{S2}^* = \left[\pi_{S2}(1-\pi_{S2}) + A_{S2}\right]$.

Now we have:

$$\begin{aligned}
& n\left[V(\hat{\pi}_{SG}) - V(\hat{\pi}_S)_N\right] = w_1w_2(\pi_{S1} + \pi_{S2} - 2\pi_{S1}\pi_{S2}) \\
& \quad + \frac{w_1w_2}{(\theta_2^* - \theta_1^*)^2} \\
& \left[\begin{array}{l} (\pi_{S1} + \pi_{S2}) \\ \{\theta_2^{*2}(1-\theta_1^*) - \theta_1^{*2}(1-\theta_2^*)\} + 2\theta_1^{*2}(1-\theta_2^*) \end{array}\right] \\
& \quad - w_1w_2\sqrt{A_{S1}^*A_{S2}^*} \\
& = w_1w_2\left[(\pi_{S1} + \pi_{S2})^2 + \left(\sqrt{A_{S1}^*} - \sqrt{A_{S2}^*}\right)^2\right] \\
& + \frac{w_1w_2}{(\theta_2^* - \theta_1^*)^2}\left[\begin{array}{l} (\pi_{S1} + \pi_{S2}) \\ \{\theta_2^{*2}(1-\theta_1^*) - \theta_1^{*2}(1-\theta_2^*)\} \\ +2\theta_1^{*2}(1-\theta_2^*) \end{array}\right. \\
& -\theta_2^{*2}(1-\theta_1^*)\pi_{S1} - \theta_1^{*2}(1-\theta_2^*)(1-\pi_{S1}) \\
& \left. -\theta_2^{*2}(1-\theta_{21}^*)\pi_{S2} - \theta_1^{*2}(1-\theta_2^*)(1-\pi_{S2})\right] \\
& = w_1w_2\left[(\pi_{S1} + \pi_{S2})^2 + \left(\sqrt{A_{S1}^*} - \sqrt{A_{S2}^*}\right)^2\right] + 0 \\
& = w_1w_2\left[(\pi_{S1} + \pi_{S2})^2 + \left(\sqrt{A_{S1}^*} - \sqrt{A_{S2}^*}\right)^2\right]
\end{aligned}$$

Which is always positive.

Theorem 5.3: The proposed estimator $\hat{\pi}_S$ under Neyman allocation is always better than that of the proposed estimator $\hat{\pi}_S$ under proportional allocation.

Proof: We have:

$$\begin{aligned}
& n\left[V(\hat{\pi}_S)_P - V(\hat{\pi}_S)_N\right] \\
& = \sum_{h=1}^{L} w_h V_{h0} - \left[\sum_{h=1}^{L} w_h\sqrt{V_{h0}}\right]^2 \\
& = \sum_{h=1}^{L} w_h\left\{V_{h0} - \sum_{h=1}^{L} w_h\sqrt{V_{h0}}\right\}^2
\end{aligned}$$

which is always positive. Hence the theorem.

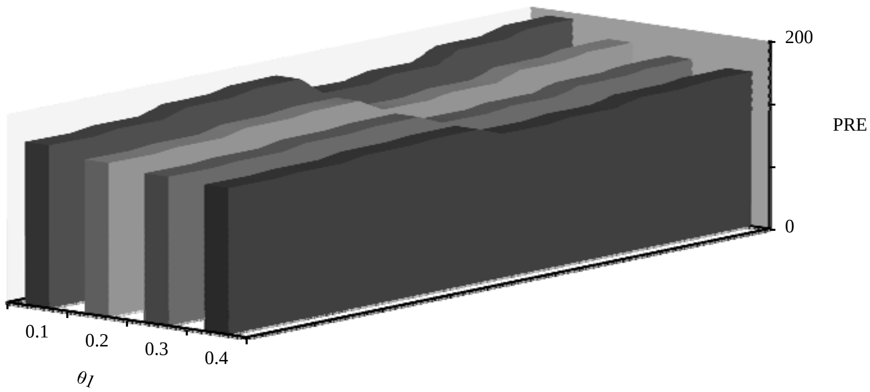

Fig. 2: Relative efficiency (%) of $\hat{\pi}_{sG}$ with respect to $(\hat{\pi}_S)_P$

Table 2: Relative efficiency (%) of $\hat{\pi}_{sG}$ with respect to $(\hat{\pi}_S)_P$

π_{S1}	π_{S2}	w_1	w_2	$\theta_1^* = 0.1$, $\theta_2^* = 0.5$	$\theta_1^* = 0.2$, $\theta_2^* = 0.6$	$\theta_1^* = 0.3$, $\theta_2^* = 0.7$	$\theta_1^* = 0.4$, $\theta_2^* = 0.8$
0.10	0.90	0.70	0.30	144.99	133.10	126.71	123.67
0.11	0.89	0.70	0.30	144.75	132.95	126.61	123.59
0.12	0.88	0.70	0.30	144.51	132.81	126.50	123.50
0.13	0.87	0.60	0.40	139.63	129.67	124.21	121.50
0.14	0.86	0.60	0.40	139.50	129.58	124.14	121.44
0.15	0.85	0.60	0.40	139.37	129.49	124.08	121.39
0.16	0.84	0.40	0.60	130.21	123.03	119.01	116.93
0.17	0.83	0.40	0.60	130.35	123.13	119.09	117.00
0.18	0.82	0.40	0.60	130.48	123.22	119.17	117.07
0.19	0.81	0.30	0.70	126.38	120.21	116.75	114.92
0.20	0.80	0.30	0.70	126.66	120.42	116.91	115.06
0.21	0.79	0.30	0.70	126.93	120.62	117.07	115.21
0.22	0.78	0.70	0.30	142.05	131.26	125.40	122.54
0.23	0.77	0.70	0.30	141.80	131.10	125.28	122.43
0.24	0.76	0.70	0.30	141.54	130.93	125.15	122.33
0.25	0.75	0.60	0.40	138.08	128.61	123.40	120.80
0.26	0.74	0.60	0.40	137.95	128.52	123.34	120.74
0.27	0.73	0.60	0.40	137.82	128.43	123.27	120.68
0.28	0.72	0.40	0.60	131.83	124.21	119.95	117.76
0.29	0.71	0.40	0.60	131.97	124.30	120.03	117.83
0.30	0.70	0.40	0.60	132.10	124.40	120.11	117.90
0.31	0.69	0.30	0.70	129.67	122.63	118.70	116.65
0.32	0.68	0.30	0.70	129.94	122.83	118.85	116.79
0.33	0.67	0.30	0.70	130.21	123.03	119.01	116.93

To have the tangible idea about the performance of the proposed estimator $\hat{\pi}_S$ (under different allocations) over the estimator $\hat{\pi}_{SG}$ due to Singh and Grewal[3]. It have computed the PRE($(\hat{\pi}_S)_P$, $\hat{\pi}_{SG}$), PRE($(\hat{\pi}_S)_N$, $\hat{\pi}_{SG}$) and PRE($(\hat{\pi}_S)_N$, $(\hat{\pi}_S)_P$) for different values of w_1, w_2; $\hat{\pi}_{S1}$, $\hat{\pi}_{S2}$ and findings are publicized in Table 2-4 using following equation:

$$\text{PRE}((\hat{\pi}_S)_P, \hat{\pi}_{SG}) = \frac{V(\hat{\pi}_{SG})}{V(\hat{\pi}_S)_P} \times 100$$

$$= \frac{[\pi_S(1-\pi_S)+B]}{(w_1 A_{S1}^* + w_2 A_{S2}^*)} \times 100$$

$$\text{PRE}((\hat{\pi}_S)_N, \hat{\pi}_{SG}) = \frac{V(\hat{\pi}_{SG})}{V(\hat{\pi}_S)_N} \times 100$$

$$= \frac{[\pi_S(1-\pi_S)+B]}{\left[(w_1\sqrt{A_{S1}^*} + w_2\sqrt{A_{S2}^*})\right]^2} \times 100$$

$$\text{PRE}((\hat{\pi}_S)_N, (\hat{\pi}_S)_P) = \frac{V(\hat{\pi}_S)_P}{V(\hat{\pi}_S)_N} \times 100$$

$$= \frac{(w_1 A_{S1}^* + w_1 A_{S2}^*)}{\left[(w_1\sqrt{A_{S1}^*} + w_2\sqrt{A_{S2}^*})\right]^2} \times 100$$

It is observed from Table 2, 3 that the relative efficiency (%) are greater than 100 which follows that

Table 3: Relative efficiency of $\hat{\pi}_{SG}$ with respect to $(\hat{\pi}_S)_N$

π_{S1}	π_{S2}	w_1	w_2	$\theta_1^* = 0.1$ $\theta_2^* = 0.5$	$\theta_1^* = 0.2$ $\theta_2^* = 0.6$	$\theta_1^* = 0.3$ $\theta_2^* = 0.7$	$\theta_1^* = 0.4$ $\theta_2^* = 0.8$
0.10	0.90	0.70	0.30	115.62	120.84	126.42	130.76
0.11	0.89	0.70	0.30	114.93	120.43	126.14	130.50
0.12	0.88	0.70	0.30	114.33	120.07	125.90	130.29
0.13	0.87	0.60	0.40	119.97	126.84	133.35	138.19
0.14	0.86	0.60	0.40	119.33	126.35	132.94	137.79
0.15	0.85	0.60	0.40	118.75	125.91	132.56	137.42
0.16	0.84	0.40	0.60	123.85	133.01	140.95	146.64
0.17	0.83	0.40	0.60	123.32	132.50	140.46	146.13
0.18	0.82	0.40	0.60	122.82	132.02	139.98	145.63
0.19	0.81	0.30	0.70	123.33	133.43	141.97	147.99
0.20	0.80	0.30	0.70	122.93	133.00	141.52	147.51
0.21	0.79	0.30	0.70	122.55	132.58	141.08	147.04
0.22	0.78	0.70	0.30	111.48	118.67	125.28	129.86
0.23	0.77	0.70	0.30	111.40	118.68	125.34	129.94
0.24	0.76	0.70	0.30	111.34	118.70	125.42	130.03
0.25	0.75	0.60	0.40	115.10	123.09	130.22	135.14
0.26	0.74	0.60	0.40	114.89	122.94	130.09	135.03
0.27	0.73	0.60	0.40	114.70	122.80	129.99	134.93
0.28	0.72	0.40	0.60	118.85	128.06	135.99	141.47
0.29	0.71	0.40	0.60	118.53	127.73	135.66	141.12
0.30	0.70	0.40	0.60	118.22	127.42	135.33	140.78
0.31	0.69	0.30	0.70	119.15	128.82	137.05	142.71
0.32	0.68	0.30	0.70	118.85	128.48	136.68	142.31
0.33	0.67	0.30	0.70	118.55	128.14	136.31	141.91

Table 4: Relative efficiency of $(\hat{\pi}_S)_P$ with respect to $(\hat{\pi}_S)_N$

π_{S1}	π_{S2}	w_1	w_2	$\theta_1^* = 0.1$ $\theta_2^* = 0.5$	$\theta_1^* = 0.2$ $\theta_2^* = 0.6$	$\theta_1^* = 0.3$ $\theta_2^* = 0.7$	$\theta_1^* = 0.4$ $\theta_2^* = 0.8$
0.10	0.90	0.70	0.30	274.11	324.46	358.85	363.42
0.11	0.89	0.70	0.30	274.40	324.78	359.19	363.79
0.12	0.88	0.70	0.30	274.69	325.09	359.53	364.16
0.13	0.87	0.60	0.40	283.00	333.55	368.26	374.31
0.14	0.86	0.60	0.40	283.07	333.64	368.37	374.41
0.15	0.85	0.60	0.40	283.14	333.74	368.47	374.51
0.16	0.84	0.40	0.60	296.82	347.80	383.06	391.31
0.17	0.83	0.40	0.60	296.50	347.50	382.75	390.94
0.18	0.82	0.40	0.60	296.20	347.20	382.44	390.57
0.19	0.81	0.30	0.70	301.71	352.92	388.40	397.38
0.20	0.80	0.30	0.70	301.23	352.45	387.91	396.80
0.21	0.79	0.30	0.70	300.76	351.97	387.42	396.23
0.22	0.78	0.70	0.30	277.82	328.44	363.05	368.09
0.23	0.77	0.70	0.30	278.15	328.79	363.41	368.50
0.24	0.76	0.70	0.30	278.48	329.14	363.78	368.92
0.25	0.75	0.60	0.40	284.11	334.88	369.72	375.80
0.26	0.74	0.60	0.40	284.23	335.02	369.86	375.95
0.27	0.73	0.60	0.40	284.35	335.15	370.01	376.11
0.28	0.72	0.40	0.60	293.31	344.35	379.55	387.11
0.29	0.71	0.40	0.60	293.03	344.09	379.27	386.78
0.30	0.70	0.40	0.60	292.77	343.82	379.00	386.46
0.31	0.69	0.30	0.70	296.17	347.34	382.66	390.65
0.32	0.68	0.30	0.70	295.73	346.89	382.19	390.11
0.33	0.67	0.30	0.70	295.28	346.44	381.73	389.57

the proposed estimator $\hat{\pi}_S$ under proportional and Neyman allocations are more efficient than Singh and Grewal[3] estimator $\hat{\pi}_{SG}$ with considerable gain in efficiency. Table 4 also showed that the values of PRE($(\hat{\pi}_S)_N$, $(\hat{\pi}_S)_P$) are greater than 100 which means that the proposed estimator $\hat{\pi}_S$ under Neyman allocation is more efficient than that of under proportional allocation. These facts can be seen from Fig. 2-4.

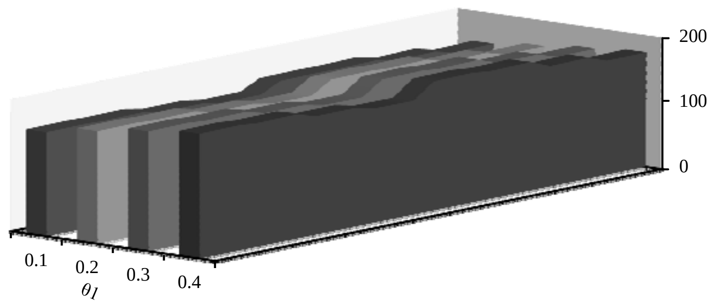

Fig. 3: Relative efficiency (%) of $\hat{\pi}_{sG}$ with respect to $(\hat{\pi}_S)_N$

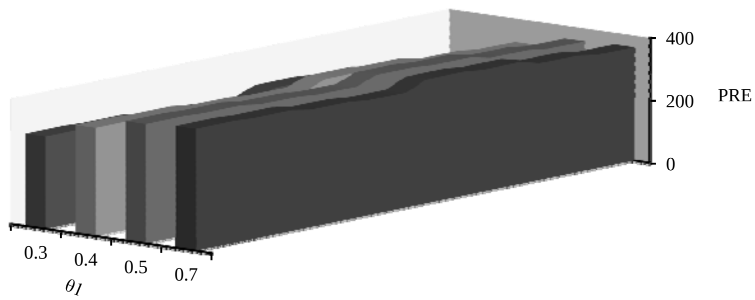

Fig. 4: Relative efficiency (%) of $(\hat{\pi}_S)_P$ with respect to $(\hat{\pi}_S)_N$

Thus our recommendation is to prefer the proposed study over Singh and Grewal[3].

CONCLUSION

This paper addresses the problem of estimating the population proportion of sensitive attribute $\hat{\pi}_S$ based on stratified sampling schemes (stratified sampling and stratified double sampling). It has been shown theoretically as well as numerically that the proposed model is more efficient than Singh and Grewal's randomized response model.

SIGNIFICANCE STATEMENTS

This study discovers a new stratified randomized response model and random sampling is generally obtained by dividing the population into non overlapping groups called strata and selecting a simple random sample from each stratum. An RR technique using a stratified random sampling gives the group characteristics related to each stratum estimator. Also, stratified sample protect a researcher from the possibility of obtaining a poor sample. This study will help the researchers to uncover the critical areas related to randomized response technique applying geometric distribution. For the future research, researcher can be considering a new theory for randomized response model with hyper geometric distribution.

REFERENCES

1. Warner, S.L., 1965. Randomized response: A survey technique for eliminating evasive answer bias. J. Am. Stat. Assoc., 60: 63-69.
2. Kuk, A.Y.C., 1990. Asking sensitive questions indirectly. Biometrika, 77: 436-438.
3. Singh, S. and I.G. Grewal, 2013. Geometric distribution as a randomization device: Implemented to the Kuk's model. Int. J. Contemp. Math. Sci., 8: 243-248.
4. Land, M., S. Singh and S.A. Sedory, 2012. Estimation of a rare sensitive attribute using poisson distribution. Statistics: J. Theor. Applied Stat., 46: 351-360.
5. Lee, G.S., D. Uhm and J.M. Kim, 2013. Estimation of a rare sensitive attribute in a stratified sample using poisson distribution. Statistics: J. Theor. Applied Stat., 47: 575-589.
6. Fox, J.A. and P.E. Tracy, 1986. Randomized Response: A Method of Sensitive Surveys. SEGE Publications, Newbury Park, CA.
7. Tarray, T.A., 2017. Scrutinize on Stratified Randomized Response Technique. GRIN Verlag, Munich, ISBN: 9783668554580.
8. Naik, P.A. and K.R. Pardasani, 2018. Three-dimensional finite element model to study effect of RyR calcium channel, ER leak and SERCA pump on calcium distribution in oocyte cell. Int. J. Comput. Methods, Vol. 15, No. 6. 10.1142/S0219876218500913.

9. Naik, P.A. and K.R. Pardasani, 2013. Finite element model to study effect of buffers in presence of voltage gated Ca^{2+} channels on calcium distribution in oocytes for one dimensional unsteady state case. Int. J. Modern Biol. Med., 4: 190-203.
10. Singh, H.P. and T.A. Tarray, 2013. An alternative to Kim and Warde's mixed randomized response technique. Stataistica, 73: 379-402.
11. Singh, H.P. and T.A. Tarray, 2015. Two-stage stratified partial randomized response strategies. Commun. Stat.-Theory Methods. 10.1080/03610926.2013.804571
12. Singh, H.P. and T.A. Tarray, 2014. An improvement over Kim and Elam stratified unrelated question randomized response model using neyman allocation. Sankhya B, 77: 91-107.
13. Tarray, T.A. and H.P. Singh, 2015. Some improved additive randomized response models utilizing higher order moments ratios of scrambling variable. Model Assist. Stat. Applic., 10: 361-383.
14. Tarray, T.A. and H.P. Singh, 2017. An adept-stratified Kuk's randomized response model using Neyman allocation. Commun. Stat.-Theory Methods, 46: 2870-2881.
15. Tarray, T.A. and H.P. Singh, 2018. A randomization device for estimating a rare sensitive attribute in stratified sampling using Poisson distribution. Afrika Matematika, 29: 407-423.
16. Yuzbasi, S. and M. Karacayir, 2018. A numerical approach for solving high-order linear delay volterra integro-differential equations. Int. J. Comput. Methods, Vol. 15, No. 5. 10.1142/S0219876218500421.
17. Zaizai, Y., W. Jingyu, L. Junfeng and W. Hua, 2008. Ratio imputation method for handling item-nonresponse in Eichhorn model. Model Assist. Stat. Applic., 3: 89-98.
18. Cochran, W.G., 1977. Sampling Techniques. 3rd Edn., John Wiley and Sons Inc., New York, ISBN-13: 978-0-471-16240-7.