Journal of Modern Applied Statistical Methods

Volume 21 | Issue 1 Article 4

Partial Randomized Response Model for Estimating a Rare Sensitive Attribute in Probability Proportional to Size Measures Using Poisson Distribution

Preeti Patidar
Department of Mathematics, SSBSR, Sharda University, India,
p.patidarstatujn@gmail.com

Housila P Singh
School of Studies in Statistics, Vikram University, India,
hpsujn@gmail.com

Recommended Citation

Preeti Patidar, Housila P Singh (2022). Partial Randomized Response Model for Estimating a Rare Sensitive Attribute in Probability Proportional to Size Measures Using Poisson Distribution. Journal of Modern Applied Statistical Methods, 21(1), https://doi.org/10.56801/Jmasm.V21.i1.4

Journal of Modern Applied Statistical Methods
2022, Vol. 21, No. 1,
Doi: 10.56801/Jmasm.V21.i1.4

ISSN 1538 – 9472

Partial Randomized Response Model for Estimating a Rare Sensitive Attribute in Probability Proportional to Size Measures Using Poisson Distribution

Preeti Patidar
Department of Mathematics, SSBSR,
Sharda University, India

Housila P Singh
School of Studies in Statistics,
Vikram University, India

In this paper motivated by Lee et al (2013) and Singh et al (2017,2020), we have suggested the estimation procedures of mean number of persons possessing sensitive attribute using Narjis and Shabbir (2020) randomized response model for the population which comprises of some clusters and the population is stratified with some clusters in each stratum. The estimator for the mean number of persons possessing sensitive attribute under a Poisson distribution, its variance, and the estimator of the variance are proposed under two-stage and stratified two-stage sampling schemeswhen the parameter of the rare sensitive attribute is pretended to be known and unknown. We employ the sampling scheme with probability proportional to size to select the first-stage units and simple random sampling with replacement to select the second-stage units.Theperformance of the suggested estimation procedures are demonstrated through numerical illustration over Singh and Suman (2019) estimators.

Keywords: Poisson distribution, Randomized response model, Probability proportional to size, Rare sensitive attribute, Unrelated non-sensitive attribute, Two-stage sampling, Stratified two-stage sampling.

1. Introduction

Concerns in the context of privacy and confidentiality often provide substantial non-responses and false responses, especially in surveys that ask direct questions about sensitive attributes such as use of illegal cannabis plant, cheating in exams, extra marital affairs, domestic violence, sexual behavior, mental disorder, criminal history, tax evasion, drug abuse, gambling, abortion and others. The randomized response technique is quite effective in reducing the non-response rate and inflated response bias.Warner (1965) was first to introduce randomized response technique (RRT) which uses a randomization device bearing two statements, one on sensitive character and other on its complement and sample units are drawn by simple random sampling with replacement (SRSWR) procedure.

The RRT models can be divided into three categories such as full, partial and optional. In full RRT model, all respondents are urged to use the randomization device and give the response as per the question (i.e. sensitive and non-sensitive) occurred on randomization device, whereas, in partial RRT model, an additional stage or randomization device is added to incorporate the element of truthful response. In two stage or partial RRT model, the first randomization device has two options: (i) Do you belong to the sensitive group? and (ii) Go to the second randomization device, with known probabilities U and (1-U) respectively. The second randomization device is exactly same as provided in full RRT model. In optional RRT model, respondents are urged to give the truthful answer if he/she considers the question to be non-sensitive and use the randomization device if he/she considers the question sensitive and give the answer after using randomization device [see, Narjis and Shabbir (2019), p.1].

We consider the problem where the number of persons possessing a rare sensitive attribute is very small and the large sample size is necessary for providing the enough précised estimate of this number. For instance, the proportion of AIDS patients who continue having affairs with strangers, the proportion of persons who have witnessed a murder, the proportion of persons who are told by their doctors that they will not survive long due to ghastly disease etc. Now days the communication system is increasing rapidly, so it is possible and easy to conduct such large randomized surveys over the internet or telephone etc. Land et al (2012) was first to consider the problem of estimating the mean total number of persons who possesses a rare sensitive attribute in the population. Later various authors including Singh and Tarray (2014, 2017), Tarray and Singh (2015), Singh et al (2019), Tarray et al (2019) among others have tackled this problem.

Lee et al (2013) have proposed a variant of Land et al's (2012) randomized response model when a population consists of some clusters and the population is stratified with some clusters in each stratum. In this paper following Lee et al (2013) we have made an effort to extend Narjis and Shabbir (2019) partial randomized response model when a population comprises of some clusters and population is stratified with some clusters in each stratum. We have derived the estimator for the mean number of persons who possess a rare sensitive attribute along with its variance and the variance estimator when the parameter of a rare unrelated attribute is supposed to be known and not known. The clusters are drawn with and without replacement.

2. Probability proportional to size (PPS) sampling scheme for estimating the rare sensitive parameter under Poisson distribution

We consider a finite population $\Omega = (\Omega_1, \Omega_2, ..., \Omega_N)$ of N clusters, known as first-stage units. The size of the ith cluster is M_i $(i = 1, 2, ..., N)$ termed as second-stage units. Selection of n first- stage units (i.e. of n cluster) are made employing PPS sampling scheme with probabilities p_i $(i = 1, 2, ..., n)$. At the second –stage we draw m_i $(i = 1, 2, ..., n)$ second-stage units from the ith selected first-stage unit using

SRSWR scheme. We designate.

π_a : The true proportion of persons having a rare sensitive attribute A,

π_y : The true proportion of persons having a unrelated rare non-sensitive attribute Y,

π_{ia} : The true proportion of persons with the rare sensitive attribute in the ith cluster,

π_{iy} : The true proportion of persons with the unrelated rare non-sensitive attribute in the ith cluster.

We also denote

$$M_0 = \sum_{i=1}^{N} M_i \text{ and } m = \sum_{i=1}^{n} m_i \text{ .}$$

[see, Lee et al (2013) and Singh et al (2020)].

3. Estimation of a rare sensitive attribute in a two-stage unrelated question RRT model

We have discussed the estimation methods for mean number of persons possessing a rare sensitive attribute using Narjis and Shabbir (2020) randomized device when the clusters are drawn with and without replacement depending on the cluster sizes and with equal probability (i.e. Simple random sampling with replacement). We have investigated the unbiased estimators for the mean number of individuals and their properties are studied when the unrelated rare innocuous attribute Y is assumed to be known and unknown. Response from the elementary units in the second stage samples are obtained on employing randomized response device of Narjis and Shabbir (2020).

3.1 When the unrelated rare innocuous attribute is known

When the population proportion π_y of persons having the unrelated rare attribute is known, respondents are requested to use the randomization device and answer without revealing their having the attribute or not. In the proposed model, each selected respondent in the sample from the ith cluster has been given two randomization devices (R_1, R_2).

Assuming the proportion of rare non-sensitive unrelated attribute is known, the responses from the elementary units in the second stage sample were collected using the Mangat (1992) randomization device which comprises the following statements for the ith cluster (i.e. the first-stage randomization device, R_1 has two kinds of statements for the ith cluster):

(i) Do you possess rare sensitive attribute A?

(ii) Go to the randomization device R_2

with probabilities U_i and $(1-U_i)$ respectively.

The second-stage randomization device, R_2 consists of three statements:

(i) Do you have the rare sensitive attribute A?

(ii) Do you have the rare non-sensitive unrelated attribute Y?

(iii) Draw one more card

with corresponding probabilities P_{1i}, P_{2i} and P_{3i} respectively such that $\sum_{j=1}^{3} P_{ji} = 1,\ j = 1,2,3.$ If statement (iii) appeared then respondent repeat the process without replacing the card. Again, if in second draw the statement (iii) appeared the respondent is asked to report his/her actual status about sensitive attribute A. The respondent should answer the question with "yes" (or "no"), if his/her actual status matches (un-matches) with the statement on the card respectively. The investigator does not know the respondent's answers, whether its from the sensitive question or from the actual status because interviewee performed randomization process confidentialy. Thus the privacy of interviewee(s) is protected and they responds without any fear [see, Narjis and Shabbir (2020), p.2].

The probability of obtaining answer "yes" from the respondent in the i^th^ cluster is :

$$\theta_{i0} = U_i \pi_{ia} + (1-U_i)\left[(P_{1i}\pi_{ia} + P_{2i}\pi_{iy})\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2 \frac{k_i}{(k_i-1)}\pi_{ia} \right] \tag{1}$$

where the randomization device R_2 consists of a deck of k_i cards provided to the respondents selected from the i^{th} cluster.

Consider selecting a large sample of persons from the i^{th} cluster in the population such that $m_i \to \infty, \pi_{ia} \to 0, \pi_{iy} \to 0$, then $m_i\pi_{ia} = \lambda_{ia}, m_i\pi_{iy} = \lambda_{iy}$ *and* $\theta_{i0} \to 0$,then $m_i\theta_{i0} = \lambda_{i0}$ (finite),

$$\lambda_{i0} = U_i \lambda_{ia} + (1-U_i)\left[(P_{1i}\lambda_{ia} + P_{2i}\lambda_{iy})\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2 \frac{k_i}{(k_i-1)}\lambda_{ia} \right], \tag{2}$$

Let $y_{i1}, y_{i2},..., y_{im_i}$ be a random sample drawn from i^{th} cluster which follows Poisson distribution with mean λ_{i0}. Then the estimator for the mean number of individuals with the rare sensitive characteristics, $\lambda_{ia}(\lambda_{ia} = m_i\pi_{ia})$ is defined as

$$\hat{\lambda}_{ia} = \frac{\left[(1/m_i)\sum_{j=1}^{m_i} y_{ij} - (1-U_i)P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}\right]}{\left[U_i + (1-U_i)\left\{P_{1i} + P_{3i}\frac{k_i}{(k_i-1)}(P_{1i}+P_{3i})\right\}\right]} \quad (3)$$

where $\lambda_{iy}\left(\lambda_{iy} = m_i\pi_{iy}\right)$ is the mean number of individuals who have the rare non-sensitive unrelated attribute in the i^{th} cluster. In the two-stage procedure, the estimator for the mean number of persons with the rare sensitive attribute is given by

$$\hat{\lambda}_{appzwr} = \frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{ia}}{p_i} = \frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}\left[\frac{(1/m_i)\sum_{j=1}^{m_i} y_{ij} - (1-U_i)P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\lambda_{iy}\right\}}{U_i + (1-U_i)\left\{P_{1i} + P_{3i}\frac{k_i}{(k_i-1)}(P_{1i}+P_{3i})\right\}}\right] \quad (4)$$

where M_i is the size of the i^{th} cluster and $M_0 = \sum_{i-1}^{n} M_i$; and p_i is the probability of drawing i^{th}(i=1, 2,3,…,N) cluster from the population in first stage sample.

Theorem 3.1 The estimator $\hat{\lambda}_{appzwr}$ for the mean number of persons who have the rare sensitive attribute is unbiased.

Proof: Since y_{ij} follows Poisson distribution with parameter

$$\lambda_{i0} = U_i\lambda_{ia} + (1-U_i)\left[\left(P_{1i}\lambda_{ia} + P_{2i}\lambda_{iy}\right)\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right]$$

Thus $E\left(\hat{\lambda}_{appzwr}\right) = E_1E_2\left(\hat{\lambda}_{appzwr}\right)$, where E_1 and E_2 are the expectations over the first and second stage samples respectively. Further we have

$$E_1E_2\left(\hat{\lambda}_{appzwr}\right) = E_1E_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{ia}}{p_i}\right] = E_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}E_2\left(\hat{\lambda}_{ia}\right)\right],$$

Now, $E_2\left(\hat{\lambda}_{appzwr}\right) = E_2\left[\left((1/m_i)\sum_{j=1}^{m_i} y_{ij} - B_i\right)\Big/D_i\right]$,

where $B_i = (1-U_i)P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}$ and

$$D_i = U_i + (1-U_i)\left\{P_{1i} + P_{3i}\frac{k_i}{(k_i-1)}(P_{1i}+P_{3i})\right\}.$$

We have

$$E_2\left(\hat{\lambda}_{appzwr}\right)=\frac{1}{D_i}\left[\frac{1}{m_i}\sum_{j=1}^{m_i}E_2\left(y_{ij}\right)-B_i\right]=\frac{1}{D_i}\left[\frac{1}{m_i}\sum_{j=1}^{m_i}\lambda_{i0}-B_i\right],$$

$$=\frac{1}{D_i}\left[\lambda_{i0}-B_i\right].$$

Thus finally, we have

$$E_1E_2\left(\hat{\lambda}_{appzwr}\right)=E_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\lambda_{ia}}{p_i}\right],$$

$$=\frac{1}{M_0}\sum_{i=1}^{n}p_i\ \frac{M_i\lambda_{ia}}{p_i},$$

$$=\frac{1}{M_0}\sum_{i=1}^{n}M_i\lambda_{ia}=\lambda_a.$$

Therefore, $\hat{\lambda}_{appzwr}$ is an unbiased estimator of λ_a.

Theorem 3.2 The variance of the unbiased estimator $\hat{\lambda}_{appzwr}$ is

$$V\left(\hat{\lambda}_{appzwr}\right)=\frac{1}{nM_0^2}\left[\sum_{i=1}^{N}p_i\left(\frac{M_i\lambda_{ia}}{p_i}-M_0\lambda_a\right)^2+\sum_{i=1}^{N}\frac{M_i^2\Phi_i}{p_im_i}\right], \tag{5}$$

where

$$\Phi_i=\left[\frac{\lambda_{ia}}{D_i}+\frac{(1-U_i)P_{2i}\left\{1+P_{3i}\dfrac{k_i}{(k_i-1)}\right\}\lambda_{iy}}{D_i^2}\right]. \tag{6}$$

Proof- Let V_1 be the variance over the first-stage sample and V_2 be the variance over the second -stage sample. The variance of $\hat{\lambda}_{appzwr}$ is given by

$$V\left(\hat{\lambda}_{appzwr}\right)=V_1E_2\left(\hat{\lambda}_{appzwr}\right)+E_1V_2\left(\hat{\lambda}_{appzwr}\right). \tag{7}$$

Now,

$$V_1E_2\left(\hat{\lambda}_{appzwr}\right)=V_1E_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{ia}}{p_i}\right],$$

$$=V_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\lambda_{ia}}{p_i}\right],$$

$$=\frac{1}{nM_0^2}\sum_{i=1}^{N}p_i\left(\frac{M_i\lambda_{ia}}{p_i}-M_0\lambda_a\right)^2. \tag{8}$$

Next,

$$E_1V_2\left(\hat{\lambda}_{appzwr}\right)=E_1V_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{ia}}{p_i}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2\ V_2\left(\hat{\lambda}_{ia}\right)}{p_i^2}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2}V_2\left\{\frac{(1/m_i)\sum_{j=1}^{m_i}y_{ij}-B_i}{D_i}\right\}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2 D_i^2}.\frac{1}{m_i^2}\sum_{j=1}^{m_i}V_2\left(y_{ij}\right)\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2 D_i^2}.\frac{1}{m_i^2}\sum_{j=1}^{m_i}\lambda_{i0}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2 D_i^2}.\frac{\lambda_{i0}}{m_i}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2 D_i^2 m_i}.\{U_i\lambda_{ia}\right.$$

$$\left.\left.+(1-U_i)\left[\left(P_{1i}\lambda_{ia}+P_{2i}\lambda_{iy}\right)\left(1+P_{3i}\frac{k_i}{(k_i-1)}\right)+P_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right]\right\}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2 m_i}.\left\{\frac{\lambda_{ia}}{D_i}+\frac{(1-U_i)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}}{D_i^2}\right\}\right],$$

$$= \frac{1}{nM_0^2}\sum_{i=1}^{N}\frac{M_i^2}{p_i m_i}.\left\{\frac{\lambda_{ia}}{D_i}+\frac{(1-U_i)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}}{D_i^2}\right\},$$

$$= \frac{1}{nM_0^2}\sum_{i=1}^{N}\frac{M_i^2\Phi_i}{p_i m_i}\ . \tag{9}$$

Inserting (8) and (9) in (7) we get the variance of the unbiased estimator $\hat{\lambda}_{appzwr}$. This completes the proof of the theorem.

Theorem 3.3The unbiased estimate of the variance of the suggested estimator $\hat{\lambda}_{appzwr}$ is given by

$$\hat{V}\left(\hat{\lambda}_{appzwr}\right)=\frac{1}{n(n-1)M_0^2}\sum_{i=1}^{n}\left(\frac{M_i\hat{\lambda}_{ia}}{p_i}-\hat{\lambda}_{appzwr}\right)^2 \tag{10}$$

Proof is simple so omitted.

Estimation when the first-stage sample is selected with PPS sampling

The size of the cluster is known and when n clusters are selected with replacement depending on size of each cluster, M_i, the probability p_i should be considered as $\frac{M_i}{M_0}\left(i.e.\ p_i=\frac{M_i}{M_0}\right)$for the i[th]

cluster. Then it is known as PPS. From the PPS, the unbiased estimator for λ_a is given by

$$\hat{\lambda}_{appswr}=\frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_{ia}\ , \tag{11}$$

using $p_i=\frac{M_i}{M_0}$, and the variance is given by

$$V\left(\hat{\lambda}_{appswr}\right)=\frac{1}{nM_0}\left[\sum_{i=1}^{N}M_i\ \left(\lambda_{ia}-\lambda_a\right)^2+\sum_{i=1}^{N}\frac{M_i\Phi_i}{m_i}\right]. \tag{12}$$

Further the estimator for the variance of $\hat{\lambda}_{appswr}$ is given by

$$\hat{V}\left(\hat{\lambda}_{appswr}\right)=\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\hat{\lambda}_{ia}-\frac{\hat{\lambda}_{appswr}}{M_0}\right)^2. \tag{13}$$

Estimation when the first-stage sample is selected with probability proportional to size without replacement (PPSWOR)

When n clusters are selected without replacement from N clusters with size $M_i(i=1,2,...,N)$ each and δ_i is an inclusion probability of a unit i in a sample set without replacement, then the estimator for λ_a, which is the parameter of the rare sensitive attribute, is given by

$$\hat{\lambda}_{appswor}=\frac{1}{M_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{ia}}{\delta_i}. \tag{14}$$

The variance of estimator $\hat{\lambda}_{appswor}$ is given by

$$V\left(\hat{\lambda}_{appswor}\right)=\frac{1}{M_0^2}\left[\sum_{i=1}^{N}\sum_{j>1}^{N}\left(\delta_i\delta_j-\delta_{ij}\right)\left(\frac{M_i\lambda_{ia}}{\delta_i}-\frac{M_j\lambda_{ja}}{\delta_j}\right)^2+\sum_{i=1}^{N}\frac{M_i^2\Phi_i}{\delta_i m_i}\right], \tag{15}$$

where δ_{ij} is an inclusion probability of units i and j in a sample set without replacement.

The estimator for the variance of $\hat{\lambda}_{appswor}$ is given by

$$\hat{V}\left(\hat{\lambda}_{appswor}\right)=\frac{1}{M_0^2}\left[\sum_{i=1}^{n}\sum_{j>1}^{n}\frac{\left(\delta_i\delta_j-\delta_{ij}\right)}{\delta_{ij}}\left(\frac{M_i\hat{\lambda}_{ia}}{\delta_i}-\frac{M_j\hat{\lambda}_{ja}}{\delta_j}\right)^2+\sum_{i=1}^{n}\frac{M_i^2}{\delta_i}\cdot\frac{\hat{\Phi}_i}{\left(m_i-1\right)}\right], \tag{16}$$

where $\hat{\Phi}_i=\left[\dfrac{\hat{\lambda}_{ia}}{D_i}+\dfrac{\left(1-U_i\right)P_{2i}\left\{1+P_{3i}\dfrac{k_i}{\left(k_i-1\right)}\right\}\lambda_{iy}}{D_i^2}\right].$

Estimation when the first-stage sample is selected using SRSWR scheme

We consider the estimation when the first-stage sample is selected using SRSWR scheme. In this case, the selection probability for all the selected clusters in the first stage is $p_i=\frac{1}{N}\left(i=1,2,...,n\right)$. The estimator of the parameter λ_a; when the first–stage sample units are chosen with equal probability and with replacement in two-stage sampling, is

$$\hat{\lambda}_{awr}=\frac{N}{nM_0}\sum_{i=1}^{n}M_i\hat{\lambda}_{ia}\ . \tag{17}$$

The variance of the estimator $\hat{\lambda}_{awr}$ is

$$V\left(\hat{\lambda}_{awr}\right)=\frac{N}{nM_0^2}\left[\frac{N}{N-1}\sum_{i=1}^{N}\left(M_i\lambda_{ia}-\overline{M}\lambda_a\right)^2+\sum_{i=1}^{N}\frac{M_i^2}{m_i}\Phi_i\right], \tag{18}$$

and its estimate

$$\hat{V}\left(\hat{\lambda}_{awr}\right)=\frac{1}{n\left(n-1\right)M_0^2}\sum_{i=1}^{n}\left(NM_i\hat{\lambda}_{ia}-\hat{\lambda}_{awr}\right)^2\ , \tag{19}$$

where $\overline{M}=\frac{M_0}{N}$.

Comparing Probability proportional to size with replacement (PPSWR) and the equal probability two-stage sampling

From (12) and (18) we have

$$V\left(\hat{\lambda}_{awr}\right)-V\left(\hat{\lambda}_{appswr}\right)=\frac{1}{nM_0}\left[\frac{N}{M_0}\cdot\frac{N}{\left(N-1\right)}\sum_{i=1}^{N}\left(M_i\lambda_{ia}-\overline{M}\lambda_a\right)^2+\frac{N}{M_0}\cdot\sum_{i=1}^{N}\frac{M_i^2}{m_i}\Phi_i\right.$$
$$\left.-\sum_{i=1}^{N}M_i\ \left(\lambda_{ia}-\lambda_a\right)^2-\sum_{i=1}^{N}\frac{M_i}{m_i}\Phi_i\right].$$

Under the assumption $\left(N-1\right)\cong N$, and $\overline{M}=\frac{M_0}{N}$, we have

$$V\left(\hat{\lambda}_{awr}\right)-V\left(\hat{\lambda}_{appswr}\right)=\frac{1}{nM_0\overline{M}}\left[\sum_{i=1}^{N}\left(M_i\lambda_{ia}-\overline{M}\lambda_a\right)^2-\overline{M}\sum_{i=1}^{N}M_i\left(\lambda_{ia}-\lambda_a\right)^2\right.$$

$$\left.+.\sum_{i=1}^{N}\frac{M_i^2}{m_i}\Phi_i-\overline{M}\sum_{i=1}^{N}\frac{M_i}{m_i}\Phi_i\right].$$

$$=\frac{1}{nM_0\overline{M}}\left[\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\lambda_{ia}^2+\overline{M}\sum_{i=1}^{N}\left(M_i-\overline{M}\right)\left(\lambda_{ia}^2-\lambda_a^2\right)\right.$$

$$\left.+.\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\frac{\Phi_i}{m_i}-\overline{M}\sum_{i=1}^{N}\left(M_i-\overline{M}\right)\frac{\Phi_i}{m_i}\right]$$

which will be zero when $M_i=\overline{M}$. It follows that when the size of the clusters are all the same i.e. $M_i=\overline{M}$, the probability $p_i\left(=\frac{M_i}{M_0}\right)$ in PPS should be the same as $\frac{1}{N}$. Thereby meaning is that for $M_i=\overline{M}$, the estimator $\hat{\lambda}_{appswr}$ is at par with $\hat{\lambda}_{awr}$ i.e. both the estimators $\hat{\lambda}_{appswr}$ and $\hat{\lambda}_{awr}$ are equally efficient. However, when the difference $\left(M_i-\overline{M}\right)$ is larger the first term $\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\lambda_{ia}^2$, and the third term, $\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\frac{\Phi_i}{m_i}$ are increasing in the above expression. Usually the PPS is more efficient than the equal probability two-stage sampling when clusters have different sizes.

3.2 Estimation procedure of a rare attribute under two-stage sampling when the proportion of rare non-sensitive unrelated attribute is not known

It is to be mentioned that here two parameters are unknown so responses are gathered twice from each individual using two randomization devices in each cluster. These randomization devices consist of the decks of k_i similar cards as described Section 3.1. Firstly, the respondents selected from ith cluster are urged to response “yes” or “no” using the following two-stage randomization devices.

The first randomization device is given as follows:

First-stage randomization device R_{11} consists of two statements:

(i) Do you have the rare sensitive attribute A?

(ii) Go to randomization device R_{21}

with corresponding probabilities U_{1i} and $\left(1-U_{1i}\right)$respectively.

The second-stage randomization device consists of three statements:

(i) Do you have the rare sensitive attribute A ?

(ii) Do you have the rare non-sensitive unrelated attribute Y ?

(iii) Draw one more card

with corresponding probabilities P_{1i}, P_{2i} and P_{3i} respectively such that $\sum_{j=1}^{3} P_{ji} = 1$. If statement (iii) appeared on card then respondent repeat the process without replacing the card. Again, if in second draw the statement (iii) appeared the respondent is asked to report his/her actual status about sensitive attribute *A*.

Next, the respondent is urged again to answer one of the same questions using second randomization device.

The second randomization device R_{12} consists of two statements:

(i) Do you have the rare sensitive attribute *A*?

(ii) Go to randomization device R_{22}

with corresponding probabilities U_{2i} and $(1-U_{2i})$ respectively.

The second-stage randomization device R_{22} consists of three statements:

(i) Do you have the rare sensitive attribute A ?

(ii) Do you have the rare non-sensitive unrelated attribute Y ?

(iii) Draw one more card

with corresponding probabilities Q_{1i}, Q_{2i} and Q_{3i} respectively such that $\sum_{j=1}^{3} Q_{ji} = 1$.

If statement (iii) appeared on card then respondent repeat the process without replacing the card. Again, if in second draw the statement (iii) appeared the respondent is asked to report his/her actual status about sensitive attribute *A*.

Based on responses gathered using two randomization devices; the probabilities that respondents in the i^th cluster answer "yes" are

$$\theta_{i1} = U_{1i}\pi_{ia} + (1-U_{1i})\left[(P_{1i}\pi_{ia} + P_{2i}\pi_{iy})\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2\frac{k_i}{(k_i-1)}\pi_{ia}\right] \quad (20)$$

and

$$\theta_{i2} = U_{2i}\pi_{ia} + (1-U_{2i})\left[(Q_{1i}\pi_{ia} + Q_{2i}\pi_{iy})\left(1 + Q_{3i}\frac{k_i}{(k_i-1)}\right) + Q_{3i}^2\frac{k_i}{(k_i-1)}\pi_{ia}\right]. \quad (21)$$

Equation (20) and (21) can be rewritten as by assuming that as $m_i \to \infty,\ \theta_{i1} \to 0,\ \theta_{i2} \to 0$, then $m_i\theta_{i1} = \lambda_{i1}, and\ m_i\theta_{i2} = \lambda_{i2}$ respectively:

$$\lambda_{i1} = U_{1i}\lambda_{ia} + (1-U_{1i})\left[(P_{1i}\lambda_{ia} + P_{2i}\lambda_{iy})\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right] \quad (22)$$

and

$$\lambda_{i2} = U_{2i}\lambda_{ia} + (1-U_{2i})\left[(Q_{1i}\lambda_{ia} + Q_{2i}\lambda_{iy})\left(1 + Q_{3i}\frac{k_i}{(k_i-1)}\right) + Q_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right]. \quad (23)$$

After deriving the results, the following two equations are obtained as

$$\frac{1}{m_i}\sum_{j=1}^{m_i} y_{i1j} = U_{1i}\hat{\lambda}_{iau} + (1-U_{1i})\left[\left(P_{1i}\hat{\lambda}_{iau} + P_{2i}\hat{\lambda}_{iyu}\right)\left(1 + P_{3i}\frac{k_i}{(k_i-1)}\right) + P_{3i}^2\frac{k_i}{(k_i-1)}\hat{\lambda}_{iau}\right] \tag{24}$$

and

$$\frac{1}{m_i}\sum_{j=1}^{m_i} y_{i2j} = U_{2i}\hat{\lambda}_{iau} + (1-U_{2i})\left[\left(Q_{1i}\hat{\lambda}_{iau} + Q_{2i}\hat{\lambda}_{iyu}\right)\left(1 + Q_{3i}\frac{k_i}{(k_i-1)}\right) + Q_{3i}^2\frac{k_i}{(k_i-1)}\hat{\lambda}_{iau}\right], \tag{25}$$

where y_{i1j} and y_{i2j} denote the observed values in the first and the second responses from the $j^{th}\ (j = 1, 2, ..., m_i)$ respondents in the $i^{th}\ (i = 1, 2, ..., n)$ clusters. Solving equations (24) and (25) the estimators for λ_{ia} and λ_{iy} are:

$$\hat{\lambda}_{iau} = \frac{1}{C_{1i}m_i}\left[(1-U_{2i})Q_{2i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\}\sum_{j=1}^{m_i} y_{i1j} - (1-U_{1i})P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\sum_{j=1}^{m_i} y_{i2j}\right] \tag{26}$$

$$\begin{aligned}\hat{\lambda}_{iyu} = \frac{1}{D_{1i}m_i}\Bigg[&\left\{U_{2i} + (1-U_{2i})\left[Q_{1i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} + Q_{3i}^2\frac{k_i}{(k_i-1)}\right]\right\}\sum_{j=1}^{m_i} y_{i1j} \\ &-\left\{U_{1i} + (1-U_{1i})\left[P_{1i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\} + P_{3i}^2\frac{k_i}{(k_i-1)}\right]\right\}\sum_{j=1}^{m_i} y_{i2j}\Bigg],\end{aligned} \tag{27}$$

where

$$\begin{aligned}C_{1i} = {} & U_{1i}(1-U_{2i})Q_{2i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} - U_{2i}(1-U_{1i})P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\} \\ & + (1-U_{1i})(1-U_{2i})(P_{1i}Q_{2i} - P_{2i}Q_{1i})\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} \\ & + (1-U_{1i})(1-U_{2i})\frac{k_i}{(k_i-1)}\left[P_{3i}^2Q_{2i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} - Q_{3i}^2P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\right] \neq 0 .\end{aligned} \tag{28}$$

$$\begin{aligned}D_{1i} = {} & U_{2i}(1-U_{1i})P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\} - U_{1i}(1-U_{2i})Q_{2i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} \\ & + (1-U_{1i})(1-U_{2i})(P_{2i}Q_{1i} - P_{1i}Q_{2i})\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\} \\ & + (1-U_{1i})(1-U_{2i})\frac{k_i}{(k_i-1)}\left[Q_{3i}^2P_{2i}\left\{1 + P_{3i}\frac{k_i}{(k_i-1)}\right\} - P_{3i}^2Q_{2i}\left\{1 + Q_{3i}\frac{k_i}{(k_i-1)}\right\}\right] \neq 0 .\end{aligned} \tag{29}$$

Thus the estimator for the parameter λ_a of the rare sensitive attribute is given by

$$\hat{\lambda}_{appzwru} = \frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{iau}}{p_i}. \tag{30}$$

Theorem 3.4The proposed estimator $\hat{\lambda}_{appzwru}$ for the mean number of persons who have the rare sensitive attribute is unbiased.

Proof: Since λ_{i1j} and λ_{i2j} are iid Poisson variates with parameters λ_{i1} and λ_{i2} respectively, therefore, we have

$$E\left(\hat{\lambda}_{iau}\right) = E_1E_2\left(\hat{\lambda}_{appzwru}\right),$$

$$= E_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}E_2\left(\hat{\lambda}_{iau}\right)\right]. \tag{31}$$

Now,

$$E_2\left(\hat{\lambda}_{iau}\right) = E_2\left[\frac{1}{C_{1i}m_i}\left\{d_{2i}\sum_{j=1}^{m_i}y_{i1j} - C_{2i}\sum_{j=1}^{m_i}y_{i2j}\right\}\right],$$

$$= \frac{1}{C_{1i}m_i}\left[d_{2i}\sum_{j=1}^{m_i}E_2\left(y_{i1j}\right) - C_{2i}\sum_{j=1}^{m_i}E_2\left(y_{i2j}\right)\right],$$

$$= \frac{1}{C_{1i}m_i}\left[d_{2i}\sum_{j=1}^{m_i}\lambda_{i1} - C_{2i}\sum_{j=1}^{m_i}\lambda_{i2}\right],$$

$$= \frac{1}{C_{1i}m_i}\left[m_id_{2i}\lambda_{i1} - m_iC_{2i}\lambda_{i2}\right],$$

$$= \frac{\left[d_{2i}\lambda_{i1} - C_{2i}\lambda_{i2}\right]}{C_{1i}}, \tag{32}$$

where $d_{2i} = (1-U_{2i})Q_{2i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\}$, $C_{2i} = (1-U_{1i})P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}$.

Inserting the values of λ_{i1} and λ_{i2} from equations (22) and (23) respectively in (32) we get

$$E_2\left(\hat{\lambda}_{iau}\right) = \lambda_{ia}. \tag{33}$$

Putting (3.33) in (3.31) we have

$$E\left(\hat{\lambda}_{iau}\right) = E_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}\lambda_{ia}\right],$$

$$= \frac{1}{M_0}\sum_{i=1}^{N}p_i\frac{M_i}{p_i}\lambda_{ia},$$

$$= \frac{1}{M_0}\sum_{i=1}^{N}M_i\lambda_{ia} = \lambda_a.$$

Theorem 3.5The variance of the estimator $\hat{\lambda}_{appzwru}$ is given by

$$V\left(\hat{\lambda}_{appzwru}\right)=\frac{1}{nM_0^2}\left[\sum_{i=1}^{N} p_i \left(\frac{M_i\lambda_{ia}}{p_i}-M_0\lambda_a\right)^2+\sum_{i=1}^{N}\frac{M_i^2}{p_i}\frac{\Phi_i^{(12)}}{C_{1i}^2 m_i}\right], \tag{34}$$

where

$$\Phi_i^{(12)}=\left[\left\{d_{2i}^2 C_{3i}+C_{2i}^2 d_{3i}-2C_{2i}C_{3i}d_{2i}d_{3i}\right\}\lambda_{ia}+\left\{d_{2i}^2 C_{2i}+C_{2i}^2 d_{2i}-2C_{2i}^2 d_{2i}^2\right\}\lambda_{iy}\right],$$

$$C_{2i}=\left(1-U_{1i}\right)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\},$$

$$C_{3i}=U_{1i}+\left(1-U_{1i}\right)\left[P_{1i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}+P_{3i}^2\frac{k_i}{(k_i-1)}\right],$$

$$d_{2i}=\left(1-U_{2i}\right)Q_{2i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\},$$

$$d_{3i}=U_{2i}+\left(1-U_{2i}\right)\left[Q_{1i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\}+Q_{3i}^2\frac{k_i}{(k_i-1)}\right].$$

Proof: The variance is decomposed by

$$V\left(\hat{\lambda}_{appzwru}\right)=V_1E_2\left(\hat{\lambda}_{appzwru}\right)+E_1V_2\left(\hat{\lambda}_{appzwru}\right). \tag{35}$$

Since $y_{i1j}\sim iid\ Poisson\left(\lambda_{i1}\right)$ and $y_{i2j}\sim iid\ Poisson\left(\lambda_{i2}\right)$, $V\left(y_{i1j}\right)=E\left(y_{i1j}\right)=\lambda_{i1}$ and $V\left(y_{i2j}\right)=E\left(y_{i2j}\right)=\lambda_{i2}$.

We have

$$V_1E_2\left(\hat{\lambda}_{appzwru}\right)=V_1E_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{iau}}{p_i}\right],$$

$$=V_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}E_2\left(\hat{\lambda}_{iau}\right)\right],$$

$$=V_1\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\lambda_{ia}}{p_i}\right],$$

$$=\frac{1}{nM_0^2}\sum_{i=1}^{N}p_i\left(\frac{M_i\lambda_{ia}}{p_i}-M_0\lambda_a\right)^2 \tag{36}$$

and

$$E_1V_2\left(\hat{\lambda}_{appzwru}\right)=E_1V_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i\hat{\lambda}_{iau}}{p_i}\right],$$

$$=E_1V_2\left[\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}\frac{1}{C_{1i}m_i}\left\{d_{2i}\sum_{j=1}^{m_i}y_{i1j}-C_{2i}\sum_{j=1}^{m_i}y_{i2j}\right\}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2}\left\{\frac{1}{C_{1i}^2 m_i^2}\left[d_{2i}^2\sum_{j=1}^{m_i}V_2(y_{i1j})+C_{2i}^2\sum_{j=1}^{m_i}V_2(y_{i2j}) - 2d_{2i}C_{2i}\sum_{j=1}^{m_i}Cov(y_{i1j},y_{i2j})\right]\right\}\right],$$

$$= E_1\left[\frac{1}{(nM_0)^2}\sum_{i=1}^{n}\frac{M_i^2}{p_i^2}\left\{\frac{1}{C_{1i}^2 m_i^2}\left[d_{2i}^2\sum_{j=1}^{m_i}\lambda_{1i}+C_{2i}^2\sum_{j=1}^{m_i}\lambda_{2i}-2d_{2i}C_{2i}\sum_{j=1}^{m_i}\lambda_{i12}\right]\right\}\right],$$

$$= \frac{1}{(nM_0)^2}\sum_{i=1}^{N}\frac{M_i^2}{p_i}\left[\frac{1}{C_{1i}^2 m_i}\left\{d_{2i}^2\sum_{j=1}^{m_i}\lambda_{1i}+C_{2i}^2\sum_{j=1}^{m_i}\lambda_{2i}-2d_{2i}C_{2i}\sum_{j=1}^{m_i}\lambda_{i12}\right\}\right], \tag{37}$$

where

$$\lambda_{i1} = V(y_{i1j}) = \left[U_{1i}\lambda_{ia} + (1-U_{1i})\left[(P_{1i}\lambda_{ia}+P_{2i}\lambda_{iy})\left(1+P_{3i}\frac{k_i}{(k_i-1)}\right)+P_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right]\right], \tag{38}$$

$$\lambda_{i2} = V(y_{i2j}) = \left[U_{2i}\lambda_{ia} + (1-U_{2i})\left[(Q_{1i}\lambda_{ia}+Q_{2i}\lambda_{iy})\left(1+Q_{3i}\frac{k_i}{(k_i-1)}\right)+Q_{3i}^2\frac{k_i}{(k_i-1)}\lambda_{ia}\right]\right], \tag{39}$$

and $\lambda_{i12} = Cov(y_{i1j}, y_{i2j}) = E(y_{i1j}, y_{i2j}) - E(y_{i1j})E(y_{i2j})$

$$= \left[\left\{U_{1i} + (1-U_{1i})\left[P_{1i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}+P_{3i}^2\frac{k_i}{(k_i-1)}\right]\right\} * \left\{U_{2i} + (1-U_{2i})\left[Q_{1i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\}+Q_{3i}^2\frac{k_i}{(k_i-1)}\right]\right\}\right]\lambda_{ia} + \left[(1-U_{1i})P_{2i}\left(1+P_{3i}\frac{k_i}{(k_i-1)}\right)\right]\left[(1-U_{2i})Q_{2i}\left(1+Q_{3i}\frac{k_i}{(k_i-1)}\right)\right]\lambda_{iy}\ . \tag{40}$$

Inserting the values from (36) and (37) in (35), we obtained the expression of the variance of the estimator $\hat{\lambda}_{appzwru}$ as given in (34).

The estimator of an unbiased estimator of the variance of $\hat{\lambda}_{appzwru}$ is given by

$$\hat{V}(\hat{\lambda}_{appzwru}) = \frac{1}{n(n-1)M_0^2}\sum_{i=1}^{n}\left(\frac{M_i\hat{\lambda}_{iau}}{p_i} - \hat{\lambda}_{appzwru}\right)^2 . \tag{41}$$

Estimation when the first-stage sample is selected using PPSWR.

In the PPS, the selecting probability p_i for cluster i is defined as $\frac{M_i}{M_0}\left(i.e.\ p_i = \frac{M_i}{M_0}\right)$.

The estimator for the parameter of the rare sensitive attribute; λ_a is defined as

$$\hat{\lambda}_{appswru} = \frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}_{iau} \ , \tag{42}$$

and its variance is obtained as

$$V\left(\hat{\lambda}_{appswru}\right) = \frac{1}{nM_0}\left[\sum_{i=1}^{N} M_i \ \left(\lambda_{ia} - \lambda_a\right)^2 + \sum_{i=1}^{N}\frac{M_i \Phi_i^{(12)}}{C_{1i}^2 m_i}\right]. \tag{43}$$

Further the estimator for the variance of $\hat{\lambda}_{appswru}$ is given by

$$\hat{V}\left(\hat{\lambda}_{appswru}\right) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\hat{\lambda}_{iau} - \frac{\hat{\lambda}_{appswru}}{M_0}\right)^2 \ . \tag{44}$$

3.3 Estimation when the first-stage sample is selected using SRSWR.

In this situation, the probability of selecting the clusters in the first stage is $p_i = \frac{1}{N}(i = 1, 2, ..., n)$. The estimator of the parameter λ_a; when the first–stage sample units are selected with equal probability and with replacement in two-stage sampling, is

$$\hat{\lambda}_{awru} = \frac{N}{nM_0}\sum_{i=1}^{n} M_i \hat{\lambda}_{iau} \ . \tag{45}$$

The variance of the estimator $\hat{\lambda}_{awru}$ is

$$V\left(\hat{\lambda}_{awru}\right) = \frac{N}{nM_0^2}\left[\frac{N}{N-1}\sum_{i=1}^{N}\left(M_i \lambda_{ia} - \overline{M}\lambda_a\right)^2 + \sum_{i=1}^{N} M_i^2 \ \frac{\Phi_i^{(12)}}{C_{1i}^2 m_i}\right]. \tag{46}$$

and its estimate

$$\hat{V}\left(\hat{\lambda}_{awru}\right) = \frac{1}{n(n-1)M_0^2}\sum_{i=1}^{n}\left(NM_i \hat{\lambda}_{iau} - \hat{\lambda}_{awru}\right)^2 , \tag{47}$$

where $\overline{M} = \frac{M_0}{N}$.

Estimation by PPSWOR

After*n* clusters are drawn without replacement (WOR) from *N* clusters with size M_i each, m_i samples are selected in cluster *i* randomly with replacement. Let δ_i is an inclusion probability of a unit *i* in a sample set without replacement, and δ_{ij} is an inclusion probability of units *i* and *j* in a sample set without replacement. When the parameter of the rare unrelated attribute is unknown, the estimator for the rare sensitive attribute is defined by

$$\hat{\lambda}_{appsworu} = \frac{1}{M_0}\sum_{i=1}^{n}\frac{M_i \hat{\lambda}_{iau}}{\delta_i} \ . \tag{48}$$

The variance of estimator $\hat{\lambda}_{appsworu}$ is given by

The variance of estimator $\hat{\lambda}_{appsworu}$ is given by

$$V\left(\hat{\lambda}_{appsworu}\right)=\frac{1}{M_0^2}\left[\sum_{i=1}^{N}\sum_{j>1}^{N}\left(\delta_i\delta_j-\delta_{ij}\right)\left(\frac{M_i\lambda_{ia}}{\delta_i}-\frac{M_j\lambda_{ja}}{\delta_j}\right)^2+\sum_{i=1}^{N}\frac{M_i^2}{\delta_i}\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}\right]. \tag{49}$$

The estimator for the variance of $\hat{\lambda}_{appsworu}$ is given by

$$\hat{V}\left(\hat{\lambda}_{appsworu}\right)=\frac{1}{M_0^2}\left[\sum_{i=1}^{n}\sum_{j>1}^{n}\frac{\left(\delta_i\delta_j-\delta_{ij}\right)}{\delta_{ij}}\left(\frac{M_i\hat{\lambda}_{iau}}{\delta_i}-\frac{M_j\hat{\lambda}_{jau}}{\delta_j}\right)^2+\sum_{i=1}^{n}\frac{M_i^2}{\delta_i}.\frac{\hat{\Phi}_i^{(12)}}{C_{1i}^2\left(m_i-1\right)}\right], \tag{50}$$

where $\hat{\Phi}_i^{(12)}=\left[\left\{d_{2i}^2C_{3i}+C_{2i}^2d_{3i}-2C_{2i}C_{3i}d_{2i}d_{3i}\right\}\hat{\lambda}_{iau}+\left\{d_{2i}^2C_{2i}+C_{2i}^2d_{2i}-2C_{2i}^2d_{2i}^2\right\}\hat{\lambda}_{iyu}\right].$ (51)

Comparing PPSWR and the equal probability two-stage sampling

Assuming $(N-1)\cong N$, From (46) and (49) we have

$$V\left(\hat{\lambda}_{awru}\right)-V\left(\hat{\lambda}_{appswru}\right)=\frac{N}{nM_0^2}\left[\sum_{i=1}^{N}\left(M_i\lambda_{ia}-\overline{M}\lambda_a\right)^2+\sum_{i=1}^{N}M_i^2\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}\right]$$

$$-\frac{1}{nM_0}\left[\sum_{i=1}^{N}M_i\left(\lambda_{ia}-\lambda_a\right)^2+\sum_{i=1}^{N}M_i\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}\right],$$

$$=\frac{1}{nM_0\overline{M}}\left[\sum_{i=1}^{N}\left(M_i\lambda_{ia}-\overline{M}\lambda_a\right)^2+.\sum_{i=1}^{N}M_i^2\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}-\overline{M}\sum_{i=1}^{N}M_i\left(\lambda_{ia}-\lambda_a\right)^2-\overline{M}\sum_{i=1}^{N}M_i\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}\right]$$

,

$$=\frac{1}{nM_0\overline{M}}\left[\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\lambda_{ia}^2+\overline{M}\sum_{i=1}^{N}\left(M_i-\overline{M}\right)\left(\lambda_{ia}^2-\lambda_a^2\right)\right.$$

$$\left.+\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}+\overline{M}\sum_{i=1}^{N}\left(M_i-\overline{M}\right)\frac{\Phi_i^{(12)}}{C_{1i}^2m_i}\right]. \tag{52}$$

From (52) it is observed that the variance $\hat{\lambda}_{awru}$ is same as the variance of $\hat{\lambda}_{appswru}$ when $M_i=\overline{M}=\frac{M_0}{N}$ that is, when all clusters have the same size. The PPS with $p_i=\frac{1}{N}$ has the same variance as the equal probability two-stage sampling even when the parameter of the rare unrelated attribute is not known. If the difference in the cluster sizes is large, the first term $\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\lambda_{ia}^2$, and the third term ,

$\sum_{i=1}^{N}\left(M_i-\overline{M}\right)^2\frac{\Phi_i^{(12)}}{C_{1i}^2 m_i}$ in (52) are increasing.Usually the estimation by PPS is better than the equal probability two-stage sampling when clusters have different sizes.

4. Estimation with PPS method of a rare sensitive attribute under stratified two-stage sampling using a randomized response model.

We consider a population which is supposed to be stratified into *L* strata such that the h^{th} stratum has N_h, $h=1,2,...,L$, cluster, which are the first-stage units. Each cluster has size M_{hi} $(i=1,2,...,N)$ in stratum *h*. In stratum h, n_h clusters (which are at the first units) are selected from N_h clusters with probability p_{hi}. At the second stage, we select m_{hi} $(i=1,2,...,n_h)$, second-stage units from the i^{th} clusterdrawn from the h^{th} stratum, using SRSWR scheme.

4.1 PPS and Equal Probability Two-Stage Sampling for Stratification (when the rare unrelated attribute is known).

In this section, we assume that n_h clusters are drawn from h^{th} stratum using probability proportional to size with replacement (PPSWR) sampling scheme and π_{hiy}, the proportion of the rare non-sensitive unrelated attribute in the h^{th} stratum is supposed to be known. When the randomized response model described in Section 3 is applied, the probability that respondents answer 'yes' in the i^{th} cluster of the stratum *h* is defined as

$$\theta_{hi0}=U_{hi}\pi_{hia}+\left(1-U_{hi}\right)\left[\left(P_{1hi}\pi_{hia}+P_{2hi}\pi_{hiy}\right)\left(1+P_{3hi}\frac{k_{hi}}{\left(k_{hi}-1\right)}\right)+P_{3hi}^2\frac{k_{hi}}{\left(k_{hi}-1\right)}\pi_{hia}\right], \tag{53}$$

where U_{hi} is the probability of the question being asked ;Question (i): 'Do you possess the rare sensitive attribute *A*? in the first stage randomization device of Section 3.1 and the symbols P_{1hi}, P_{2hi} and P_{3hi} are the probabilities of presenting the statements (i), (ii) and (iii) in the randomization device used in the i^{th} cluster of the h^{th} stratum $\left(P_{1hi}+P_{2hi}+P_{3hi}=1\right)$. π_{hia} and π_{hiy} are the population proportions of the rare sensitive attribute, *A*, and the rare unrelated attribute, *Y*, respectively in the i^{th} cluster of the h^{th} stratum. π_{hiy} is assumed to be known.In the i^{th} cluster of stratum *h* since the two attribute *A* and *Y* are rare, $m_{hi}\theta_{hi0}=\lambda_{hi0}>0$, $m_{hi}\pi_{hia}=\lambda_{hia}>0$ and $m_{hi}\pi_{hiy}=\lambda_{hiy}>0$ are finite for $m_{hi}\to\infty$ as $\theta_{hi0}\to 0$, $\pi_{hia}\to 0$ and $\pi_{hiy}\to 0$.

Let $y_{hi1},y_{hi2},...,y_{him_{hi}}$ be m_{hi} random samples from the Poisson distribution with mean λ_{hi0} in the cluster of stratum *h*. Then the estimator $\hat{\lambda}_{hia}$ of the mean total number of persons bearing the rare sensitive attribute in the i^{th} cluster of the h^{th} stratum is defined as

$$\hat{\lambda}_{hia} = \frac{\left[(1/m_{hi}) \sum_{j=1}^{m_{hi}} y_{hij} - (1-U_{hi}) P_{2hi} \left\{ 1 + P_{3hi} \frac{k_{hi}}{(k_{hi}-1)} \right\} \lambda_{hiy} \right]}{\left[U_{hi} + (1-U_{hi}) \left\{ P_{1hi} + P_{3hi} \frac{k_{hi}}{(k_{hi}-1)} (P_{1hi} + P_{3hi}) \right\} \right]}. \quad (54)$$

Now, an estimator for the rare sensitive attribute, λ_{ha} in stratum *h* is given by

$$\hat{\lambda}_{ha} = \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hia}}{p_{hi}} , \quad (55)$$

where $M_{h0} = \sum_{i-1}^{n_h} M_{hi}$.

Under the stratified two-stage sampling design, the final estimator $\hat{\lambda}_{asppzwr}$ for the rare sensitive attribute, λ_{ha}, is given by

$$\hat{\lambda}_{asppzwr} = \sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hia}}{p_{hi}}, \quad (56)$$

where $W_h = \frac{N_h}{N}$, $N = \sum_{h=1}^{L} N_h$ and p_{hi} is the initial probability of selecting the ith cluster, which
is a first-stage unit in the h^{th} stratum.

Properties of the estimator $\hat{\lambda}_{asppzwr}$

Theorem-4.1-The estimator $\hat{\lambda}_{asppzwr}$ is unbiased.

Proof: Since $y_{hij} \sim iid\ Poisson(\lambda_{hi0}), E(y_{hij}) = \lambda_{hi0}$,

$$\lambda_{hi0} = U_{hi}\lambda_{hia} + (1-U_{hi}) \left[(P_{1hi}\lambda_{hia} + P_{2hi}\lambda_{hiy}) \left(1 + P_{3hi} \frac{k_{hi}}{(k_{hi}-1)} \right) + P_{3hi}^2 \frac{k_{hi}}{(k_{hi}-1)} \lambda_{hia} \right], \quad (57)$$

and the estimator $\hat{\lambda}_{hia}$ is an unbiased estimator of $\lambda_{hia} (i = 1, 2, ..., n_h;\ j = 1, 2, ..., m_{hi})$. Therefore,

$$E(\hat{\lambda}_{asppzwr}) = E_1 E_2 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hia}}{p_{hi}} \right],$$

$$= E_1 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \lambda_{hia}}{p_{hi}} \right],$$

$$= \sum_{h=1}^{L} W_h \frac{1}{M_{h0}} \sum_{i=1}^{N_h} p_{hi} \frac{M_{hi}\lambda_{hia}}{p_{hi}} ,$$
$$= \sum_{h=1}^{L} W_h \frac{1}{M_{h0}} \sum_{i=1}^{N_h} M_{hi}\lambda_{hia} ,$$
$$= \sum_{h=1}^{L} W_h \lambda_{ha} = \lambda_a .$$

This completes the proof of the theorem.

Theorem 4.2 The variance of the estimator $\hat{\lambda}_{asppzwr}$ is given by

$$V\left(\hat{\lambda}_{asppzwr}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}^2} \left[\sum_{i=1}^{N_h} p_{hi} \left(\frac{M_{hi}\lambda_{hia}}{p_{hi}} - M_{h0}\lambda_{ha} \right)^2 + \sum_{i=1}^{N_h} \frac{M_{hi}^2}{p_{hi} m_{hi}} \Phi_{hi} \right], \tag{58}$$

where $\Phi_{hi} = \left[\frac{\lambda_{hia}}{D_{hi}} + \frac{(1-U_{hi})P_{2hi}\left\{1 + P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\lambda_{hiy}}{D_{hi}^2} \right]$

and $D_{hi} = \left[U_{hi} + (1-U_{hi})\left\{ P_{1hi} + P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}(P_{1hi} + P_{3hi}) \right\} \right]$.

Proof: The variance of $\hat{\lambda}_{asppzwr}$ is decomposed by

$$V\left(\hat{\lambda}_{asppzwr}\right) = V_1 E_2\left(\hat{\lambda}_{asppzwr}\right) + E_1 V_2\left(\hat{\lambda}_{asppzwr}\right) . \tag{59}$$

We have

$$V_1 E_2\left(\hat{\lambda}_{asppzwr}\right) = V_1 E_2 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\hat{\lambda}_{hia}}{p_{hi}} \right],$$
$$= V_1 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\lambda_{hia}}{p_{hi}} \right],$$
$$= \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}^2} \sum_{i=1}^{N_h} p_{hi} \left(\frac{M_{hi}\lambda_{hia}}{p_{hi}} - M_{h0}\lambda_{ha} \right)^2 . \tag{60}$$

As $y_{hij} \sim iid\ Poisson(\lambda_{hi0}), V(y_{hij}) = \lambda_{hi0}$, the second term is

$$E_1 V_2\left(\hat{\lambda}_{asppzwr}\right) = E_1 V_2 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\hat{\lambda}_{hia}}{p_{hi}} \right],$$
$$= E_1 \left[\sum_{h=1}^{L} W_h^2 \frac{1}{(n_h M_{h0})^2} \sum_{i=1}^{n_h} \frac{M_{hi}^2}{p_{hi}^2} V_2\left(\hat{\lambda}_{hia}\right) \right],$$
$$= E_1 \left[\sum_{h=1}^{L} W_h^2 \frac{1}{(n_h M_{h0})^2} \sum_{i=1}^{n_h} \frac{M_{hi}^2}{p_{hi}^2} V_2 \left\{ \frac{1}{D_{hi}} \left[(1/m_{hi}) \sum_{j=1}^{m_{hi}} y_{hij} - B_{hi} \right] \right\} \right],$$

$$= E_1\left[\sum_{h=1}^{L} W_h^2 \frac{1}{(n_h M_{h0})^2}\sum_{i=1}^{n_h}\frac{M_{hi}^2}{p_{hi}^2 D_{hi}^2}.\left\{\frac{1}{m_{hi}^2}\sum_{j=1}^{m_{hi}} V_2(y_{hij})\right\}\right],$$
$$= E_1\left[\sum_{h=1}^{L} W_h^2 \frac{1}{(n_h M_{h0})^2}\sum_{i=1}^{n_h}\frac{M_{hi}^2}{p_{hi}^2 D_{hi}^2}.\frac{1}{m_{hi}^2}\sum_{j=1}^{m_{hi}} \lambda_{hi0}\right],$$
$$= E_1\left[\sum_{h=1}^{L} W_h^2 \frac{1}{(n_h M_{h0})^2}\sum_{i=1}^{n_h}\frac{M_{hi}^2}{p_{hi}^2}.\frac{1}{m_{hi} D_{hi}^2}\lambda_{hi0}\right],$$
$$= \left[\sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}^2}\sum_{i=1}^{N_h}\frac{M_{hi}^2}{p_{hi}}.\frac{\Phi_{hi}}{m_{hi}}\right], \because \frac{\lambda_{hi0}}{D_{hi}^2} = \Phi_{hi}; \tag{61}$$
where $B_{hi} = (1-U_{hi})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\lambda_{hiy}$.

Adding (60) and (61), we get the variance of the unbiased estimator $\hat{\lambda}_{asppzwr}$ as given in (58). Thusthe theorem is proved.

The unbiased estimate of the variance of $\hat{\lambda}_{asppzwr}$ is given by

$$\hat{V}(\hat{\lambda}_{asppzwr}) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h(n_h-1)M_{h0}^2}\sum_{i=1}^{n_h}\left(\frac{M_{hi}\hat{\lambda}_{hia}}{p_{hi}} - \hat{\lambda}_{ha}\right)^2. \tag{62}$$

Estimation when the first-stage sample is selected using PPSWR

When n_h clusters are selected with replacement (WR) from N_h clusters depending on the cluster size, M_{hi} in stratum *h*, with selection probabilities $p_{hi}(i = 1, 2, ..., n_h)$, where $p_{hi} = \frac{M_{hi}}{M_{h0}}$. It is known as PPSWR. Hence the unbiased estimator of λ_a is given by

$$\hat{\lambda}_{asppswr} = \sum_{h=1}^{L} W_h \frac{1}{n_h}\sum_{i=1}^{n_h}\hat{\lambda}_{hia}, \tag{63}$$

and its variance is obtained as

$$V(\hat{\lambda}_{asppswr}) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}}\left[\sum_{i=1}^{N_h} M_{hi}\,(\lambda_{hia} - \lambda_{ha})^2 + \sum_{i=1}^{N_h}\frac{M_{hi}\Phi_{hi}}{m_{hi}}\right]. \tag{64}$$

An estimator for the variance of the estimator $\hat{\lambda}_{asppswr}$, is

$$\hat{V}(\hat{\lambda}_{asppswr}) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h(n_h-1)}\sum_{i=1}^{n_h}\left(\hat{\lambda}_{hia} - \frac{\hat{\lambda}_{ha}}{M_{h0}}\right)^2. \tag{65}$$

Estimation when the first-stage sample is selected using SRSWR scheme.

When the first–stage sample units are drawn with equal probability and with replacement in two-stage sampling, the selection probabilities for all the selected clusters at first-stage from the h^{th} stratum is $p_{hi} = \frac{1}{N_h}$ and the estimator of λ_a is

$$\hat{\lambda}_{assrswr} = \sum_{h=1}^{L} W_h \frac{N_h}{n_h M_{h0}} \sum_{i=1}^{n_h} M_{hi} \hat{\lambda}_{hia} ,$$

$$= \sum_{h=1}^{L} W_h \hat{\lambda}_{ha} , \tag{66}$$

where $\hat{\lambda}_{ha} = \frac{N_h}{n_h M_{h0}} \sum_{i=1}^{n_h} M_{hi} \hat{\lambda}_{hia}$.

The variance of the estimator $\hat{\lambda}_{assrswr}$ is

$$V\left(\hat{\lambda}_{assrswr}\right) = \sum_{h=1}^{L} W_h^2 \frac{N_h}{n_h M_{h0}^2} \left[\frac{N_h}{(N_h - 1)} \sum_{i=1}^{N_h} \left(M_{hi}\lambda_{hia} - \overline{M}_h \lambda_{ha} \right)^2 + \sum_{i=1}^{N_h} \frac{M_{hi}^2}{m_{hi}} \Phi_{hi} \right] \tag{67}$$

and its estimate is given by

$$\hat{V}\left(\hat{\lambda}_{assrswr}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h (n_h - 1) M_{h0}^2} \sum_{i=1}^{n_h} \left(N_h M_{hi} \hat{\lambda}_{hia} - \hat{\lambda}_{ha} \right)^2 , \tag{68}$$

where $\overline{M}_h = \frac{M_{h0}}{N_h}$.

Estimation by PPSWOR for stratification

Let δ_{hi} be the probability that the i^{th} unit belongs to the first-stage sample and δ_{hij} the probability that both the i^{th} and j^{th} units belong to this sample, using PPSWOR from the h^{th} stratum. The estimator for the rare sensitive attribute in stratum h is given by

$$\hat{\lambda}_{happswor} = \frac{1}{M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\hat{\lambda}_{hia}}{\delta_{hi}} . \tag{69}$$

The unbiased estimator for the rare sensitive attribute (i.e. of λ_a) is given by

$$\hat{\lambda}_{asppswor} = \sum_{h=1}^{L} W_h \frac{1}{M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\hat{\lambda}_{hia}}{\delta_{hi}} . \tag{70}$$

The variance of the estimator $\hat{\lambda}_{asppswor}$ is given by

$$V\left(\hat{\lambda}_{asppswor}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{M_{h0}^2} \left[\sum_{i=1}^{N_h} \sum_{j>1}^{N_h} \left(\delta_{hi}\delta_{hj} - \delta_{hij} \right) \left(\frac{M_{hi}\lambda_{hia}}{\delta_{hi}} - \frac{M_{hj}\lambda_{hja}}{\delta_{hj}} \right)^2 + \sum_{i=1}^{N_h} \frac{M_{hi}^2}{m_{hi}\delta_{hi}} \Phi_{hi} \right]. \tag{71}$$

The estimator for the variance of $\hat{\lambda}_{asppswor}$ is given by

$$\hat{V}\left(\hat{\lambda}_{asppswor}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{M_{h0}^2}\left[\sum_{i=1}^{n_h}\sum_{j>1}^{n_h}\frac{\left(\delta_{hi}\delta_{hj}-\delta_{hij}\right)}{\delta_{hij}}\left(\frac{M_{hi}\hat{\lambda}_{hia}}{\delta_{hi}}-\frac{M_{hj}\hat{\lambda}_{hja}}{\delta_{hj}}\right)^2+\sum_{i=1}^{n_h}\frac{M_{hi}^2}{p_{hi}^*}.\frac{\hat{\Phi}_{hi}}{\left(m_{hi}-1\right)}\right], \tag{72}$$

where $\hat{\Phi}_{hi}=\left[\dfrac{\hat{\lambda}_{hia}}{D_{hi}}+\dfrac{\left(1-U_{hi}\right)P_{2hi}\left\{1+P_{3hi}\dfrac{k_{hi}}{\left(k_{hi}-1\right)}\right\}\lambda_{hiy}}{D_{hi}^2}\right]$

Comparing PPSWR and the equal probability two-stage sampling for stratification. From (64) and (67) we have

$$V\left(\hat{\lambda}_{assrswr}\right)-V\left(\hat{\lambda}_{asppswr}\right)=\sum_{h=1}^{L}W_h^2\left[\frac{N_h}{n_hM_{h0}^2}\left\{\frac{N_h}{\left(N_h-1\right)}\sum_{i=1}^{N_h}\left(M_{hi}\lambda_{hia}-\overline{M}_h\lambda_{ha}\right)^2+\sum_{i=1}^{N_h}\frac{M_{hi}^2}{m_{hi}}\Phi_{hi}\right\}-\frac{1}{n_hM_{h0}}\left\{\sum_{i=1}^{N_h}M_{hi}\left(\lambda_{hia}-\lambda_{ha}\right)^2+\sum_{i=1}^{N_h}\frac{M_{hi}}{m_i}\Phi_{hi}\right\}\right]. \tag{73}$$

For $\left(N_h-1\right)\cong N_h$, (73) boils down to

$$V\left(\hat{\lambda}_{assrswr}\right)-V\left(\hat{\lambda}_{asppswr}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_hM_{h0}\overline{M}_h}\left[\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)^2\lambda_{hia}^2+\overline{M}_h\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)\left(\lambda_{hia}^2-\lambda_{ha}^2\right)+\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)^2\frac{\Phi_{hi}}{m_{hi}}+\overline{M}_h\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)\frac{\Phi_{hi}}{m_{hi}}\right]. \tag{74}$$

It is observed from (74) that

$V\left(\hat{\lambda}_{assrswr}\right)=V\left(\hat{\lambda}_{asppswr}\right)$

when each cluster has the same size, $M_{hi}=\overline{M}_h=\dfrac{M_{h0}}{N_h}$ in stratum *h*, let $p_{hi}=\dfrac{1}{N_h}$ in the PPS. Also in (74) the first term $\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)^2\lambda_{hia}^2$, and the third term, $\sum_{i=1}^{N_h}\left(M_{hi}-\overline{M}_h\right)^2\dfrac{\Phi_{hi}}{m_{hi}}$ are increasing in case the clusters have different sizes. Usually the PPS is better than the equal probability two-stage sampling in stratified two-stage sampling.

4.2 Estimation method of rare-attribute when the rare non-sensitive unrelated attribute is unknown in stratified population.

Estimation by PPSWR for Stratification.
In this section, it is supposed that n_h clusters are selected from h[th] stratum (h=1, 2,…,L) using PPSWR sampling design and π_{hiy} is supposed to be unknown. The randomization device in the i[th] cluster of the h[th] stratum is the same as in Section 3.2. The probabilities that respondents in the i[th] cluster of the h[th] stratum answer "yes" are

$$\theta_{hi1} = U_{h1i}\pi_{hia} + (1-U_{h1i})\left[\left(P_{1hi}\pi_{hia} + P_{2hi}\pi_{hiy}\right)\left(1 + P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + P_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\pi_{hia}\right] \quad (75)$$

and

$$\theta_{hi2} = U_{h2i}\pi_{hia} + (1-U_{h2i})\left[\left(Q_{1hi}\pi_{hia} + Q_{2hi}\pi_{hiy}\right)\left(1 + Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + Q_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\pi_{hia}\right], \quad (76)$$

where U_{h1i} is the probability of the question being asked as described in (i) of Section 3.1 in first-stage randomization device and $(P_{1hi}, P_{2hi}, P_{3hi})$ are the probabilities of presenting the statements (i), (ii), and (iii) in the second stage randomization device, when this device is used in the i[th] cluster of the h[th] stratum, such that $P_{1hi} + P_{2hi} + P_{3hi} = 1$, π_{hia} is the true proportion of rare sensitive attribute A_h and π_{hiy} is the true proportion of rare non-sensitive unrelated attribute Y_h in the i[th] cluster of the h[th] stratum. Further U_{h2i} is the probability of the question being asked as described in (i) of Section 3.1 in first-stage randomization device (of second randomization device) and $(Q_{1hi}, Q_{2hi}, Q_{3hi})$ are the probabilities of presenting the statements (i), (ii), and (iii) in the second-stage randomization device (of second randomization device), when this device is used in the i[th] cluster of the h[th] stratum, such that $Q_{1hi} + Q_{2hi} + Q_{3hi} = 1$,and π_{hia} and π_{hiy} are the population proportions of the rare sensitive and unrelated attributes, respectively. Now π_{hiy} is assumed to be unknown. Since the two attribute are in the population, therefore, $m_{hi}\theta_{hi1} = \lambda_{hi1}$, *and* $m_{hi}\theta_{hi2} = \lambda_{hi2}$ are finite as $m_{ih} \to \infty$, $\theta_{hi1} \to 0$, $\theta_{hi2} \to 0$,where

$$\lambda_{hi1} = U_{h1i}\lambda_{hia} + (1-U_{h1i})\left[\left(P_{1hi}\lambda_{hia} + P_{2hi}\lambda_{hiy}\right)\left(1 + P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + P_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\lambda_{hia}\right] \quad (77)$$

and

$$\lambda_{hi2} = U_{h2i}\lambda_{hia} + (1-U_{h2i})\left[\left(Q_{1hi}\lambda_{hia} + Q_{2hi}\lambda_{hiy}\right)\left(1 + Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + Q_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\lambda_{hia}\right]. \quad (78)$$

Similar to the previous section, the following equations are obtained:

$$\frac{1}{m_{hi}}\sum_{j=1}^{m_{hi}} y_{hi1j} = U_{h1i}\hat{\lambda}_{hia} + (1-U_{h1i})\left[\left(P_{1hi}\hat{\lambda}_{hia} + P_{2hi}\hat{\lambda}_{hiy}\right)\left(1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + P_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\hat{\lambda}_{hia}\right] \tag{79}$$

and

$$\frac{1}{m_{hi}}\sum_{j=1}^{m_{hi}} y_{hi2j} = U_{h2i}\hat{\lambda}_{hia} + (1-U_{h2i})\left[\left(Q_{1hi}\hat{\lambda}_{hia} + Q_{2hi}\hat{\lambda}_{hiy}\right)\left(1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right) + Q_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\hat{\lambda}_{hia}\right], \tag{80}$$

where y_{hi1j} and y_{hi2j} are the first and the second answers of the j^{th} $(j = 1, 2, ..., m_{hi})$ respondents in the i^{th} cluster of the h^{th} stratum.

From (79) and (80), the estimators for the mean number of persons who possess the rare sensitive and unrelated attribute in the i^{th} cluster of stratum h are given by

$$\begin{aligned}\hat{\lambda}_{hiau} = \frac{1}{C_{h1i}m_{hi}}\Bigg[&(1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\sum_{j=1}^{m_{hi}} y_{hi1j} \\ &-(1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\sum_{j=1}^{m_i} y_{hi2j}\Bigg],\end{aligned} \tag{81}$$

$$\begin{aligned}\hat{\lambda}_{hiyu} = \frac{1}{D_{h1i}m_{hi}}\Bigg[&\left\{U_{h2i} + (1-U_{h2i})\left[Q_{1hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} + Q_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\right]\right\}\sum_{j=1}^{m_{hi}} y_{hi1j} \\ &-\left\{U_{h1i} + (1-U_{h1i})\left[P_{1hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} + P_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\right]\right\}\sum_{j=1}^{m_{hi}} y_{hi2j}\Bigg],\end{aligned} \tag{82}$$

where

$$\begin{aligned}C_{h1i} =\ & U_{h1i}(1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} - U_{h2i}(1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} \\ &+(1-U_{h1i})(1-U_{h2i})(P_{1hi}Q_{2hi} - P_{2hi}Q_{1hi})\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} \\ &+(1-U_{h1i})(1-U_{h2i})\frac{k_{hi}}{(k_{hi}-1)}\left[P_{3hi}^2Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} - Q_{3hi}^2P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\right] \neq 0\end{aligned} \tag{83}$$

$$\begin{aligned}D_{h1i} =\ & U_{h2i}(1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} - U_{h1i}(1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} \\ &+(1-U_{h1i})(1-U_{h2i})(P_{2hi}Q_{1hi} - P_{1hi}Q_{2hi})\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} \\ &+(1-U_{h1i})(1-U_{h2i})\frac{k_{hi}}{(k_{hi}-1)}\left[Q_{3hi}^2P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} - P_{3hi}^2Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\right] \neq 0.\end{aligned} \tag{84}$$

The estimator of the mean total number λ_{ha} of persons having rare sensitive attribute in h[th] stratum is

$$\hat{\lambda}_{hasppzwru} = \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hiau}}{p_{hi}}, \qquad (85)$$

where p_{hi} is the selection probability of the i^{th} cluster in h[th] stratum under PPSWR scheme and $M_{h0} = \sum_{i=1}^{n_h} M_{hi}$. Thus the final estimator $\hat{\lambda}_{asppzwru}$ of the mean total number λ_a of persons having a rare sensitive attribute in the population under stratified two-stage sampling scheme is

$$\hat{\lambda}_{asppzwru} = \sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hiau}}{p_{hi}}, \qquad (86)$$

where $W_h = \frac{N_h}{N}$ and $N = \sum_{h=1}^{L} N_h$.

Properties of the estimator $\hat{\lambda}_{asppzwru}$

Theorem 4.3 The estimator $\hat{\lambda}_{asppzwru}$ of the mean total number of persons is having the rare sensitive attribute.

Proof: Since λ_{hi1j} and λ_{hi2j} are iid Poisson variates with parameters λ_{hi1} and λ_{hi2} respectively, therefore, we have

$$E\left(\hat{\lambda}_{asppzwru}\right) = E_1 E_2 \left(\hat{\lambda}_{asppzwru}\right),$$

$$= E_1 E_2 \left[\sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi} \hat{\lambda}_{hiau}}{p_{hi}} \right],$$

Now,

$$E_2\left(\hat{\lambda}_{hiau}\right) = E_2 \left[\frac{1}{C_{h1i} m_{hi}} \left\{ d_{h2i} \sum_{j=1}^{m_{hi}} y_{hi1j} - C_{h2i} \sum_{j=1}^{m_{hi}} y_{hi2j} \right\} \right],$$

$$= \frac{1}{C_{h1i} m_{hi}} \left[d_{h2i} \sum_{j=1}^{m_{hi}} E_2\left(y_{hi1j}\right) - C_{h2i} \sum_{j=1}^{m_{hi}} E_2\left(y_{hi2j}\right) \right],$$

$$= \frac{1}{C_{h1i} m_{hi}} \left[d_{h2i} \sum_{j=1}^{m_{hi}} \lambda_{hi1} - C_{h2i} \sum_{j=1}^{m_{hi}} \lambda_{hi2} \right],$$

$$= \frac{1}{C_{h1i} m_{hi}} \left[m_{hi} d_{h2i} \lambda_{hi1} - m_{hi} C_{h2i} \lambda_{hi2} \right],$$

$$= \frac{\left[d_{h2i} \lambda_{hi1} - C_{h2i} \lambda_{hi2} \right]}{C_{h1i}}, \qquad (87)$$

where $d_{h2i} = (1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\dfrac{k_{hi}}{(k_{hi}-1)}\right\}$ and

$C_{h2i} = (1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\dfrac{k_{hi}}{(k_{hi}-1)}\right\}$.

Inserting the values of λ_{hi1} and λ_{hi2} from equations (77) and (78) respectively in (87) we get

$E_2\left(\hat{\lambda}_{hiau}\right) = \lambda_{hia}$.

Thus, $E_1E_2\left(\hat{\lambda}_{asppzwru}\right) = E_1E_2\left[\sum_{h=1}^{L}W_h\dfrac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\dfrac{M_{hi}\hat{\lambda}_{hiau}}{p_{hi}}\right]$,

$$= E_1\left[\sum_{h=1}^{L}W_h\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}}{p_{hi}}E_2\left(\hat{\lambda}_{hiau}\right)\right],$$

$$= E_1\left[\sum_{h=1}^{L}W_h\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}\lambda_{hia}}{p_{hi}}\right],$$

$$= \sum_{h=1}^{L}W_h\frac{1}{M_{h0}}\sum_{i=1}^{N_h}p_{hi}\frac{M_{hi}\lambda_{hia}}{p_{hi}},$$

$$= \sum_{h=1}^{L}W_h\left(\frac{1}{M_{h0}}\sum_{i=1}^{N_h}M_{hi}\lambda_{hia}\right),$$

$$= \sum_{h=1}^{L}W_h\lambda_{ha} = \lambda_a. \qquad (88)$$

Thus the theorem is proved.

Theorem 4.4 The variance of the estimator $\hat{\lambda}_{asppzwru}$ is given by

$$V\left(\hat{\lambda}_{asppzwru}\right) = \sum_{h=1}^{L}W_h^2\frac{1}{n_hM_{h0}^2}\left[\sum_{i=1}^{N_h}p_{hi}\left(\frac{M_{hi}\lambda_{hia}}{p_{hi}} - M_{h0}\lambda_{ha}\right)^2 + \sum_{i=1}^{N_h}\frac{M_{hi}^2}{p_{hi}}\frac{\Phi_{hi}^{(12)}}{C_{h1i}^2m_{hi}}\right], \qquad (89)$$

where

$\Phi_{hi}^{(12)} = \left[\left\{d_{h2i}^2C_{h3i} + C_{h2i}^2d_{h3i} - 2C_{h2i}C_{h3i}d_{h2i}d_{h3i}\right\}\lambda_{hia} + \left\{d_{h2i}^2C_{h2i} + C_{h2i}^2d_{h2i} - 2C_{h2i}^2d_{h2i}^2\right\}\lambda_{hiy}\right]$

,

$$C_{h3i} = U_{h1i} + (1-U_{h1i})\left[P_{1hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} + P_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\right],$$

$$d_{h3i} = U_{h2i} + (1-U_{h2i})\left[Q_{1hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} + Q_{3hi}^2\frac{k_{hi}}{(k_{hi}-1)}\right].$$

Proof: The variance of the estimator $\hat{\lambda}_{asppzwru}$ is decomposed as

$$V\left(\hat{\lambda}_{asppzwru}\right) = V_1E_2\left(\hat{\lambda}_{asppzwru}\right) + E_1V_2\left(\hat{\lambda}_{asppzwru}\right). \qquad (90)$$

The first term of (90) is simplified to

$$V_1E_2\left(\hat{\lambda}_{asppzwru}\right)=V_1E_2\left[\sum_{h=1}^{L}W_h\,\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}\hat{\lambda}_{hiau}}{p_{hi}}\right],$$

$$=V_1\left[\sum_{h=1}^{L}W_h\,\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}}{p_{hi}}E_2\left(\hat{\lambda}_{hiau}\right)\right],$$

$$=V_1\left[\sum_{h=1}^{L}W_h\,\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}\lambda_{hia}}{p_{hi}}\right],$$

$$=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_hM_{h0}^2}\sum_{i=1}^{N}p_{hi}\left(\frac{M_{hi}\lambda_{hia}}{p_{hi}}-M_{h0}\lambda_{ha}\right)^2. \tag{91}$$

The second term of (90) is simplified to

$$E_1V_2\left(\hat{\lambda}_{asppzwru}\right)=E_1V_2\left[\sum_{h=1}^{L}W_h\,\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}\hat{\lambda}_{hiau}}{p_{hi}}\right],$$

$$=E_1\left[\sum_{h=1}^{L}W_h^2\,\frac{1}{(n_hM_{h0})^2}\sum_{i=1}^{n_h}\frac{M_{hi}^2}{p_{hi}^2}V_2\left(\hat{\lambda}_{hiau}\right)\right]. \tag{92}$$

Now,

$$V_2\left(\hat{\lambda}_{hiau}\right)=V_2\left[\frac{1}{C_{h1i}m_{hi}}\left\{d_{h2i}\sum_{j=1}^{m_{hi}}y_{hi1j}-C_{h2i}\sum_{j=1}^{m_{hi}}y_{hi2j}\right\}\right]$$

$$=\frac{1}{C_{h1i}^2m_{hi}^2}\left[d_{h2i}^2\sum_{j=1}^{m_{hi}}V_2\left(y_{hi1j}\right)+C_{h2i}^2\sum_{j=1}^{m_{hi}}V_2\left(y_{hi2j}\right)-2d_{h2i}C_{h2i}\sum_{j=1}^{m_{hi}}Cov\left(y_{hi1j},y_{hi2j}\right)\right],$$

$$=\frac{1}{C_{h1i}^2m_{hi}^2}\left[d_{h2i}^2\sum_{j=1}^{m_{hi}}\lambda_{h1i}+C_{h2i}^2\sum_{j=1}^{m_{hi}}\lambda_{h2i}-2d_{h2i}C_{h2i}\sum_{j=1}^{m_{hi}}\lambda_{hi12}\right],$$

$$=\frac{1}{C_{h1i}^2m_{hi}}\left[d_{h2i}^2\lambda_{h1i}+C_{h2i}^2\lambda_{h2i}-2d_{h2i}C_{h2i}\lambda_{hi12}\right], \tag{93}$$

where $\lambda_{hi12}=Cov\left(y_{hi1j},y_{hi2j}\right)=E\left(y_{hi1j}y_{hi2j}\right)-E\left(y_{hi1j}\right)E\left(y_{hi2j}\right)$

$$=\left[\left\{U_{h1i}+\left(1-U_{h1i}\right)\left[P_{1hi}\left\{1+P_{3hi}\,\frac{k_{hi}}{(k_{hi}-1)}\right\}+P_{3hi}^2\,\frac{k_{hi}}{(k_{hi}-1)}\right]\right\}\right.$$

$$*\left\{U_{h2i}+\left(1-U_{h2i}\right)\left[Q_{1hi}\left\{1+Q_{3hi}\,\frac{k_{hi}}{(k_{hi}-1)}\right\}+Q_{3hi}^2\,\frac{k_{hi}}{(k_{hi}-1)}\right]\right\}\Bigg]\lambda_{hia}$$

$$+\left[\left(1-U_{h1i}\right)P_{2hi}\left\{1+P_{3hi}\,\frac{k_{hi}}{(k_{hi}-1)}\right\}\right]\left[\left(1-U_{h2i}\right)Q_{2hi}\left(1+Q_{3hi}\,\frac{k_{hi}}{(k_{hi}-1)}\right)\right]\lambda_{hiy}$$

and $\left(\lambda_{h1i},\lambda_{h2i}\right)$ are same as define earlier.

Putting the values from (83) in (91) and then substituting from (90) and (92) in (100), we obtained the expression of the variance of the estimator $\hat{\lambda}_{asppzwru}$ as given in (89).

The estimator of an unbiased estimator of the variance of $\hat{\lambda}_{asppzwru}$ is given by

$$\hat{V}\left(\hat{\lambda}_{asppzwru}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_h\left(n_h-1\right)M_{h0}^2}\sum_{i=1}^{n_h}\left(\frac{M_{hi}\hat{\lambda}_{hiau}}{p_{hi}}-\hat{\lambda}_{hau}\right)^2. \quad (94)$$

Estimation when the first-stage sample is selected using SRSWR scheme

Let the first–stage sample units are drawn with equal probability and with replacement in two-stage stratified sampling, the probabilities of selecting the i^{th} cluster from the h^{th} stratum is $p_{hi}=\frac{1}{N_h}$. Then the unbiased estimator for the parameter λ_a is defined by

$$\hat{\lambda}_{aswru}=\sum_{h=1}^{L}W_h\,\frac{N_h}{n_hM_{h0}}\sum_{i=1}^{n_h}M_{hi}\hat{\lambda}_{hiau}\ . \quad (95)$$

The variance of the estimator $\hat{\lambda}_{aswru}$ is

$$V\left(\hat{\lambda}_{aswru}\right)=\sum_{h=1}^{L}W_h^2\,\frac{N_h}{n_hM_{h0}^2}\left[\frac{N_h}{\left(N_h-1\right)}\sum_{i=1}^{N_h}\left(M_{hi}\lambda_{hia}-\overline{M}_h\lambda_{ha}\right)^2+\sum_{i=1}^{N_h}\frac{M_{hi}^2}{m_{hi}C_{h1i}^2}\Phi_{hi}^{(12)}\right] \quad (96)$$

and its estimate is given by

$$\hat{V}\left(\hat{\lambda}_{aswru}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_h\left(n_h-1\right)M_{h0}^2}\sum_{i=1}^{n_h}\left(N_hM_{hi}\hat{\lambda}_{hiau}-\hat{\lambda}_{hau}\right)^2. \quad (97)$$

Estimation when the first-stage sample is selected using PPS sampling.

When n_h clusters are drawn with replacement depending on cluster size, we consider the PPS. Let $p_{hi}=\frac{M_{hi}}{M_{h0}}$. The unbiased estimator for the mean number of persons who possess the rare sensitive attribute λ_a is

$$\hat{\lambda}_{asppswru}=\sum_{h=1}^{L}W_h\,\frac{1}{n_h}\sum_{i=1}^{n_h}\hat{\lambda}_{hiau}\ , \quad (98)$$

and its variance is obtained as

$$V\left(\hat{\lambda}_{asppswru}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_hM_{h0}}\left[\sum_{i=1}^{N_h}M_{hi}\ \left(\lambda_{hia}-\lambda_{ha}\right)^2+\sum_{i=1}^{N_h}\frac{M_{hi}\Phi_{hi}^{(12)}}{C_{h1i}^2m_{hi}}\right]. \quad (99)$$

Also, the estimator for the variance of $\hat{\lambda}_{asppswru}$ is given by

$$\hat{V}\left(\hat{\lambda}_{asppswru}\right)=\sum_{h=1}^{L}W_h^2\,\frac{1}{n_h\left(n_h-1\right)}\sum_{i=1}^{n_h}\left(\hat{\lambda}_{hiau}-\frac{\hat{\lambda}_{hau}}{M_{h0}}\right)^2. \quad (100)$$

Estimation by PPSWR for Stratification

Let δ_{hi} be the probability that the i^{th} unit belongs to the first-stage sample and δ_{hij} the probability that both the $i^{th}$and j^{th} units belong to first-stage sample, using probability proportional to size sampling without replacement from the h^{th} stratum. The unbiased estimator of λ_a is given by

$$\hat{\lambda}_{asppsworu} = \sum_{h=1}^{L} W_h \hat{\lambda}_{hau}, \tag{101}$$

where the estimator λ_{hau} in stratum *h* is defined by

$$\hat{\lambda}_{hau} = \frac{1}{M_{h0}} \sum_{i=1}^{n_h} \frac{M_{hi}\hat{\lambda}_{hiau}}{\delta_{hi}}. \tag{102}$$

The variance of estimator $\hat{\lambda}_{asppsworu}$ is given by

$$V\left(\hat{\lambda}_{asppswor}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{M_{h0}^2}\left[\sum_{i=1}^{N_h}\sum_{j>1}^{N_h}\left(\delta_{hi}\delta_{hj} - \delta_{hij}\right)\left(\frac{M_{hi}\lambda_{hia}}{\delta_{hi}} - \frac{M_{hj}\lambda_{hja}}{\delta_{hj}}\right)^2 + \sum_{i=1}^{N_h}\frac{M_{hi}^2\Phi_{hi}^{(12)}}{\delta_{hi}C_{h1i}^2 m_{hi}}\right]. \tag{103}$$

Further the estimator for the variance of $\hat{\lambda}_{asppsworu}$ is given by

$$V\left(\hat{\lambda}_{asppsworu}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{M_{h0}^2}\left[\sum_{i=1}^{n_h}\sum_{j>1}^{n_h}\frac{\left(\delta_{hi}\delta_{hj} - \delta_{hij}\right)}{\delta_{hij}}\left(\frac{M_{hi}\hat{\lambda}_{hiau}}{\delta_{hi}} - \frac{M_{hj}\hat{\lambda}_{hjau}}{\delta_{hj}}\right)^2 + \sum_{i=1}^{n_h}\frac{M_{hi}^2\hat{\Phi}_{hi}^{(12)}}{\left(m_{hi}-1\right)\delta_{hi}C_{h1i}^2}\right], \tag{104}$$

where

$$\hat{\Phi}_{hi}^{(12)} = \left[\left\{d_{h2i}^2 C_{h3i} + C_{h2i}^2 d_{h3i} - 2C_{h2i}C_{h3i}d_{h2i}d_{h3i}\right\}\hat{\lambda}_{hia} + \left\{d_{h2i}^2 C_{h2i} + C_{h2i}^2 d_{h2i} - 2C_{h2i}^2 d_{h2i}^2\right\}\hat{\lambda}_{hiy}\right].$$

Comparing PPSWR and the equal probability two-stage sampling for stratification
From (96) and (99) we have

$$V\left(\hat{\lambda}_{aswru}\right) - V\left(\hat{\lambda}_{asppswru}\right) = \sum_{h=1}^{L} W_h^2\left[\frac{N_h}{n_h M_{h0}^2}\left\{\frac{N_h}{\left(N_h - 1\right)}\sum_{i=1}^{N_h}\left(M_{hi}\lambda_{hia} - \overline{M}_h\lambda_{ha}\right)^2 + \sum_{i=1}^{N_h}\frac{M_{hi}^2\Phi_{hi}^{(12)}}{C_{h1i}^2 m_{hi}}\right\}\right.$$
$$\left. - \frac{1}{n_h M_{h0}}\left\{\sum_{i=1}^{N_h} M_{hi}\left(\lambda_{hia} - \lambda_{ha}\right)^2 + \sum_{i=1}^{N_h}\frac{M_{hi}\Phi_{hi}^{(12)}}{C_{h1i}^2 m_{hi}}\right\}\right], \tag{105}$$

Under the assumption $\left(N_h - 1\right) \cong N_h$, and $\overline{M}_h = \dfrac{M_{h0}}{N_h}$, expression (105) reduces to

$$V\left(\hat{\lambda}_{aswru}\right) - V\left(\hat{\lambda}_{asppswru}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}\overline{M}_h}\left[\sum_{i=1}^{N_h}\left(M_{hi} - \overline{M}_h\right)^2\lambda_{hia}^2 + \overline{M}_h\sum_{i=1}^{N_h}\left(M_{hi} - \overline{M}_h\right)\left(\lambda_{hia}^2 - \lambda_{ha}^2\right)\right.$$
$$\left. + \sum_{i=1}^{N_h}\left(M_{hi} - \overline{M}_h\right)^2\frac{\Phi_{hi}^{(12)}}{C_{h1i}^2 m_{hi}} + \overline{M}_h\sum_{i=1}^{N_h}\left(M_{hi} - \overline{M}_h\right)\frac{\Phi_{hi}^{(12)}}{C_{h1i}^2 m_{hi}}\right]. \tag{106}$$

It is observed from (4.54) that if $M_{hi} = \overline{M}_h$, then $V\left(\hat{\lambda}_{aswru}\right) = V\left(\hat{\lambda}_{asppswru}\right)$. Further putting $M_{hi} = \overline{M}_h = \frac{M_{h0}}{N_h}$, the probability $p_{hi} = \frac{M_{hi}}{M_{h0}}$ in PPS is equal to $\frac{1}{N_h}$, which is the probability in the stratified PPS with the equal Probability two-stage sampling.

5. Numerical Study

We have evaluated and compared the suggested estimation methods with those of Singh and Suman (2019) estimators.

5.1 Comparison of the suggested procedure with Singh and Suman (2019) estimators

In this section we present Singh and Suman (2019) estimators under two-stage sampling and stratified two-stage sampling schemes, where first-stage samples are selected from the clustered population using probability proportional to size with replacement (PPSWR) sampling scheme.

(a) When the proportion of persons having the unrelated rare attribute is known under two-stage sampling scheme:

$$\hat{\lambda}^{(1)}_{appswr} = \frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}\hat{\lambda}^{(1)}_{ia}, \tag{107}$$

where $\hat{\lambda}^{(1)}_{ia} = \frac{1}{J_i}\left[(1/m_i)\sum_{j=1}^{m_i} y_{ij} - (1-U_i)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}\right]$

$$J_i = \left[U_i + (1-U_i)P_{1i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\right] \text{ and } P_i = \frac{M_i}{M_0}.$$

Thus

$$\hat{\lambda}^{(1)}_{appswr} = \frac{1}{n}\sum_{i=1}^{n}\hat{\lambda}^{(1)}_{ia} \tag{108}$$

The Variance of $\hat{\lambda}^{(1)}_{appswr}$ is given by

$$V\left(\hat{\lambda}^{(1)}_{appswr}\right) = \frac{1}{nM_0}\left[\sum_{i=1}^{N} M_i\ (\lambda_{ia}-\lambda_a)^2 + \sum_{i=1}^{N}\frac{M_i\eta_i}{m_i}\right], \tag{109}$$

where $\eta_i = \left[\frac{\lambda_{ia}}{J_i} + \frac{(1-U_i)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}}{J_i^2}\right]$ (110)

From (12) and (109) we have

$$V\left(\hat{\lambda}^{(1)}_{appswr}\right)-V\left(\hat{\lambda}_{appswr}\right)=\frac{1}{nM_0}\left[\sum_{i=1}^{N}\frac{M_i}{m_i}\left(\eta_i-\Phi_i\right)\right],$$

$$=\frac{1}{nM_0}\sum_{i=1}^{N}\frac{M_i}{m_i}\frac{(1-U_i)P_{3i}^2k_i}{(k_i-1)J_iD_i}\left\{\lambda_{ia}+\left(\frac{1}{J_i}+\frac{1}{D_i}\right)(1-U_i)P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\lambda_{iy}\right\}$$

$> 0.$

It follows that the proposed estimator $\hat{\lambda}_{appswr}$ is better than the $\hat{\lambda}^{(1)}_{appswr}$.

5.2 When the proportion of persons having the unrelated rare attribute is unknown under two-stage sampling design:

$$\hat{\lambda}^{(1)}_{appswru}=\frac{1}{nM_0}\sum_{i=1}^{n}\frac{M_i}{p_i}\frac{1}{m_iJ_{1i}}\left[d_{2i}\sum_{j=1}^{m_i}y_{i1j}-C_{2i}\sum_{j=1}^{m_i}y_{i2j}\right], \tag{111}$$

$$J_{1i}=\left[U_{1i}(1-U_{2i})Q_{2i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\}-U_{2i}(1-U_{1i})P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\right.$$
$$\left.+(1-U_{1i})(1-U_{2i})(P_{1i}Q_{2i}-P_{2i}Q_{1i})\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\}\right]$$

$$d_{2i}=(1-U_{2i})Q_{2i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\},$$

$$C_{2i}=(1-U_{1i})P_{2i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}.$$

The variance of $\hat{\lambda}^{(1)}_{appswru}$ is given by

$$V\left(\hat{\lambda}^{(1)}_{appswru}\right)=\frac{1}{nM_0}\left[\sum_{i=1}^{N}M_i\left(\lambda_{ia}-\lambda_a\right)^2+\sum_{i=1}^{N}\frac{M_i\eta_i^{(12)}}{J_{1i}^2m_i}\right] \tag{112}$$

$$\eta_i^{(12)}=\left[\left\{d_{2i}^2C_{3i}^{(1)}+C_{2i}^2d_{3i}^{(1)}-2C_{2i}C_{3i}^{(1)}d_{2i}d_{3i}^{(1)}\right\}\lambda_{ia}+\left\{d_{2i}^2C_{2i}+C_{2i}^2d_{2i}-2C_{2i}^2d_{2i}^2\right\}\lambda_{iy}\right], \tag{113}$$

$$d_{3i}^{(1)}=U_{2i}+(1-U_{2i})Q_{1i}\left\{1+Q_{3i}\frac{k_i}{(k_i-1)}\right\},$$

$$C_{3i}^{(1)}=U_{1i}+(1-U_{1i})P_{1i}\left\{1+P_{3i}\frac{k_i}{(k_i-1)}\right\}.$$

5.3 When the proportion of persons having the unrelated rare attribute is known under two-stage sampling design:

$$\hat{\lambda}^{(1)}_{asppswr}=\sum_{h=1}^{L}W_h\frac{1}{n_hM_{h0}}\sum_{i=1}^{n_h}\frac{M_{hi}}{p_{hi}}\hat{\lambda}^{(1)}_{hia},$$

where $\hat{\lambda}^{(1)}_{hia}=\frac{1}{J_{hi}}\left[(1/m_{hi})\sum_{j=1}^{m_{hi}}y_{hij}-(1-U_{hi})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\lambda_{hiy}\right]$

$$J_{hi} = \left[U_{hi} + (1-U_{hi})P_{1hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\right] \text{ and } P_{hi} = \frac{M_{hi}}{M_{h0}}..$$

The Variance of $\hat{\lambda}^{(1)}_{asppswr}$ is given by

$$V\left(\hat{\lambda}^{(1)}_{asppswr}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}}\left[\sum_{i=1}^{N} M_{hi}\ (\lambda_{hia}-\lambda_{ha})^2 + \sum_{i=1}^{N}\frac{M_{hi}\eta_{hi}}{m_{hi}}\right],$$

where $\eta_{hi} = \left[\frac{\lambda_{hia}}{J_{hi}} + \frac{(1-U_{hi})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\lambda_{hiy}}{J_{hi}^2}\right]$.

5.4 When the proportion of persons having the unrelated rare attribute is unknown under stratified two-stage sampling design:

$$\hat{\lambda}^{(1)}_{asppsru} = \sum_{h=1}^{L} W_h \frac{1}{n_h M_{h0}} \sum_{i=1}^{n_h}\frac{M_{hi}}{p_{hi}}\hat{\lambda}_{hiau}\ , \tag{114}$$

where

$$\hat{\lambda}_{hiau} = \frac{1}{J_{h1i}m_{hi}}\left[(1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\sum_{j=1}^{m_{hi}} y_{hi1j} - (1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\sum_{j=1}^{m_{hi}} y_{hi2j}\right], \tag{115}$$

$$J_{hi} = \left[U_{h1i}(1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} - U_{h2i}(1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\} + (1-U_{h1i})(1-U_{h2i})(P_{1hi}Q_{2hi}-P_{2hi}Q_{1hi})\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}\right],$$

$$p_{hi} = \frac{M_{hi}}{M_{h0}}.$$

The Variance of $\hat{\lambda}^{(1)}_{asppsru}$ is given by

$$V\left(\hat{\lambda}^{(1)}_{asppsru}\right) = \sum_{h=1}^{L} W_h^2 \frac{1}{n_h M_{h0}}\left[\sum_{i=1}^{N_h} M_{hi}\ (\lambda_{hia}-\lambda_{ha})^2 + \sum_{i=1}^{N_h}\frac{M_{hi}\eta_{hi}}{m_{hi}J_{h1i}^2}\right],$$

where

$$\eta_{hi} = \left[\left\{d_{h2i}^2 C_{h3i}^{(1)} + C_{h2i}^2 d_{h3i}^{(1)} - 2C_{h2i}C_{h3i}^{(1)}d_{h2i}d_{h3i}^{(1)}\right\}\lambda_{hia} + \left\{d_{h2i}^2 C_{h2i} + C_{h2i}^2 d_{h2i} - 2C_{h2i}^2 d_{h2i}^2\right\}\lambda_{hiy}\right] \tag{116}$$

$$C_{h2i} = (1-U_{h1i})P_{2hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\},$$

$$d_{h2i} = (1-U_{h2i})Q_{2hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\},$$

$$C_{3hi}^{(1)} = U_{h1i} + (1-U_{h1i})P_{1hi}\left\{1+P_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\},$$

$$d_{3hi}^{(1)} = U_{h2i} + (1-U_{h2i})Q_{1hi}\left\{1+Q_{3hi}\frac{k_{hi}}{(k_{hi}-1)}\right\}.$$

The percent relative efficiencies (PRE's) of the proposed estimators $\hat{\lambda}_{appswr}$, $\hat{\lambda}_{appswru}$, $\hat{\lambda}_{asppswr}$ and $\hat{\lambda}_{asppswru}$ with respect to Singh and Suman (2019) estimators $\hat{\lambda}_{appswr}^{(1)}$, $\hat{\lambda}_{appswru}^{(1)}$, $\hat{\lambda}_{asppswr}^{(1)}$ and $\hat{\lambda}_{asppswru}^{(1)}$ are respectively defined by

$$e_1 = PRE\left(\hat{\lambda}_{appswr}, \hat{\lambda}_{appswr}^{(1)}\right) = \frac{V\left(\hat{\lambda}_{appswr}^{(1)}\right)}{V\left(\hat{\lambda}_{appswr}\right)} * 100, \tag{117}$$

$$e_2 = PRE\left(\hat{\lambda}_{appswru}, \hat{\lambda}_{appswru}^{(1)}\right) = \frac{V\left(\hat{\lambda}_{appswru}^{(1)}\right)}{V\left(\hat{\lambda}_{appswru}\right)} * 100, \tag{118}$$

$$e_3 = PRE\left(\hat{\lambda}_{asppswr}, \hat{\lambda}_{asppswr}^{(1)}\right) = \frac{V\left(\hat{\lambda}_{asppswr}^{(1)}\right)}{V\left(\hat{\lambda}_{asppswr}\right)} * 100 \ , \tag{119}$$

$$e_4 = PRE\left(\hat{\lambda}_{asppswru}, \hat{\lambda}_{asppswru}^{(1)}\right) = \frac{V\left(\hat{\lambda}_{asppswru}^{(1)}\right)}{V\left(\hat{\lambda}_{asppswru}\right)} * 100, \tag{120}$$

To carry out the numerical comparison under two stage sampling design, we consider a population of five clusters (N=5) with sizes $M_i = (1000, 2000, 2000, 3000, 4000)$ for $i = 1,2,3,4,5$. Two clusters (n=2) are selected using the PPSWR sampling scheme depending on the cluster sizes as $p_i = \frac{M_i}{M_0}$, where $M_0 = \sum_{i=1}^{5} M_i = 12000$ [see, Lee et al (2013)]. It is assumed that

(a) $\lambda_{1y} = \lambda_{2y} = \lambda_{3y} = \lambda_{4y} = \lambda_{5y} = 1$ for the rare unrelated attribute,

(b) the values of P_{1i}, P_{2i} *and* U_i are same in all clusters; $P_{1i} = P_1$, $P_{2i} = P_2 = \frac{(1-P_1)}{3}$, $P_{3i} = (1-P_{1i}-P_{2i}) = P_3 = (1-P_1-P_2) = \frac{2(1-P_1)}{3}$, [see Nargis and Shabbir (2019)], $U_i = U$ for $i = 1,2,3,4,5$.

(c) the total number of cards, $k_i = k = 100, i = 1,2,3,4,5.$ in a deck for each cluster.

In the stratified PPS, a population is stratified into two strata (h=2), and there are two clusters $N_1 = 2$ with sizes $M_{1i} = (1000, 2000)$ for $i = 1,2$ in stratum 1, and three clusters $N_2 = 3$ with sizes $M_{2i} = (2000, 3000, 4000)$ for $i = 1,2,3$ in stratum 2. We select a cluster from each stratum $(n_1 = n_2 = 1)$. In both procedures, we suppose that the samples in each cluster are drawn with 10% and the parameters for the rare unrelated attribute which was assumed to be known are equal to 1.
We have taken $M_0 = 12000$ and $P_{1hi}, P_{2hi}, P_{3hi}, U_{h1i}$ *and* U_{h2i} are equal for all clusters and strata

$$P_{111} = P_{121} = P_{1,}\ P_{211} = P_{221} = P_2,\ P_{311} = P_{321} = P_3,$$
$$Q_{111} = Q_{121} = Q_{1,}\ Q_{211} = Q_{221} = Q_2,\ Q_{311} = Q_{321} = Q_3,$$
$$U_{111} = U_{112} = U_{211} = U_{212} = U_{213} = U_1,$$
$$U_{121} = U_{122} = U_{221} = U_{222} = U_{223} = U_2,$$
$$W_1 = \frac{N_1}{N},\ W_2 = \frac{N_2}{N},\ k_{hi} = k = 100,\ \lambda_{iy} = 1;\ i = 1 to 5$$

Findings are shown in Tables 1 to 4.

Table 1. PRE's of the estimator $\hat{\lambda}_{appswr}$ with respect to the estimator $\hat{\lambda}^{(1)}_{appswr}$ for $\lambda_{iy} = 1;\ i = 1,2,3,4,5$ and $k = 100$.

U	Mean total number of persons possessing the rare sensitive attribute *A*					P_1 P_2 P_3	0.1 0.3 0.6	0.2 0.267 0.533	0.4 0.2 0.4	0.6 0.133 0.267	0.8 0.067 0.133
	λ_{1a}	λ_{2a}	λ_{3a}	λ_{4a}	λ_{5a}	λ_a		PRE ($= e_1$)			
0.01	1	1	1	1	1	1	629.18	259.94	138.31	111.01	102.11
	1	1	1	1	2		123.61	105.75	100.98	100.22	100.04
	1	1	1	2	1		131	107.57	101.3	100.29	100.05
	1	1	2	1	1		145.09	111.08	101.91	100.43	100.07
	1	2	1	1	1		145.09	111.08	101.91	100.43	100.07
	2	1	1	1	1		182.71	120.67	103.64	100.83	100.13
0.05	1	1	1	1	1	1	480.49	235.3	135.14	110.36	102.01
	1	1	1	1	2		115.96	104.64	100.88	100.21	100.03
	1	1	1	2	1		120.98	106.11	101.16	100.27	100.04
	1	1	2	1	1		130.57	108.96	101.71	100.4	100.07
	1	2	1	1	1		130.57	108.96	101.71	100.4	100.07
	2	1	1	1	1		156.36	116.76	103.26	100.77	100.13
0.1	1	1	1	1	1	1	368.79	211.1	131.53	109.58	101.88
	1	1	1	1	2		110.47	103.6	100.76	100.19	100.03
	1	1	1	2	1		113.77	104.74	101.01	100.25	100.04
	1	1	2	1	1		120.11	106.96	101.49	100.37	100.06
	1	2	1	1	1		120.11	106.96	101.49	100.37	100.06

	2	1	1	1	1		137.28	113.07	102.84	100.7	100.12
0.3	1	1	1	1	1	1	195.24	154.55	120.04	106.79	101.42
	1	1	1	1	2		102.83	101.43	100.43	100.12	100.02
	1	1	1	2	1		103.73	101.89	100.57	100.16	100.03
	1	1	2	1	1		105.49	102.79	100.84	100.24	100.05
	1	2	1	1	1		105.49	102.79	100.84	100.24	100.05
	2	1	1	1	1		110.35	105.29	101.6	100.47	100.09
0.5	1	1	1	1	1	1	141.21	127.45	111.95	104.44	100.98
	1	1	1	1	2		100.97	100.6	100.23	100.08	100.02
	1	1	1	2	1		101.28	100.79	100.3	100.1	100.02
	1	1	2	1	1		101.89	101.17	100.45	100.15	100.03
	1	2	1	1	1		101.89	101.17	100.45	100.15	100.03
	2	1	1	1	1		103.61	102.24	100.86	100.29	100.06

Table 2. PRE's of the estimator $\hat{\lambda}_{appswru}$ with respect to the estimator $\hat{\lambda}^{(1)}_{appswru}$ for $\lambda_{iy} = 1;\ i = 1, 2, 3, 4, 5$ and $k = 100$

U_1	U_2	Mean total number of persons possessing the rare sensitive attribute *A* ↓					P_1 P_2 P_3 Q_1 Q_1 Q_1	0.1 0.3 0.6 0.1 0.3 0.6	0.1 0.3 0.6 0.3 0.233 0.466	0.2 0.267 0.533 0.1 0.3 0.6	0.2 0.267 0.533 0.3 0.233 0.466	0.4 0.2 0.4 0.1 0.3 0.6	0.4 0.2 0.4 0.3 0.233 0.466
		λ_{1a}	λ_{2a}	λ_{3a}	λ_{4a}	λ_{5a}	λ_a	PRE ($= e_2$)					
0.6	0.2	1	1	1	1	1	1	109.61	132.4	107.11	118.6	103.98	107.61
		1	1	1	1	2		100.38	102.86	100.22	100.93	100.09	100.2
		1	1	1	2	1		100.5	103.71	100.29	101.22	100.12	100.26
		1	1	2	1	1		100.73	105.28	100.43	101.78	100.17	100.39
		1	2	1	1	1		100.73	105.28	100.43	101.78	100.17	100.39
		2	1	1	1	1		101.35	109.12	100.81	103.26	100.34	100.74
0.8		1	1	1	1	1		106.88	110.8	105.28	107.77	102.85	103.76
		1	1	1	1	2		100.14	100.24	100.1	100.16	100.05	100.07
		1	1	1	2	1		100.18	100.32	100.13	100.21	100.07	100.09
		1	1	2	1	1		100.27	100.47	100.2	100.31	100.1	100.13
		1	2	1	1	1		100.27	100.47	100.2	100.31	100.1	100.13
		2	1	1	1	1		100.52	110.8	100.38	100.59	100.19	100.25
0.6	0.4	1	1	1	1	1		114.83	277.49	110.12	131.39	105.28	110.56
		1	1	1	1	2		101.12	221.1	100.47	104.93	100.14	100.38
		1	1	1	2	1		101.46	232.12	100.62	106.26	100.18	100.5
		1	1	2	1	1		102.1	245.34	100.9	108.59	100.27	100.73
		1	2	1	1	1		102.1	245.34	100.9	108.59	100.27	100.73
		2	1	1	1	1		103.7	261.49	101.67	113.64	100.51	101.37
0.8		1	1	1	1	1	1	108.85	115.2	106.57	110.43	103.34	104.65
		1	1	1	1	2		100.19	100.41	100.13	100.24	100.06	100.09
		1	1	1	2	1		100.26	100.54	100.17	100.31	100.08	100.11
		1	1	2	1	1		100.38	100.8	100.26	100.46	100.12	100.17
		1	2	1	1	1		100.38	100.8	100.26	100.46	100.12	100.17
		2	1	1	1	1		100.73	101.52	100.5	100.89	100.22	100.32

0.6	0.7	1	1	1	1	1	1	121.72	110.51	142.4	114.4	104.39	146.88
		1	1	1	1	2		102.27	100.35	114.08	100.65	100.6	120.6
		1	1	1	2	1		102.93	100.46	116.99	100.85	100.77	124.12
		1	1	2	1	1		104.14	100.68	121.43	101.25	101.09	129.08
		1	2	1	1	1		104.14	100.68	121.43	101.25	101.09	129.08
		2	1	1	1	1		107.04	101.29	129	102.31	101.84	136.62
0.8		1	1	1	1	1		121.07	186.13	114.46	130.07	106.17	110.59
		1	1	1	1	2		101.07	147.55	100.47	103.01	100.13	100.28
		1	1	1	2	1		101.4	153.89	100.62	103.89	100.17	100.37
		1	1	2	1	1		102.04	162.17	100.91	105.5	100.25	100.54
		1	2	1	1	1		102.04	162.17	100.91	105.5	100.25	100.54
		2	1	1	1	1		103.73	173.48	101.72	109.39	100.48	101.03

Table 3. PRE's of the estimator $\hat{\lambda}_{asppswr}$ with respect to the estimator $\hat{\lambda}^{(1)}_{asppswr}$ for $\lambda_{iy} = 1;\ i = 1,2,3,4,5$ and $k = 100$.

U	Mean total number of persons possessing the rare sensitive attribute *A*					P_1 P_2 P_3	0.1 0.3 0.6	0.2 0.267 0.533	0.4 0.2 0.4	0.6 0.133 0.267	0.8 0.067 0.133
	λ_{1a}	λ_{2a}	λ_{3a}	λ_{4a}	λ_{5a}	λ_a	PRE $(= e_3)$				
0.01	1	1	1	1	1	1	629.18	259.94	138.31	111.01	102.11
	1	1	1	1	2		126.48	106.44	101.1	100.25	100.04
	1	1	1	2	1		134.7	108.47	101.45	100.33	100.05
	1	1	2	1	1		150.31	112.36	102.13	100.48	100.08
	1	2	1	1	1		139.08	109.62	101.67	100.38	100.06
	2	1	1	1	1		172.44	118.09	103.19	100.73	100.12
0.05	1	1	1	1	1	1	480.49	235.3	135.14	110.36	102.01
	1	1	1	1	2		117.9	105.2	100.98	100.23	100.04
	1	1	1	2	1		123.48	106.84	101.29	100.3	100.05
	1	1	2	1	1		134.11	109.99	101.91	100.45	100.07
	1	2	1	1	1		126.5	107.78	101.49	100.35	100.06
	2	1	1	1	1		149.34	114.67	102.85	100.68	100.11
0.1	1	1	1	1	1	1	368.79	211.1	131.53	109.58	101.88
	1	1	1	1	2		111.73	104.03	100.85	100.21	100.03
	1	1	1	2	1		115.41	105.3	101.12	100.28	100.05
	1	1	2	1	1		122.44	107.76	101.66	100.41	100.07
	1	2	1	1	1		117.44	106.05	101.29	100.32	100.05
	2	1	1	1	1		132.63	111.44	102.48	100.62	100.1
0.3	1	1	1	1	1	1	195.24	154.54	120.04	106.79	101.42
	1	1	1	1	2		103.16	101.6	100.48	100.14	100.03
	1	1	1	2	1		104.17	102.11	100.63	100.18	100.03
	1	1	2	1	1		106.12	103.1	100.93	100.27	100.05
	1	2	1	1	1		104.77	102.42	100.73	100.21	100.04
	2	1	1	1	1		109.06	104.63	101.41	100.41	100.08
0.5	1	1	1	1	1	1	141.21	127.45	111.95	104.44	100.98
	1	1	1	1	2		101.08	100.67	100.25	100.08	100.02

	1	1	1	2	1		101.43	100.88	100.33	100.11	100.02
	1	1	2	1	1		102.11	101.3	100.5	100.17	100.03
	1	2	1	1	1		101.65	101.02	100.39	100.13	100.03
	2	1	1	1	1		103.16	101.96	100.75	100.25	100.05

Table 4. PRE's of the estimator $\hat{\lambda}_{asppswru}$ with respect to the estimator $\hat{\lambda}^{(1)}_{asppswru}$ for $\lambda_{iy} = 1;\ i = 1,2,3,4,5$ and $k = 100$.

U_1	U_2	Mean total number of persons possessing the rare sensitive attribute *A*					P_1 P_2 P_3 Q_1 Q_1 Q_1	0.1 0.3 0.6 0.1 0.3 0.6	0.1 0.3 0.6 0.3 0.233 0.466	0.2 0.267 0.533 0.1 0.3 0.6	0.2 0.267 0.533 0.3 0.233 0.466	0.4 0.2 0.4 0.1 0.3 0.6	0.4 0.2 0.4 0.3 0.233 0.466
		λ_{1a}	λ_{2a}	λ_{3a}	λ_{4a}	λ_{5a}	λ_a	PRE ($= e_4$)					
0.6	0.2	1	1	1	1	1	1	109.61	132.4	107.11	118.6	103.98	107.61
		1	1	1	1	2		100.61	104.57	100.36	101.51	100.14	100.32
		1	1	1	2	1		100.8	105.82	100.47	101.96	100.19	100.42
		1	1	2	1	1		101.15	108.02	100.68	102.8	100.28	100.62
		1	2	1	1	1		100.5	103.66	100.29	101.21	100.12	100.26
		2	1	1	1	1		100.95	106.61	100.57	102.29	100.23	100.51
0.8		1	1	1	1	1		106.88	110.8	105.28	107.77	102.85	103.76
		1	1	1	1	2		100.33	100.57	100.24	100.37	100.12	100.16
		1	1	1	2	1		100.43	100.75	100.31	100.49	100.15	100.21
		1	1	2	1	1		100.62	101.09	100.45	100.71	100.22	100.3
		1	2	1	1	1		100.16	100.28	100.12	100.18	100.06	100.08
		2	1	1	1	1		100.32	100.55	100.23	100.36	100.11	100.15
0.6	0.4	1	1	1	1	1		114.83	277.49	110.12	131.39	105.28	110.56
		1	1	1	1	2		101.78	239.51	100.75	107.5	100.22	100.61
		1	1	1	2	1		102.29	247.81	100.98	109.28	100.29	100.8
		1	1	2	1	1		103.19	257.16	101.41	112.18	100.42	101.16
		1	2	1	1	1		101.47	232.15	100.63	106.23	100.18	100.5
		2	1	1	1	1		102.69	253.85	101.19	110.54	100.36	100.96
0.8		1	1	1	1	1	1	108.85	115.2	106.57	110.43	103.34	104.65
		1	1	1	1	2		100.46	100.97	100.31	100.56	100.14	100.2
		1	1	1	2	1		100.6	101.27	100.41	100.74	100.18	100.27
		1	1	2	1	1		100.87	101.82	100.6	101.07	100.27	100.39
		1	2	1	1	1		100.23	100.48	100.16	100.28	100.07	100.1
		2	1	1	1	1		100.45	100.93	100.31	100.54	100.14	100.2
0.6		1	1	1	1	1		121.72	110.51	142.4	114.4	104.39	146.88
		1	1	1	1	2		103.54	100.57	119.33	101.05	100.85	126.71
		1	1	1	2	1		104.49	100.75	122.47	101.36	101.08	130.07
		1	1	2	1	1		106.12	101.08	126.83	101.96	101.47	134.41
		1	2	1	1	1		102.95	100.47	117.03	100.86	100.86	124.24
		2	1	1	1	1		105.26	100.9	124.9	101.63	101.52	132.84
0.8	0.7	1	1	1	1	1	1	121.07	186.13	114.46	130.07	106.17	110.59
		1	1	1	1	2		102.43	165.09	101.1	106.46	100.3	100.65

		1	1	1	2	1		103.12	169.55	101.43	108.05	100.39	100.85
		1	1	2	1	1		104.37	174.67	102.04	110.69	100.57	101.22
		1	2	1	1	1		101.26	151.58	100.56	103.5	100.15	100.33
		2	1	1	1	1		102.39	166.29	101.08	106.34	100.3	100.64

Tables 1 and 3 demonstrate the PREs of the estimators $\hat{\lambda}_{appswr}$ and $\hat{\lambda}_{appswru}$ under two-stage sampling scheme for the known and unknown unrelated rare attribute.

Tables 2 and 4 demonstrate the PREs of the estimators $\hat{\lambda}_{asppswr}$ and $\hat{\lambda}_{asppswru}$ under two-stage stratified sampling scheme for the known and unknown unrelated rare attribute.

It is observed from Tables 1-4 that the PRE's exceed 100% for all chosen parametric values which indicates that proposed model and estimation procedures perform better than the Singh and Suman (2019) RRT model.

It is further observed from Tables 1 and 3 that the value of PREs increase for decreasing value of *U* with considerable gain in efficiency for smaller values of *U*.

From Tables 1 and 3, it is clearly seen that there is higher PRE's for smaller values of P_i 's for different sets of mean total number of persons possessing the rare sensitive attribute *A* i.e. λ_{ia}.

Tables 2 and 4 exhibit that there is gain in PREs for smaller values of both P_{hi} 's and Q_{hi} 's. It is also observed that we get higher PREs when the value of P_{hi} 's is smaller than Q_{hi} 's. Thus if $P_{hi}\text{'s} < Q_{hi}\text{'s}$, the proposed estimators $\hat{\lambda}^{(1)}_{appswru}$ and $\hat{\lambda}^{(1)}_{asppswru}$ perform better than the $\hat{\lambda}_{appswru}$ and $\hat{\lambda}_{asppswru}$ respectively.

6. Conclusion

In this paper we have proposed alternative estimation procedures to estimate the mean number of persons having a rare sensitive attribute under a Poisson distribution using two-stage and stratified two-stage sampling schemes. Probability proportional to size sampling scheme has been used to drawn the first-stage units and simple random sampling with replacement is used to select the second-stage units. We have obtained the variances of the suggested estimators and their estimates, if the parameter of the unrelated non-sensitive attribute is either known or unknown. It is proved that the estimators using Narjis and Shabbir (2020) randomized response model under two stage and stratified two stage sampling schemes are advantageous in terms of PRE's when compared with Singh and Suman (2019). With theoretical and empirical results associated with proposed estimators we also observed that in proposed estimation procedures design parameters play important roles in increasing or decreasing the efficiency. Thus, the envisaged two-stage unrelated question RRT is therefore recommended to the survey practitioners for its use in practice.

References

Land, M., Singh, S. and Sedory, S.A. (2012): Estimation of a rare sensitive attribute using Poisson distribution. Statistics,46, 351–360.

Lee, G.S., Uhm, D. and Kim, J.M. (2013): Estimation of a rare sensitive attribute in probability proportional to size using Poisson distribution. Statistics, DOI: 10.1080/02331888.2012.760091

Mangat, N.S. (1992): Two-stage randomized response sampling procedure using unrelated question. Journal of the Indian Society of Agricultural Statistics,44(1), 82-87.

Narjis, G. and Shabbir, J. (2019): An efficient partial randomized response model for estimating a rare sensitive attribute using Poisson distribution. Communications in Statistics- Theory and Methods, DOI: 10.1080/03610926.2019.1628992.

Narjis, G. and Shabbir, J. (2020): A two-stage unrelated question randomized response model for estimating the rare sensitive parameter under Poisson distribution. Communications in Statistics -Theory and Methods, DOI:10.1080/03610926.2020.1769671.

Singh, G. N. and Suman, S. (2019): An ameliorated two-stage randomized response model for estimating a rare stigmatized characteristic using Poisson distribution Kuwait Journal of Science,46, 21–30..

Singh, G. N., Singh, C. and Suman, S. (2019): Revisit of a randomized response model for estimating a rare sensitive attribute under probability proportional to size sampling using Poisson probability distributionCommunications in Statistics- Theory and Methods, DOI: 10.1080/03610926.2018.1508718.

Singh, G. N., Singh, C., Suman, S. and Kumar, A. (2017): A two-stage unrelated randomized response model for estimating the rare sensitive attribute in probability proportional to size sampling using Poisson distribution. Communications in Statistics- Theory and Methods, DOI 10.1080/03610926.2017.1361992.

Singh, G. N., Suman, S. and Singh, C. (2020): Estimation of a rare sensitive attribute in two-stage sampling using a randomized response model under Poisson distribution. Mathematical Population Studies,27(2), 81-114.

Singh, H.P. and Tarray, T.A. (2014): A stratified mangat and singh's optional randomized response model using proportional and optimal allocation. Statistica,73(1),65-83.

Singh, H. P. and Tarray, T. A. (2017): A two-stage land et al.'s randomized response model for estimating a rare sensitive attribute using Poisson distributionCommunications in Statistics- Theory and Methods, 46 (1),389–405

Tarray, T. A. and Singh, H. P. (2015): A randomized response model for estimating a rare sensitive attribute in stratified sampling using Poisson distribution. Model Assisted Statistics and Applications,10(4), 345–360.

Tarray, T. A., Singh, H. P. and Masood, S. (2019): An endowed randomized response model for estimating a rare sensitive attribute using Poisson distribution. Trends in Applied Sciences Research, 14 (1), 1-6.

Warner, S.L. (1965): Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60, 63-69.