Hacettepe Journal of Mathematics and Statistics
Volume 40 (1) (2011), 115 – 124

RATIO ESTIMATOR FOR THE POPULATION MEAN AT THE CURRENT OCCASION IN THE PRESENCE OF NON-RESPONSE IN SUCCESSIVE SAMPLING

Esra Satıcı* and Cem Kadılar†‡

Received 14 : 10 : 2009 : Accepted 14 : 06 : 2010

Abstract

In this article, a new estimator for the population mean at the current occasion in successive sampling having two occasions is proposed under the missing data case. The minimum mean square error and optimum fraction of a fresh sample of the proposed estimator are obtained. This estimator is compared with the estimator suggested by Singh and Priyanka (Singh, G. N., Priyanka, K. *Effect of non-response on current occasion in search of good rotation patterns on successive occasions*, Statistics in Transition **8** (2), 273–292, 2007), and the efficient condition for the proposed estimator is found. In a numerical example, the mean square errors of these estimators are also computed according to various values for the fraction of missing data, coefficient of correlation, fraction of sub sampling and fraction of fresh sample. The results of the numerical example show that the proposed estimator is more efficient than Singh-Priyanka estimator for all values of the parameters.

Keywords: Missing data, Successive sampling, Ratio estimator, Optimal replacement policy.

2000 AMS Classification: 62 D 05.

---

*General Directorate of Highways, Directorate of Strategy Development, Ankara, Turkey. E-mail: esra.satici@gmail.com

†Department of Statistics, Hacettepe University, 06800 Beytepe, Ankara, Turkey. E-mail: kadilar@hacettepe.edu.tr

‡Corresponding Author.

1. Introduction and notations

It is often seen that a population having a large number of elements remains unchanged on several occasions but the value of the unit's changes. Successive (rotation) sampling provides a strong tool for generating reliable estimates at different occasions. Successive sampling has been extensively used to provide more efficient estimates of population characteristics such as the mean. In successive sampling, it is common practice to use the information collected on a previous occasion to improve the precision of the estimates for the current occasion.

The problem of sampling on two successive occasions with a partial replacement of sampling units was first considered by Jessen [4] in the analysis of a survey which collected farm data. The theory of successive sampling was extended by Patterson [5], Sen [7, 8, 9], Chaturvedi and Tripathi [1]. Later, Singh *et al.* [12] and Singh and Singh [11] used the auxiliary information on the current occasion in successive sampling. Utilizing the auxiliary information on both the occasions, Singh [10] and Singh and Priyanka [14] proposed chain type ratio and regression type estimators in the successive sampling, respectively.

It is well known that missing data is a common problem in sample surveys. Hansen and Hurwitz [3] suggested a technique for handling non-response. More recently, Singh and Priyanka [13] used the Hansen and Hurwitz technique for estimation of the population mean on the current occasion in the context of sampling on two occasions. However, in this article, using Hansen and Hurwitz [3] technique, we propose a ratio-type estimator to estimate the population mean at the current occasion in the presence of non-response on two occasions of the successive sampling, and compare the proposed estimator with the Singh-Priyanka estimator.

Let $U = (U_1, U_2, \ldots, U_N)$ be a finite population of N units sampled over two occasions. Let a simple random sample of size n be selected on the first occasion. The character under study is denoted by x (y) on the first (second) occasion, respectively. It is assumed that information on an auxiliary variable z, is available for both of the occasions. A random sub-sample of $m = n\lambda$ units is retained (matched) for use on the current (second) occasion. It is assumed that there is non-response on the current occasion, so that the population can be divided into two classes with a response size of N_1 at the first attempt and a non-response size of N_2. On the current occasion, a simple random sample of $u = n - m = n\mu$ units are drawn afresh from the entire population. In this way, the size of the second sample is also n. The terms λ and μ, $(\lambda + \mu = 1)$, are the fractions of matched and fresh samples on the second occasion, respectively. It is assumed that u_1 units respond and u_2 units do not respond in the unmatched portion of the second sample.

Let a random sub-sample of u_{2h} units be drawn from the non-response size of u_2 on the current occasion. The following notations are used in the rest of the article:

$\bar{X}, \bar{Y}, \bar{Z}$: The population means of the variables x, y, and z, respectively.

$\bar{x}_n$, $\bar{z}_n$, $\bar{y}_m$, $\bar{x}_m$, $\bar{z}_m$, $\bar{y}_{u1}$, $\bar{y}_{u2h}$, $\bar{z}_u$: The sample means of the related variables according to the sample sizes shown in the subscripts.

ρ_{yx}, ρ_{xz}, ρ_{yz}: The population correlation coefficients between the variables shown in the subscripts.

S_x^2, S_y^2, S_z^2: The population variances of x, y, and z, respectively.

$W = \frac{N_2}{N}$: The proportion of non-response units in the population.

2. Suggested estimator

Singh and Priyanka [13] suggested an estimator depending on two estimators for the population mean of the study variable, $\bar{Y}$, on the second occasion. The first of these, which is based on the Hansen and Hurwitz [3] estimator using a sample of size u drawn afresh on the second occasion, is given by

$$T_{\mathrm{SP1}} = \bar{y}_u^* = \frac{u_1\bar{y}_{u1} + u_2\bar{y}_{u2h}}{u}. \tag{1}$$

The second estimator, which is a regression type estimator using a sample of size m common occasions, is defined as

$$T_{\mathrm{SP2}} = \bar{y}_m + b_{yx}\left(\bar{x}_n - \bar{x}_m\right), \tag{2}$$

where b_{yx} is the sample regression coefficient.

Considering a convex linear combination of the estimators, T_{SP1} and T_{SP2}, Singh and Priyanka [13] obtained the estimator of $\bar{Y}$ as follows:

$$T_{\mathrm{SP}} = \varphi T_{\mathrm{SP1}} + (1-\varphi)T_{\mathrm{SP2}}, \tag{3}$$

where φ is an unknown constant to be determined under a certain criterion.

By modifying the estimator in (1), we propose a ratio estimator which is based on a sample of size u drawn afresh on the second occasion and given by,

$$T_{\mathrm{SK1}} = \frac{\bar{y}_u^*}{\bar{z}_u^*}\bar{Z}, \tag{4}$$

where $\bar{z}_u^* = \frac{u_1\bar{z}_{u1}+u_2\bar{z}_{u2h}}{u}$.

Combining the estimators T_{SK1} and T_{SP2} similar to Singh and Priyanka [13], we obtain the estimator of $\bar{Y}$ as follows:

$$T_{\mathrm{SK}} = \gamma T_{\mathrm{SK1}} + (1-\gamma)T_{\mathrm{SP2}}, \tag{5}$$

where γ is an unknown constant to be determined under a certain criterion.

2.1. Theorem. *The Mean Square Error (MSE) of T_{SK} is given by*

$$\mathrm{MSE}(T_{\mathrm{SK}}) = \gamma^2\mathrm{MSE}(T_{\mathrm{SK1}}) + (1-\gamma)^2\mathrm{MSE}(T_{\mathrm{SP2}}), \tag{6}$$

where

$$\mathrm{MSE}(T_{\mathrm{SK1}}) = \left[\left(\frac{1}{u}-\frac{1}{N}\right)+\frac{W(k-1)}{u}\right]\left(S_y^2 - 2R_z\rho_{yz}S_yS_z + R_z^2S_z^2\right),$$
$$\mathrm{MSE}(T_{\mathrm{SP2}}) = \left[\left(\frac{1}{m}-\frac{1}{n}\right)\left(1-\rho_{yx}^2\right)+\left(\frac{1}{n}-\frac{1}{N}\right)\right]S_y^2,$$

and $R_x = \frac{\bar{Y}}{\bar{X}}$, $R_z = \frac{\bar{Y}}{\bar{Z}}$, $k = \frac{u_2}{u_{2h}}$.

Proof. Since the samples are independent, we ignore the covariance term and the MSE of TSK can be given by

$$\mathrm{MSE}(T_{\mathrm{SK}}) = \gamma^2\mathrm{MSE}(T_{\mathrm{SK1}}) + (1-\gamma)^2\mathrm{MSE}(T_{\mathrm{SP2}}), \tag{7}$$

where $\mathrm{MSE}(T_{\mathrm{SK1}})$ can be found using the Taylor Series Method to the first degree of approximation, namely

$$h\left(\bar{y}_u^*, \bar{z}_u^*\right) - h\left(\bar{Y}, \bar{Z}\right) \cong \left.\frac{\partial h\,(a,b)}{\partial a}\right|_{\bar{Y},\bar{Z}}\left(\bar{y}_u^* - \bar{Y}\right) + \left.\frac{\partial h\,(a,b)}{\partial b}\right|_{\bar{Y},\bar{Z}}\left(\bar{z}_u^* - \bar{Z}\right), \tag{8}$$

where

$$h\left(\bar{y}_u^*, \bar{z}_u^*\right) = T_{\mathrm{SK1}},$$
$$\left.\frac{\partial h\ (a,b)}{\partial a}\right|_{\bar{Y},\bar{Z}} = \left.\frac{\partial T_{\mathrm{SK1}}}{\partial \bar{y}_u^*}\right|_{\bar{Y},\bar{Z}} = 1,$$
$$\left.\frac{\partial h\ (a,b)}{\partial b}\right|_{\bar{Y},\bar{Z}} = \left.\frac{\partial T_{\mathrm{SK1}}}{\partial \bar{z}_u^*}\right|_{\bar{Y},\bar{Z}} = -\frac{\bar{Y}}{\bar{Z}} = -R_z.$$

Using (8), we can write

$$T_{\mathrm{SK1}} - \bar{Y} \cong \left(\bar{y}_u^* - \bar{Y}\right) - \frac{\bar{Y}}{\bar{Z}}\left(\bar{z}_u^* - \bar{Z}\right),$$
$$\left(T_{\mathrm{SK1}} - \bar{Y}\right)^2 \cong \left(\bar{y}_u^* - \bar{Y}\right)^2 + \frac{\bar{Y}^2}{\bar{Z}^2}\left(\bar{z}_u^* - \bar{Z}\right)^2 - 2\frac{\bar{Y}}{\bar{Z}}\left(\bar{y}_u^* - \bar{Y}\right)\left(\bar{z}_u^* - \bar{Z}\right),$$
$$\mathrm{E}\left(T_{\mathrm{SK1}} - \bar{Y}\right)^2 \cong E\left(\bar{y}_u^* - \bar{Y}\right)^2 + \frac{\bar{Y}^2}{\bar{Z}^2}E\left(\bar{z}_u^* - \bar{Z}\right)^2 - 2\frac{\bar{Y}}{\bar{Z}}E\left(\bar{y}_u^* - \bar{Y}\right)\left(\bar{z}_u^* - \bar{Z}\right),$$

$$\mathrm{MSE}(T_{\mathrm{SK1}}) \cong V\left(\bar{y}_u^*\right) + \frac{\bar{Y}^2}{\bar{Z}^2}V\left(\bar{z}_u^*\right) - 2\frac{\bar{Y}}{\bar{Z}}\mathrm{Cov}\left(\bar{y}_u^*, \bar{z}_u^*\right), \tag{9}$$

where

$$\begin{aligned}\bar{y}_u^* &= \frac{u_1}{u}\bar{y}_{u1} + \frac{u_2}{u}\bar{y}_{u2h} - \frac{u_2}{u}\bar{y}_{u2} + \frac{u_2}{u}\bar{y}_{u2}\\ &= \frac{u_1}{u}\bar{y}_{u1} + \frac{u_2}{u}\bar{y}_{u2} + \frac{u_2}{u}\left(\bar{y}_{u2h} - \bar{y}_{u2}\right)\\ &= \bar{y}_u + \frac{u_2}{u}\left(\bar{y}_{u2h} - \bar{y}_{u2}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}V\left(\bar{y}_u^*\right) &= V\left(\bar{y}_u\right) + \frac{u_2^2}{u^2}V\left(\bar{y}_{u2h} - \bar{y}_{u2}\right) + 2\frac{u_2}{u}\mathrm{cov}\left[\bar{y}_u, \left(\bar{y}_{u2h} - \bar{y}_{u2}\right)\right]\\ &= V\left(\bar{y}_u\right) + \frac{u_2^2}{u^2}V\left(\bar{y}_{u2h} - \bar{y}_{u2}\right),\end{aligned}$$

since $\mathrm{cov}\left[\bar{y}_u, \left(\bar{y}_{u2h} - \bar{y}_{u2}\right)\right] = 0$ [2]. Hence,

$$\begin{aligned}V\left(\bar{y}_{u2h} - \bar{y}_{u2}\right) &= V\left(\bar{y}_{u2h}\right) + V\left(\bar{y}_{u2}\right) - 2\mathrm{cov}\left(\bar{y}_{u2h}, \bar{y}_{u2}\right)\\ &= V\left(\bar{y}_{u2h}\right) - V\left(\bar{y}_{u2}\right)\\ &= \left(\frac{1}{u_{2h}} - \frac{1}{N_2}\right)S_{y2}^2 - \left(\frac{1}{u_2} - \frac{1}{N_2}\right)S_{y2}^2\\ &= \left(\frac{1}{u_{2h}} - \frac{1}{u_2}\right)S_{y2}^2,\end{aligned}$$

since $\mathrm{cov}\left(\bar{y}_{u2h}, \bar{y}_{u2}\right) = V\left(\bar{y}_{u2}\right)$ [2]. Then, we get:

$$\begin{aligned}V\left(\bar{y}_u^*\right) &= \left(\frac{1}{u} - \frac{1}{N}\right)S_y^2 + \frac{u_2^2}{u^2}\left(\frac{1}{u_{2h}} - \frac{1}{u_2}\right)S_{y2}^2\\ &= \left(\frac{1}{u} - \frac{1}{N}\right)S_y^2 + \frac{u_2}{u^2}\left(\frac{u_2 - u_{2h}}{u_{2h}}\right)S_{y2}^2\\ &= \left(\frac{1}{u} - \frac{1}{N}\right)S_y^2 + \frac{w}{u}\left(k-1\right)S_{y2}^2,\end{aligned}$$

where $w = \frac{u_2}{u}$ and $k = \frac{u_2}{u_{2h}}$. Averaging over the distribution of the fraction w of non-response in the sample, $V\left(\bar{y}_u^*\right)$ can also be expressed as

$$V\left(\bar{y}_u^*\right) = \left(\frac{1}{u} - \frac{1}{N}\right)S_y^2 + \frac{W}{u}\left(k-1\right)S_{y2}^2.$$

Similarly, $V(\bar{z}_u^*)$ can be obtained as

$$V(\bar{z}_u^*) = \Big(\frac{1}{u} - \frac{1}{N}\Big)S_z^2 + \frac{W(k-1)}{u}S_{z2}^2.$$

The last term of (9) can be given by

$$\begin{aligned}
\operatorname{cov}(\bar{y}_u^*, \bar{z}_u^*) &= \operatorname{cov}\Big[\Big(\bar{y}_u + \frac{u_2}{u}(\bar{y}_{u2h} - \bar{y}_{u2})\Big), \Big(\bar{z}_u + \frac{u_2}{u}(\bar{z}_{u2h} - \bar{z}_{u2})\Big)\Big] \\
&= \operatorname{cov}(\bar{y}_u, \bar{z}_u) + \frac{u_2}{u}\operatorname{cov}[\bar{y}_u, (\bar{z}_{u2h} - \bar{z}_{u2})] + \frac{u_2}{u}\operatorname{cov}[\bar{z}_u, (\bar{y}_{u2h} - \bar{y}_{u2})] \\
&\qquad + \frac{u_2^2}{u^2}\operatorname{cov}[(\bar{y}_{u2h} - \bar{y}_{u2}), (\bar{z}_{u2h} - \bar{z}_{u2})] \\
&= \operatorname{cov}(\bar{y}_u, \bar{z}_u) + \frac{u_2^2}{u^2} cov[(\bar{y}_{u2h} - \bar{y}_{u2}), (\bar{z}_{u2h} - \bar{z}_{u2})] \\
&= \Big(\frac{1}{u} - \frac{1}{N}\Big)S_{yz}^2 + \frac{u_2^2}{u^2}\left(\frac{1}{u_{2h}} - \frac{1}{u_2}\right)S_{yz(2)}^2 \\
&= \Big(\frac{1}{u} - \frac{1}{N}\Big)S_{yz} + \frac{w(k-1)}{u}S_{yz(2)} \\
&= \Big(\frac{1}{u} - \frac{1}{N}\Big)S_{yz} + \frac{W(k-1)}{u}S_{yz(2)},
\end{aligned}$$

since $\operatorname{cov}[\bar{y}_u, (\bar{z}_{u2h} - \bar{z}_{u2})] = \operatorname{cov}[\bar{z}_u, (\bar{y}_{u2h} - \bar{y}_{u2})] = 0$ [2], and we note that $\mathrm{E}(w) = W$.

As in Singh and Priyanka [13], we also assume that $S_{y2}^2 = S_y^2$, $S_{z2}^2 = S_z^2$ and $S_{yz(2)} = S_{yz}$. Finally, (9) can be written as

$$\begin{aligned}
\mathrm{MSE}(T_{\mathrm{SK1}}) &= \left[\Big(\frac{1}{u} - \frac{1}{N}\Big) + \frac{W(k-1)}{u}\right]\left[S_y^2 + \frac{\bar{Y}^2}{\bar{Z}^2}S_z^2 - 2\frac{\bar{Y}}{\bar{Z}}S_{yz}\right] \\
&= \left[\Big(\frac{1}{u} - \frac{1}{N}\Big) + \frac{W(k-1)}{u}\right]\left(S_y^2 - 2R_z\rho_{yz}S_yS_z + R_z^2S_z^2\right), \qquad (10)
\end{aligned}$$

where $\rho_{yz} = \frac{S_{yz}}{S_yS_z}$.

Similarly, it is known that

$$\begin{aligned}
\mathrm{MSE}(T_{\mathrm{SP2}}) &= \mathrm{E}\left(T_{\mathrm{SP2}} - \bar{Y}\right)^2 = \mathrm{E}\left[\bar{y}_m + b_{yx}(\bar{x}_n - \bar{x}_m) - \bar{Y}\right]^2 \\
&= \left[\Big(\frac{1}{m} - \frac{1}{n}\Big)(1 - \rho_{yx}^2) + \Big(\frac{1}{n} - \frac{1}{N}\Big)\right]S_y^2. \qquad (11)
\end{aligned}$$

Using (10) and (11) in (7), it is clear that we can obtain $\mathrm{MSE}(T_{\mathrm{SK}})$ as in (6). □

To obtain the optimum value of γ in $MSE(T_{\mathrm{SK}})$, the MSE equation is minimized with respect to γ, and subsequently the optimum value of γ is obtained as

$$\gamma_{\mathrm{opt}} = \frac{\mathrm{MSE}(T_{\mathrm{SP2}})}{\mathrm{MSE}(T_{\mathrm{SK1}}) + \mathrm{MSE}(T_{\mathrm{SP2}})}. \qquad (12)$$

Substituting the value of γ_{opt} in (6), we get

$$\mathrm{MSE}_{\min}(T_{\mathrm{SK}}) = \frac{\mathrm{MSE}(T_{\mathrm{SK1}})\,\mathrm{MSE}(T_{\mathrm{SP2}})}{\mathrm{MSE}(T_{\mathrm{SK1}}) + \mathrm{MSE}(T_{\mathrm{SP2}})}. \qquad (13)$$

2.2. Theorem. *$MSE_{\min}(T_{\mathrm{SK}})$ is derived as*

$$\mathrm{MSE}_{\min}(T_{\mathrm{SK}}) = \frac{K_1\mu^2 + K_2\mu + K_3}{K_4\mu^2 + K_5\mu + K_6}S_y^2, \qquad (14)$$

where $\mu = \frac{u}{n}$ *is the fraction of fresh sample taken on the second occasion,*

$$K_1 = n^2 A_1 A_3 - n A_1 (1-\rho_{yx}^2),\ K_2 = N A_1 A_2 (1-\rho_{yx}^2) - n A_1 A_3 (N A_2 + n),$$
$$K_3 = N n A_1 A_2 A_3,\ K_4 = n^2 A_1 + N n (1-\rho_{yx}^2) S_y^2 - N n^2 A_3 S_y^2,$$
$$K_5 = N n^2 A_3 S_y^2 - N n A_1 A_2 - n^2 A_1,\ \text{and } K_6 = N n A_1 A_2.$$

Here $A_1 = S_y^2 - 2R_z\rho_{yz}S_yS_z + R_z^2S_z^2$, $A_2 = 1 + W(k-1)$, *and* $A_3 = \frac{1}{n} - \frac{1}{N}$.

Proof. Using (10) and (11) in (13), the minimum MSE of the proposed estimator given in (14) is easily obtained. □

2.3. Theorem. *The optimum value of* μ *and the minimum value of* MSE(T_{SK}) *with respect to both* γ *and* μ *are given by*

$$\hat{\mu} = \frac{-\Delta_2 \pm \sqrt{\Delta_2^2 - 4\Delta_1\Delta_3}}{2\Delta_1} = \mu_0 \tag{15}$$

and

$$\text{MSE}_{\min^*}(T_{\text{SK}}) = \frac{K_1\mu_0^2 + K_2\mu_0 + K_3}{K_4\mu_0^2 + K_5\mu_0 + K_6} S_y^2, \tag{16}$$

respectively, where $\Delta_1 = K_1K_5 - K_2K_4$, $\Delta_2 = 2K_1K_6 - 2K_3K_4$, *and* $\Delta_3 = K_2K_6 - K_3K_5$.

Proof. To determine the optimum value of μ, so that $\bar{Y}$ may be estimated with maximum precision, we minimize $\text{MSE}_{\min}(T_{\text{SK}})$ with respect to μ, which results in a quadratic equation as $\Delta_1\mu^2 + \Delta_2\mu + \Delta_3 = 0$. We obtain $\hat{\mu}$, say μ_0, by solving this equation for μ. We are certain that a real value of $\hat{\mu}$ exists if $\Delta_2^2 - 4\Delta_1\Delta_3 \geq 0$. For certain situations, two real values of $\hat{\mu}$ are possible and in these cases it should be remembered that $0 \leq \hat{\mu} \leq 1$, therefore all other values of $\hat{\mu}$ are inadmissible. Substituting the admissible value of $\hat{\mu}$, say μ_0, computed using (15) into (14), we get the value of $MSE_{\min^*}(T_{\text{SK}})$ given in (16). □

3. Efficiency comparison

In this section, we compare the proposed estimator given in (5) with the Singh-Priyanka estimator given in (3), as follows:

$$\begin{aligned}
&\text{MSE}_{\min}(T_{SK}) < \text{MSE}_{\min}(T_{SP}),\\
&\frac{\text{MSE}(T_{SK1})\text{MSE}(T_{SP2})}{\text{MSE}(T_{SK1}) + \text{MSE}(T_{SP2})} < \frac{\text{MSE}(T_{SP1})\text{MSE}(T_{SP2})}{\text{MSE}(T_{SP1}) + \text{MSE}(T_{SP2})},\\
&\text{MSE}(T_{SK1})[\text{MSE}(T_{SP1}) + \text{MSE}(T_{SP2})] < \text{MSE}(T_{SP1})[\text{MSE}(T_{SK1})\\
&\qquad\qquad + \text{MSE}(T_{SP2})],\\
&\text{MSE}(T_{SK1})\text{MSE}(T_{SP2}) < \text{MSE}(T_{SP1})\text{MSE}(T_{SP2}),\\
&\text{MSE}(T_{SK1}) < \text{MSE}(T_{SP1}),\\
&\left[\left(\frac{1}{u} - \frac{1}{N}\right) + \frac{W(k-1)}{u}\right](S_y^2 - 2R_z\rho_{yz}S_yS_z + R_z^2S_z^2)\\
&\qquad\qquad < \left[\left(\frac{1}{u} - \frac{1}{N}\right) + \frac{(k-1)W}{u}\right] S_y^2,\\
&S_y^2 - 2R_z\rho_{yz}S_yS_z + R_z^2S_z^2 < S_y^2,\\
&R_z^2S_z^2 < 2R_z\rho_{yz}S_yS_z,\\
&R_zS_z < 2\rho_{yz}S_y,
\end{aligned}$$

$$\frac{1}{2}R_z\frac{S_z}{S_y} < \rho_{yz},$$

$$\frac{1}{2}\frac{C_z}{C_y} < \rho_{yz}, \tag{17}$$

where we assume that population means of y and z are positive. When the condition in (17) is satisfied, the proposed ratio estimator is more efficient than the Singh-Priyanka estimator.

4. Numerical illustration

We have used data in Satıcı [6] based on 923 districts of Turkey. Considering the numbers of teachers (x) and private teaching institutions (z) as the auxiliary variables, the population mean of the number of successful students in the student selection examination for secondary schools (y) is estimated using the proposed and the Singh-Priyanka estimators. The mean square errors of these estimators are also calculated according to various values for the fraction of missing data (W), coefficient of correlation, fraction of sub-sampling (k) and fraction of fresh samples (μ), and in this way we have examined the relative efficiencies computed by

$$\mathrm{RE} = \frac{\mathrm{MSE}_{\min}\,(T_{SP})}{\mathrm{MSE}_{\min}\,(T_{SK})}. \tag{18}$$

In Table 1, we observe the population statistics about the study and the auxiliary variables for the years 2006 and 2007. Note that we take only 261 homogenous districts for the population as we use simple random sampling in this article, and that we take the size of the sample as $n = 90$.

Table 1. Data statistics

$N = 261$	$\overline{X}_{2007} = 312.33$	$\overline{Z}_{2007} = 2.05$
$\overline{Y}_{2006} = 222.58$	$\overline{X}_{2006} = 306.45$	$\overline{Z}_{2006} = 2.07$
$S_y^2 = 172386.4$	$S_x^2 = 290706.7$	$S_z^2 = 17.503$
$C_y = 1.86537$	$C_x = 1.75941$	$C_z = 2.02126$
$\rho_{yx} = 0.970$	$\rho_{yz} = 0.935$	$\rho_{xz} = 0.928$

The values of MSE of the proposed and Singh-Priyanka estimators are computed as described in Section 2 according to various values for the parameters, such as ρ_{yx}, W, k, μ, and ρ_{yz}. In these values, $\mu = 0.78$ is the optimal value of the Singh-Priyanka estimator and $\mu = 0.97$ is the optimal value of the proposed estimator. In Tables 2 and 3, relative efficiencies computed using (18) are given for $\rho_{yx} = 0.97$ and $\rho_{yx} = 0.80$, respectively. Especially, we also try $\rho_{yx} = 0.80$ since the performance of the Singh-Priyanka estimator depends on the value of ρ_{yx}.

For this numerical illustration, from Tables 2 and 3, it is clearly observed that the proposed ratio estimator is more efficient than the Singh-Priyanka estimator for all values of the parameters. Actually, this is an expected result since condition (17) is satisfied as follows:

$$\frac{1}{2}\frac{C_z}{C_y} = 0.5418 < \rho_{yz} = 0.935.$$

Table 2. The relative efficiencies for $\rho_{yx} = 0.97$ and different values of W, k, μ and ρ_{yz} taken as in Singh and Priyanka [13]

k		1.5					2		
μ		0.1	0.3	0.5	0.78	0.97	0.1	0.3	0.5
W	ρ_{yz}	RE	RE	RE	RE	RE	RE	RE	RE
0.2	0.935	1.338	1.981	2.605	3.586	5.414	1.310	1.906	2.489
	0.850	1.118	1.342	1.560	1.900	2.540	1.108	1.316	1.519
	0.985	2.429	5.156	7.795	11.951	19.695	2.313	4.837	7.305
0.4	0.935	1.310	1.906	2.489	3.424	5.283	1.267	1.785	2.301
	0.850	1.108	1.316	1.519	1.846	2.494	1.093	1.274	1.454
	0.985	2.313	4.837	7.305	11.269	19.138	2.130	4.325	6.509
0.6	0.935	1.287	1.841	2.389	3.283	9.159	1.234	1.693	2.155
	0.850	1.100	1.293	1.484	1.796	2.450	1.082	1.242	1.403
	0.985	2.215	4.563	6.880	10.667	18.163	1.991	3.934	5.892
0.8	0.935	1.267	1.785	2.301	3.156	5.042	1.209	1.620	2.039
	0.850	1.093	1.274	1.454	1.752	2.410	1.073	1.216	1.362
	0.985	2.130	4.325	6.509	10.132	18.117	1.883	3.626	5.399

Table 2. (Continued)

k		2		2.5				
μ		0.78	0.97	0.1	0.3	0.5	0.78	0.97
W	ρ_{yz}	RE	RE	RE	RE	RE	RE	RE
0.2	0.935	3.425	5.283	1.287	1.841	2.389	3.282	5.159
	0.850	1.846	2.494	1.100	1.293	1.484	1.796	2.450
	0.985	11.269	19.138	2.215	4.563	6.880	10.667	18.613
0.4	0.935	3.156	5.042	1.234	1.693	2.155	2.941	4.827
	0.850	1.752	2.410	1.082	1.242	1.403	1.677	2.335
	0.985	10.132	18.117	1.991	3.934	5.892	9.221	17.205
0.6	0.935	2.941	4.827	1.198	1.589	1.989	2.689	4.543
	0.850	1.677	2.335	1.069	1.205	1.345	1.589	2.235
	0.985	9.221	17.205	1.837	3.494	5.188	8.151	16.006
0.8	0.935	2.765	4.633	1.171	1.512	1.864	2.494	4.300
	0.850	1.616	2.267	1.060	1.179	1.302	1.521	2.151
	0.985	8.475	16.386	1.725	3.169	4.661	7.328	14.972

Table 3. The relative efficiencies for $\rho_{yx} = 0.80$ and different values of W, k, μ and ρ_{yz} taken as in Singh and Priyanka [13]

k		1.5					2		
μ		0.1	0.3	0.5	0.78	0.97	0.1	0.3	0.5
W	ρ_{yz}	RE	RE	RE	RE	RE	RE	RE	RE
	0.935	1.354	2.131	3.042	4.720	6.424	1.325	2.047	2.909
0.2	0.850	1.123	1.394	1.712	2.298	2.892	1.113	1.365	1.666
	0.985	2.497	5.790	9.646	16.754	23.970	2.376	5.432	9.082
	0.935	1.325	2.047	2.909	4.568	6.382	1.280	1.911	2.689
0.4	0.850	1.113	1.365	1.666	2.244	2.877	1.098	1.318	1.589
	0.985	2.376	5.432	9.082	16.108	23.790	2.184	4.856	8.151
	0.935	1.301	1.974	2.792	4.427	6.340	1.245	1.806	2.514
0.6	0.850	1.105	1.340	1.625	2.195	2.860	1.086	1.281	1.528
	0.985	2.273	5.124	8.588	15.513	23.610	2.039	4.413	7.412
	0.935	1.280	1.911	2.689	4.297	6.299	1.219	1.723	2.372
0.8	0.850	1.098	1.318	1.589	2.150	2.848	1.076	1.252	1.479
	0.985	2.184	4.856	8.151	14.963	23.440	1.926	4.061	6.811

Table 3. (Continued)

k		2		2.5				
μ		0.78	0.97	0.1	0.3	0.5	0.78	0.97
W	ρ_{yz}	RE	RE	RE	RE	RE	RE	RE
	0.935	4.568	6.382	1.301	1.974	2.792	4.427	6.340
0.2	0.850	2.244	2.877	1.105	1.340	1.625	2.195	2.862
	0.985	16.108	23.790	2.273	5.124	8.588	15.513	23.614
	0.935	4.297	6.299	1.245	1.806	2.514	4.065	6.219
0.4	0.850	2.150	2.848	1.086	1.281	1.528	2.070	2.820
	0.985	14.963	23.440	2.039	4.413	7.412	13.978	23.100
	0.935	4.060	6.219	1.207	1.687	2.311	3.770	6.103
0.6	0.850	2.069	2.820	1.072	1.240	1.457	1.967	2.780
	0.985	13.978	23.100	1.878	3.911	6.551	12.738	22.610
	0.935	3.863	6.141	1.180	1.599	2.156	3.530	5.992
0.8	0.850	1.990	2.793	1.063	1.209	1.403	1.882	2.741
	0.985	13.124	22.770	1.760	3.537	5.894	11.713	22.138

From Tables 2 and 3, it is also observed that the proposed estimator has the highest relative efficiency for the values $W = 0.2$, $k = 1.5$, $\mu = 0.97$. From this numerical illustration, we find that there are 4 cases increasing the efficiency of the proposed estimator:

1. When the fraction of missing data decreases,
2. When the fraction of fresh samples increases,
3. When the fraction of sub-sampling decreases,
4. When the correlation between y and z increases.

As the Singh-Priyanka estimator depends on the value of ρ_{yx}, we have examined the MSE values for different values of ρ_{yx}. We see that the efficiency of the proposed estimator is also increasing while ρ_{yx} is decreasing. Hence, it is shown that in the presence of non-response, the proposed estimator performs better than the Singh-Priyanka estimator.

5. Conclusion

We have derived a new ratio-type estimator in the presence of non-response in successive sampling. We have obtained its MSE equation and investigated the optimal replacement policy. The MSE of the proposed estimator has been compared with the Singh-Priyanka estimator in theory and the efficiency condition for the proposed estimator has been found. This theoretical condition has also been satisfied by a numerical example for various values of parameters. As a result, we can infer that the proposed estimator is more efficient than the the Singh-Priyanka estimator under the condition in (17).

References

[1] Chaturvedi, D. K. and Tripathi, T. P. *Estimation of population ratio on two occasions using multivariate auxiliary information*, J. Ind. Statistics Assoc. **21**, 113–120, 1983.

[2] Cochran, W. G. *Sampling Techniques* (John Wiley and Sons, New York, 1977).

[3] Hansen, M. H. and Hurwitz, W. N. *The problem of nonresponse in sample surveys*, American Statistical Association **41**, 517–529, 1946.

[4] Jessen, R. J. *Statistical investigation of a sample survey for obtaining farm facts*, Iowa Agricultural Experiment Station Road Bulletin **N 304**, 1–104, 1942.

[5] Patterson, H. D. *Sampling on successive occasions with partial replacement of units*, J. Roy. Statist. Assoc. **B 12**, 241–255, 1950.

[6] Satıcı, E. *Estimation of Population Mean in the Case of Missing Observations* (PhD. Thesis in Turkish, pp.100, Faculty of Science, Department of Statistics, Hacettepe University, 2009).

[7] Sen, A. R. *Successive sampling with two auxiliary variables*, Sankhya **B 33**, 371–378, 1971.

[8] Sen, A. R. *Successive sampling with* p ($p \leq 1$) *auxiliary variables*, Ann. Math. Statist. **43**, 2031–2034, 1972.

[9] Sen, A. R. *Theory and application of sampling on repated occasions with several auxiliary variables*, Biometrics **29**, 381–385, 1973.

[10] Singh, G. N. *On the use of chain-type ratio estimator in successive sampling*, Statistics in Transition **7** (1), 21–26, 2005.

[11] Singh, G. N. and Singh V. K. *On the use of auxiliary information in successive sampling*, Journal of the Indian Society Agricultural Statistics **54** (1), 1–12, 2001.

[12] Singh, V. K., Singh, G. N. and Shukla, D. *An efficient family of ratio-cum-difference type estimators in successive sampling over two occasions*, J. Sci. Res. **41 C**, 149–159, 1991.

[13] Singh, G. N. and Priyanka, K. *Effect of non-response on current occasion in search of good rotation patterns on successive occasions*, Statistics in Transition **8** (2), 273–292, 2007.

[14] Singh, G. N. and Priyanka, K. *Search of good rotation patterns to improve the precision of estimates at current occasion*, Communications in Statistics: Theory and Methods **37**, 337–348, 2008.