



Thailand Statistician
January 2024; 22(1): 40-49
<http://statassoc.or.th>
Contributed paper

A General Class of Estimators in Presence of Non-response and Measurement error under Two-Phase Successive Sampling

Sunil Kumar* and Monica Choudhary

Department of Statistics, University of Jammu, Jammu, J&K, India

* Corresponding author; Email: sunilbhoulgal06@gmail.com

Received: 25 November 2021

Revised: 10 April 2022

Accepted: 13 April 2022

Abstract

In this paper, we consider a general class of estimators for estimating the population mean in two-phase successive sampling in the presence of non-response and measurement errors. The expression of the biases and mean squared errors (MSE's) have been derived under the first degree of approximation. To check the performance, the proposed estimators are compared with the estimators under the complete response situation. A simulation study was carried out to evaluate the results. On the basis of the available results, appropriate recommendations have been made to the researchers.

Keywords: Study Variable, auxiliary variable, bias, mean square error, percent relative efficiency

1. Introduction

In sample surveys, it is assumed that all the information is available, i.e., a complete response was observed from all the selected units in the sample, which is actually not correct. The reason for incomplete responses may be due to non-response, which is a very common problem in most of the surveys. To tackle this problem, Hansen and Hurwitz (1946) suggested a technique by simply drawing a sub-sample from the non-respondents group and then estimating it. Further, this technique was extended by Cochran (1977), Khare and Srivastava (1997), Singh and Kumar (2008), Singh and Kumar (2009b), Singh and Kumar (2010), Kumar et al. (2011), Chaudhary et al. (2015) and among others. Besides non-response, measurement error is also a serious concern because of incomplete information in sample surveys. Measurement error occurs when the information provided by the respondents is not true, and it also includes observational error, instrumental error, respondent error, etc. Various researchers studied measurement error like Fuller (1995), Wang (2002), Singh and Karpe (2009), Sharma and Singh (2013), etc. The problem of non-response and measurement errors may creep into surveys at the same time. If these errors are small and negligible, then they can be ignored, but if they are not negligible, then they create a problem. A very limited number of studies are available where the interaction of both non-response and measurement errors are studied jointly. Azeem (2014), Kumar et al. (2018) proposed a class of estimators for estimating the population mean in the presence of both non-response and measurement errors in case of stratified random sampling by utilizing two auxiliary variables which are highly correlated with the variables under study. Furthermore, Azeem and Hanif (2017) and Zahid and Shabbir (2018) have also studied the effects of both non-response and measurement errors on estimating the population mean. Kumar and Sharma (2020) also studied the joint effects of non-response and measurement errors using a double sampling

scheme. In the present paper, we study the joint effects of non-response and measurement errors of population mean in two-phase successive sampling.

2. Sampling and Notations Used

Consider a finite population of size N , which has been sampled over two occasions. Let x and y be the study variables on the first and second occasions, respectively, with non-response and measurement errors assumed on both occasions. Let z be an auxiliary variable with unknown population mean and has positive correlation with variables $x(y)$ on the first (second) occasions, respectively. To estimate the population mean of auxiliary variable z , a preliminary simple random sample without replacement (SRSWOR) $S_{n'}$ of n' units is drawn on the first occasion and information on z is collected. Further, a second-phase sample of size n ($n < n'$) is drawn from the first phase (preliminary) sample units by the method of SRSWOR and the information on the study variable x is collected. A random sub sample of size m is retained (matched) from the second-phase sample selected on the first occasion for its use on the current occasion. A preliminary sample of size u' is drawn from the non-sampled units of the population by the method of SRSWOR and information on z is collected to estimate the population mean of auxiliary variable z on the current occasion. Then, a second-phase sample of size $u = (n - m) = nq(u < u')$ ($u = n$) is drawn from the first phase sample and the information on study variable y is collected. Complete response studies rarely available in practice, many survey studies suffers from the presence of non-response and measurement errors simultaneously. For non-response, dividing the population into two classes, in the first class those who respond and in the second class those who do not respond and the size of these two classes are N_1 and N_2 , respectively. Again from sample units n_1 units respond and n_2 units do not respond. Similarly, from unmatched and matched portion u_1 and m_1 units respond and u_2 and m_2 units do not respond, respectively. Again a sub-sample of size $n_{h2} (= \frac{n_2}{k}, k > 1)$ units are drawn from the non-respondents class of the sample units, where k is inverse sampling ratio. Similarly, $u_{h2} (= \frac{u_2}{k}, k > 1)$ and $m_{h2} (= \frac{m_2}{k}, k > 1)$ units are drawn from the non-respondents class of the unmatched and matched portion of the sample, respectively. Also, we consider the presence of measurement error along with the non-response associated with the sample units i.e. $V_i = x_i - X_i$, $U_i = y_i - Y_i$; for both occasions, which are random in nature with mean zero and population variance S_U^2 and S_V^2 . Here we use the following notations

$x(y)$: the study variables on the first (second) occasion.

μ_X (μ_Y): the population mean of study variables on the first (second) occasion.

μ_Z : the population mean of auxiliary variable z .

μ_{ym} , μ_{yu} , μ_{xm} , μ_{xn} , μ_{zn} and μ_{zu} : the sample means of the variables y , x and z , respectively based on respective sample size shown in their subscripts.

$\mu_{zn'}$ ($\mu_{zu'}$): the sample means of auxiliary variable based on sample of size n' and u' , respectively.

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

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

$\hat{\mu}_{xm}^*$, $\hat{\mu}_{xn}^*$ ($\hat{\mu}_{ym}^*$, $\hat{\mu}_{yn}^*$): the sample means of the respective variables based on non-response and measurement errors in the second phase sample of size u , m and n on first (second) occasions, respectively.

3. The Proposed Class of Estimators

Inspired by G. N. Singh and et al. (2017), a class of estimators, say $\hat{\xi}^*$ of the population mean μ_Y of the study variable y on the current (second) occasion when the population mean of auxiliary variable z is unknown has been proposed. The proposed class of estimators $\hat{\xi}^*$ is a convex linear combination of two distinct classes of estimators $\hat{\xi}_u^*$ and $\hat{\xi}_m^*$ are defined as

$$\hat{\xi}^* = \phi \hat{\xi}_u^* + (1 - \phi) \hat{\xi}_m^* \quad (1)$$

where $\phi(0 \leq \phi \leq 1)$ is an unknown scalar (constant) to be estimated by the minimization of the mean square error of the estimator $\hat{\xi}_u^*$.

The class of estimators $\hat{\xi}_u^*$ is based on the current sample S_u of size u drawn on the current occasion, and the class of estimators $\hat{\xi}_m^*$ is based on the sample S_m of size m drawn on matched portion of both occasions.

A class of estimators $\hat{\xi}_u^*$, which is based on unmatched sample of size u units drawn on current occasion as

$$\hat{\xi}_u^* = F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'}) \tag{2}$$

We consider the composite function $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})$ as one to one function of $\hat{\mu}_{yu}^*, \mu_{zu}$ and $\mu_{zu'}$ such that

$$F(\mu_Y, \mu_Z, \mu_Z) = \mu_Y \Rightarrow \left. \frac{\partial F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})}{\partial \hat{\mu}_{yu}^*} \right|_{(\mu_Y, \mu_Z, \mu_Z)} = 1 \tag{3}$$

and satisfy the following regularity conditions

- (i) Anything chosen in samples, $(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})$ assume values in closed convex subspace R^3 of three-dimensional real space containing the point (μ_Y, μ_Z, μ_Z) .
- (ii) The function $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})$ is continuous and bounded in R^3 .
- (iii) The first and second partial derivative of $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})$ exist, continuous and bounded in R^3 .

The estimator $\hat{\xi}_u^*$ is very wide in the sense for any parametric function, $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu'})$ satisfying above regularity conditions with $F(\mu_Y, \mu_Z, \mu_Z) = \mu_Y$ to give an estimator of μ_Y . Such as, the following ratio, product and exponential type estimators of μ_Y are the members of the classes $\hat{\xi}_u^*$.

$$\begin{aligned} t_1 &= \hat{\mu}_{yu}^* \frac{\mu_{zu'}}{\mu_{zu}}, t_2 = \hat{\mu}_{yu}^* \frac{\mu_{zu}}{\mu_{zu'}}, t_3 = \hat{\mu}_{yu}^* + \alpha_1(\mu_{zu'} - \mu_{zu}), t_4 = \hat{\mu}_{yu}^* \exp\left(\frac{(\mu_{zu'} - \mu_{zu})}{(\mu_{zu'} + \mu_{zu})}\right), \\ t_5 &= \hat{\mu}_{yu}^* \exp\left(\frac{(\mu_{zu} - \mu_{zu'})}{(\mu_{zu} + \mu_{zu'})}\right), t_6 = \hat{\mu}_{yu}^* \left(\frac{\mu_{zu}}{\mu_{zu'}}\right)^{\alpha_2}, t_7 = \hat{\mu}_{yu}^* \left[2 - \left(\frac{\mu_{zu}}{\mu_{zu'}}\right)^{\alpha_3}\right], \\ t_8 &= \hat{\mu}_{yu}^* \left[\frac{\mu_{zu'}}{\mu_{zu'} + \alpha_4(\mu_{zu} - \mu_{zu'})}\right], t_9 = \hat{\mu}_{yu}^* \left[\alpha_5 \frac{\mu_{zu'}}{\mu_{zu}} + (1 - \alpha_5) \left(\frac{\mu_{zu'}}{\mu_{zu}}\right)^2\right], \\ t_{10} &= \hat{\mu}_{yu}^* \left[\alpha_6 \frac{\mu_{zu}}{\mu_{zu'}} + (1 - \alpha_6) \left(\frac{\mu_{zu}}{\mu_{zu'}}\right)^2\right], t_{11} = \hat{\mu}_{yu}^* \left[\alpha_7 + (1 - \alpha_7) \frac{\mu_{zu}}{\mu_{zu'}}\right], \\ t_{12} &= \hat{\mu}_{yu}^* \left[\alpha_8 + (1 - \alpha_8) \frac{\mu_{zu}}{\mu_{zu'}}\right], t_{13} = \hat{\mu}_{yu}^* \left[\frac{\alpha_9 \mu_{zu'} + (1 - \alpha_9) \mu_{zu}}{\alpha_9 \mu_{zu} + (1 - \alpha_9) \mu_{zu'}}\right]. \end{aligned}$$

where $\alpha_i (i = 1, 2, 3, \dots, 9)$ are suitably chosen constants so that the mean square errors of the above estimators may be minimum.

Similarly, for matched portion on current occasion, we propose another class of estimators $\hat{\xi}_m^*$

$$\hat{\xi}_m^* = G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn'}) \tag{4}$$

For any chosen sample, $\hat{\xi}_m^* = G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn'})$ assume values in closed, bounded convex subspace R^5 of five-dimensional real space containing the point $(\mu_Y, \mu_X, \mu_X, \mu_Z, \mu_Z)$, where $G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn'})$ is a composite function $(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn'})$, such that

$$G(\mu_Y, \mu_X, \mu_X, \mu_Z, \mu_Z) = \mu_Y \Rightarrow \left. \frac{\partial G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn'})}{\partial \hat{\mu}_{ym}^*} \right|_{(\mu_Y, \mu_X, \mu_X, \mu_Z, \mu_Z)} = 1 \tag{5}$$

and satisfy the following regularity conditions that are similar to those given for function $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')$ in Eqn. (3). Following are the members of the classes $\hat{\xi}_m^*$ as

$$\begin{aligned} s_1 &= \hat{\mu}_{ym}^* \frac{\hat{\mu}_{xn}^* \mu_{zn}'}{\hat{\mu}_{xm}^* \mu_{zn}}, s_2 = \hat{\mu}_{ym}^* \frac{\hat{\mu}_{xn}^* \mu_{zn}}{\hat{\mu}_{xm}^* \mu_{zn}'}, s_3 = \hat{\mu}_{ym}^* \frac{\hat{\mu}_{xm}^* \mu_{zn}'}{\hat{\mu}_{xn}^* \mu_{zn}}, \\ s_4 &= \left[\hat{\mu}_{ym}^* + b_1(\hat{\mu}_{xn}^* - \hat{\mu}_{xm}^*) \right] \frac{\mu_{zn}'}{\mu_{zn}}, s_5 = \hat{\mu}_{ym}^* \exp\left(\frac{\hat{\mu}_{xn}^* - \hat{\mu}_{xm}^*}{\hat{\mu}_{xn}^* + \hat{\mu}_{xm}^*}\right) \frac{\mu_{zn}'}{\mu_{zn}}, \\ s_6 &= \left[\hat{\mu}_{ym}^* + b_2(\hat{\mu}_{xn}^* - \hat{\mu}_{xm}^*) \right] \exp\left(\frac{\mu_{zn}' - \mu_{zn}}{\mu_{zn}' + \mu_{zn}}\right), s_7 = \hat{\mu}_{ym}^* \exp\left(\frac{\hat{\mu}_{xn}^* - \hat{\mu}_{xm}^*}{\hat{\mu}_{xn}^* + \hat{\mu}_{xm}^*}\right) \exp\left(\frac{\mu_{zn}' - \mu_{zn}}{\mu_{zn}' + \mu_{zn}}\right), \\ s_8 &= \frac{\hat{\mu}_{ym}^*}{\hat{\mu}_{xm}^*} \left[\hat{\mu}_{xn}^* + b_3(\mu_{zn}' - \mu_{zn}) \right], s_9 = \hat{\mu}_{ym}^* + b_4 \left[\frac{\hat{\mu}_{xn}^*}{\mu_{zn}} \mu_{zn}' + \hat{\mu}_{xm}^* \right], \\ s_{10} &= \hat{\mu}_{ym}^* + b_5 \left[(\hat{\mu}_{xn}^* - \hat{\mu}_{xm}^*) + b_4(\mu_{zn} - \mu_{zn}') \right], s_{11} = \hat{\mu}_{ym}^* \frac{\hat{\mu}_{xm}^*}{[\hat{\mu}_{xn}^* + b_6(\mu_{zn}' - \mu_{zn})]}, \\ s_{12} &= \hat{\mu}_{ym}^* + b_7 \left[\hat{\mu}_{xn}^* \frac{\hat{\mu}_{zn}}{\mu_{zn}'} - \hat{\mu}_{xm}^* \right], \text{etc.} \end{aligned}$$

where $b_i (i = 1, 2, 3, \dots, 7)$ are suitably chosen constants.

To obtain the bias and mean square errors of proposed class of estimators $\hat{\xi}_u^*$ and $\hat{\xi}_m^*$ are derived up to first order of approximation such that

$$\begin{aligned} \hat{\mu}_{yu}^* &= \frac{u_1 \mu_{yu_1} + u_2 \mu_{yu_{h2}}}{u}, \hat{\mu}_{ym}^* = \frac{m_1 \mu_{ym_1} + m_2 \mu_{ym_{h2}}}{m}, \\ \hat{\mu}_{xm}^* &= \frac{m_1 \mu_{xm_1} + m_2 \mu_{xm_{h2}}}{m}, \hat{\mu}_{xn}^* = \frac{n_1 \mu_{xn_1} + n_2 \mu_{xn_{h2}}}{n}, \end{aligned}$$

where $\mu_{yu_1} = \frac{1}{u_1} \sum_{i=1}^{u_1} y_i$, $\mu_{ym_1} = \frac{1}{m_1} \sum_{j=1}^{m_1} y_j$, $\mu_{xm_1} = \frac{1}{m_1} \sum_{j=1}^{m_1} x_j$, $\mu_{xn_1} = \frac{1}{n_1} \sum_{k=1}^{n_1} x_k$, $\mu_{yu_{h2}} = \frac{1}{u_{h2}} \sum_{l=1}^{u_{h2}} y_l$, $\mu_{ym_{h2}} = \frac{1}{m_{h2}} \sum_{r=1}^{m_{h2}} y_r$, $\mu_{xm_{h2}} = \frac{1}{m_{h2}} \sum_{r=1}^{m_{h2}} x_r$, $\mu_{xn_{h2}} = \frac{1}{n_{h2}} \sum_{s=1}^{n_{h2}} x_s$.

Again, here

$$\begin{aligned} \hat{\mu}_{yu}^* &= \mu_Y + \omega_{yu}^*, \hat{\mu}_{ym}^* = \mu_Y + \omega_{ym}^*, \hat{\mu}_{xm}^* = \mu_X + \omega_{xm}^*, \hat{\mu}_{xn}^* = \mu_X + \omega_{xn}^*, \mu_{zu} = \mu_Z(1 + \varepsilon_1), \\ \mu_{zu}' &= \mu_Z(1 + \varepsilon_2), \mu_{zn} = \mu_Z(1 + \varepsilon_3), \mu_{zn}' = \mu_Z(1 + \varepsilon_4), \end{aligned}$$

$$\begin{aligned} \omega_{yu}^* &= \frac{1}{\sqrt{u}} (\omega_{yu} + \omega_{Uu}), \omega_{ym}^* = \frac{1}{\sqrt{m}} (\omega_{ym} + \omega_{Um}), \omega_{xm}^* = \frac{1}{\sqrt{m}} (\omega_{xm} + \omega_{Vm}), \omega_{xn}^* = \frac{1}{\sqrt{n}} (\omega_{xn} + \omega_{Vn}), \\ \omega_{yu} &= \frac{1}{\sqrt{u}} \sum_{i=1}^u (Y_i^* - \mu_Y), \omega_{ym} = \frac{1}{\sqrt{m}} \sum_{i=1}^m (Y_i^* - \mu_Y), \omega_{xm} = \frac{1}{\sqrt{m}} \sum_{i=1}^m (X_i^* - \mu_X), \\ \omega_{xn} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i^* - \mu_X), \omega_{Uu} = \frac{1}{\sqrt{u}} \sum_{i=1}^u U_i^*, \omega_{Um} = \frac{1}{\sqrt{m}} \sum_{i=1}^m U_i^*, \omega_{Vm} = \frac{1}{\sqrt{m}} \sum_{i=1}^m V_i^*, \\ \omega_{Vn} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n V_i^*, U_i^* = (y_i - Y_i), V_i^* = (x_i - X_i), \end{aligned}$$

such that

$$\begin{aligned} E(\omega_{yu}^*) &= E(\omega_{ym}^*) = E(\omega_{xm}^*) = E(\omega_{xn}^*) = 0; \\ E(\omega_{yu}^*)^2 &= \mu_Y^2 B^*, E(\omega_{ym}^*)^2 = \mu_Y^2 b^*, E(\omega_{xm}^*)^2 = \mu_X^2 E^*, E(\omega_{xn}^*)^2 = \mu_X^2 e^* = E(\omega_{xm}^* \omega_{xn}^*), \\ E(\omega_{ym}^* \omega_{xm}^*) &= \mu_Y \mu_X D^*, E(\omega_{ym}^* \omega_{xn}^*) = \mu_Y \mu_X d^*. \end{aligned}$$

Also, $E(\varepsilon_i) = 0; i = 1, 2, \dots, 4$.

$$\begin{aligned} E(\varepsilon_1)^2 &= f_2 C_z^2, E(\varepsilon_2)^2 = f_2' C_z^2 = E(\varepsilon_1 \varepsilon_2), E(\varepsilon_3)^2 = f C_z^2, E(\varepsilon_4)^2 = f' C_z^2 = E(\varepsilon_3 \varepsilon_4), \\ E(\omega_{yu}^* \varepsilon_1) &= \frac{1}{\mu_Z} f_2 \rho_{yz} C_y C_z, E(\omega_{yu}^* \varepsilon_2) = \frac{1}{\mu_Z} f_2' \rho_{yz} C_y C_z, \\ E(\omega_{ym}^* \varepsilon_3) &= \frac{1}{\mu_Z} f \rho_{yz} C_y C_z, E(\omega_{ym}^* \varepsilon_4) = \frac{1}{\mu_Z} f' \rho_{yz} C_y C_z, \\ E(\omega_{xm}^* \varepsilon_3) &= \frac{1}{\mu_Z} f \rho_{xz} C_x C_z = E(\omega_{xn}^* \varepsilon_3), \\ E(\omega_{xm}^* \varepsilon_4) &= \frac{1}{\mu_Z} f' \rho_{xz} C_x C_z = E(\omega_{xn}^* \varepsilon_4), \end{aligned}$$

where, $\theta = W_2(k - 1), W_2 = N_2/N,$

$$f_2 = \left(\frac{1}{u} - \frac{1}{N}\right), f_1 = \left(\frac{1}{m} - \frac{1}{N}\right), f = \left(\frac{1}{n} - \frac{1}{N}\right), f'_2 = \left(\frac{1}{u} - \frac{1}{N}\right), f'_1 = \left(\frac{1}{n} - \frac{1}{N}\right),$$

$$B^* = \{f_2(C_y^2 + C_U^2) + \frac{\theta}{u}(C_{y(2)}^2 + C_{U(2)}^2)\}, b^* = \{f_1(C_y^2 + C_U^2) + \frac{\theta}{m}(C_{y(2)}^2 + C_{U(2)}^2)\},$$

$$E^* = \{f_1(C_x^2 + C_V^2) + \frac{\theta}{m}(C_{x(2)}^2 + C_{V(2)}^2)\}, e^* = \{f(C_x^2 + C_V^2) + \frac{\theta}{n}(C_{x(2)}^2 + C_{V(2)}^2)\},$$

$$D^* = \{f_1(\rho_{yx}C_yC_x) + \frac{\theta}{m}(\rho_{yx(2)}C_{y(2)}C_{x(2)})\}, d^* = \{f(\rho_{yx}C_yC_x) + \frac{\theta}{n}(\rho_{yx(2)}C_{y(2)}C_{x(2)})\}.$$

4. Bias and Mean Square Error

The class of estimators $\hat{\xi}_u^*$ express in term of expected values, we expand $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')$ about the point (μ_Y, μ_Z, μ_Z) in third order of Taylor's series expansion and we have

$$F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}') = F(\mu_Y, \mu_Z, \mu_Z) + (\hat{\mu}_{yu}^* - \mu_Y)A_1 + (\mu_{zu} - \mu_Z)A_2 + (\mu_{zu}' - \mu_Z)A_3$$

$$+ \frac{1}{2} \left\{ (\hat{\mu}_{yu}^* - \mu_Y)^2 A_{11} + (\mu_{zu} - \mu_Z)^2 A_{22} + (\mu_{zu}' - \mu_Z)^2 A_{33} + 2(\hat{\mu}_{yu}^* - \mu_Y) \right.$$

$$\left. (\mu_{zu} - \mu_Z)A_{12} + 2(\hat{\mu}_{yu}^* - \mu_Y)(\mu_{zu}' - \mu_Z)A_{13} + 2(\mu_{zu} - \mu_Z)(\mu_{zu}' - \mu_Z)A_{23} \right\}$$

$$+ \frac{1}{6} \left\{ (\hat{\mu}_{yu}^* - \mu_Y) \frac{\partial}{\partial \hat{\mu}_{yu}^*} + (\mu_{zu} - \mu_Z) \frac{\partial}{\partial \mu_{zu}} + (\mu_{zu}' - \mu_Z) \frac{\partial}{\partial \mu_{zu}'} \right\}^3 F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}') + \dots$$

where

$$A_1 = \left. \frac{\partial F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')}{\partial \hat{\mu}_{yu}^*} \right|_{(\mu_Y, \mu_Z, \mu_Z)}, A_2 = \left. \frac{\partial F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')}{\partial \mu_{zu}} \right|_{(\mu_Y, \mu_Z, \mu_Z)}, A_3 = \left. \frac{\partial F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')}{\partial \mu_{zu}'} \right|_{(\mu_Y, \mu_Z, \mu_Z)}$$

and $(A_{11}, A_{22}, A_{33}, A_{12}, A_{13}, A_{23})$ are the second derivative of $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')$ at point (μ_Y, μ_Z, μ_Z) .

Under the mentioned conditions of $F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}')$ in Eqn. (3), it is noted that

$$F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}') = \mu_Y \Rightarrow A_1 = 1 \text{ and } A_{11} = \left. \frac{\partial^2}{\partial (\hat{\mu}_{yu}^*)^2} F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}') \right|_{(\mu_Y, \mu_Z, \mu_Z)} = 0.$$

The population mean μ_Z of the auxiliary variable z is unknown, so we have impose the constraints as

$$A_2 = -A_3.$$

Then, the expression of $\hat{\xi}_u^*$ up to the first order approximation is shown as

$$\hat{\xi}_u^* = F(\hat{\mu}_{yu}^*, \mu_{zu}, \mu_{zu}') = \mu_Y + \omega_{yu}^* A_1 + \mu_Z A_2 (\epsilon_1 - \epsilon_2)$$

$$+ \frac{1}{2} \left\{ \mu_Z^2 (\epsilon_1^2 A_{22} + \epsilon_2^2 A_{33} + 2\epsilon_1 \epsilon_2 A_{23}) + 2\omega_{yu}^* \mu_Z (\epsilon_1 A_{12} + \epsilon_2 A_{13}) \right\} \tag{6}$$

Similarly, the expression of $\hat{\xi}_m^*$ up to the first order approximation are derived as

$$\hat{\xi}_m^* = G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zm}, \mu_{zn}') = \mu_Y + \omega_{ym}^* j_1 + (\omega_{xm}^* - \omega_{xn}^*) j_2 + \mu_Z (\epsilon_3 - \epsilon_4) j_4$$

$$+ \frac{1}{2} \left\{ (\omega_{xm}^*)^2 j_{22} + (\omega_{xn}^*)^2 j_{33} + 2(\omega_{xm}^* \omega_{xn}^*) j_{23} + \mu_Z^2 (\epsilon_3^2 j_{44} + \epsilon_4^2 j_{55} + 2\epsilon_3 \epsilon_4 j_{45}) \right.$$

$$+ 2(\omega_{ym}^* \omega_{xm}^*) j_{12} + 2(\omega_{ym}^* \omega_{xn}^*) j_{13} + 2\mu_Z (\omega_{ym}^* \epsilon_3 j_{14} + \omega_{ym}^* \epsilon_4 j_{15})$$

$$\left. + 2\mu_Z (\omega_{xm}^* \epsilon_3 j_{24} + \omega_{xm}^* \epsilon_4 j_{25} + \omega_{xn}^* \epsilon_3 j_{34} + \omega_{xn}^* \epsilon_4 j_{35}) \right\} \tag{7}$$

where $G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}') = \mu_Y \Rightarrow j_1 = 1$,

$$j_2 = \frac{\partial G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}')}{\partial \hat{\mu}_{xm}^*} \Bigg|_{(\mu_Y, \mu_X, \mu_X \mu_Z, \mu_Z)},$$

$$j_3 = \frac{\partial G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}')}{\partial \hat{\mu}_{xn}^*} \Bigg|_{(\mu_Y, \mu_X, \mu_X \mu_Z, \mu_Z)},$$

$$j_4 = \frac{\partial G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}')}{\partial \mu_{zn}} \Bigg|_{(\mu_Y, \mu_X, \mu_X \mu_Z, \mu_Z)},$$

$$j_5 = \frac{\partial G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}')}{\partial \mu_{zn}'} \Bigg|_{(\mu_Y, \mu_X, \mu_X \mu_Z, \mu_Z)}.$$

Here, we imposed the constraints $j_2 = -j_3$, $j_4 = -j_5$ and $(j_{22}, j_{33}, j_{44}, j_{55}, j_{23}, j_{45}, j_{12}, j_{13}, j_{14}, j_{15}, j_{24}, j_{25}, j_{34}, j_{35})$ are the second-order partial derivatives of $G(\hat{\mu}_{ym}^*, \hat{\mu}_{xm}^*, \hat{\mu}_{xn}^*, \mu_{zn}, \mu_{zn}')$ at the point $(\mu_Y, \mu_X, \mu_X \mu_Z, \mu_Z)$.

Taking expectations on both sides of Eqns. (6) and (7), we obtain the bias and MSE of $(\hat{\xi}^*)$ upto the first order of approximation as

$$B(\hat{\xi}^*) = \phi B(\hat{\xi}_u^*) + (1 - \phi) B(\hat{\xi}_m^*)$$

$$MSE(\hat{\xi}^*) = \phi^2 MSE(\hat{\xi}_u^*) + (1 - \phi)^2 MSE(\hat{\xi}_m^*)$$

where $B(\hat{\xi}_u^*) = E(\hat{\xi}_u^* - \mu_Y) = \frac{1}{2} \mu_Z^2 C_z^2 \{f_2 A_{22} + f_2'(A_{33} + 2A_{23})\} + \rho_{yz} C_y C_z \{f_2 A_{12} + f_2' A_{13}\}$.

$$B(\hat{\xi}_m^*) = E(\hat{\xi}_m^* - \mu_Y) = \frac{1}{2} [\mu_x^2 \{E^* j_{22} + e^* j_{33} + 2e^* j_{23}\} + \mu_Z^2 C_z^2 \{f j_{44} + f'(j_{55} + 2j_{45})\}]$$

$$+ \mu_Y \mu_X \{D^* j_{12} + d^* j_{13}\} + \rho_{yz} C_y C_z \{f j_{14} + f' j_{15}\}$$

$$+ \rho_{xz} C_x C_z \{f(j_{24} + j_{34}) + f'(j_{25} + j_{35})\}$$

$$MSE(\hat{\xi}_u^*) = E(\hat{\xi}_u^* - \mu_Y)^2 = \mu_Y^2 A_1^2 B^* + \mu_Z^2 A_2^2 (f_2 - f_2') C_z^2 + 2A_1 A_2 (f_2 - f_2') \rho_{yz} C_y C_z \quad (8)$$

$$MSE(\hat{\xi}_m^*) = E(\hat{\xi}_m^* - \mu_Y)^2 = \mu_Y^2 j_1^2 b^* + \mu_X^2 j_2^2 (E^* - e^*) + \mu_Z^2 j_4^2 (f - f') C_z^2$$

$$+ 2j_1 j_2 \mu_Y \mu_X (D^* - d^*) + 2j_1 j_4 (f - f') \rho_{yz} C_y C_z \quad (9)$$

The estimator $\hat{\xi}_u^*$ and $\hat{\xi}_m^*$ are based on two independent samples of size u and m respectively, therefore the covariance has been ignored, i.e. $C(\hat{\xi}_u^*, \hat{\xi}_m^*) = 0$.

To obtain the minimum MSE of $\hat{\xi}_u^*$ and $\hat{\xi}_m^*$, we differentiate Eqns. (8) and (9) with respect to A_2 , j_2 and j_4 and equating to zero, respectively, we get

$$A_2 = -\rho_{yz} \frac{C_y}{C_z} \frac{1}{\mu_Z^2} = j_4 \text{ and } j_2 = -\frac{\mu_Y (D^* - d^*)}{\mu_X (E^* - e^*)}.$$

Substitute the value of A_2 , j_2 and j_4 in Eqns. (8) and (9), then we get the minimum mean square error of the proposed estimators $\hat{\xi}^*$ as

$$MSE(\hat{\xi}^*)_{min} = \phi^2 MSE(\hat{\xi}_u^*)_{min} + (1 - \phi)^2 MSE(\hat{\xi}_m^*)_{min} \quad (10)$$

where

$$MSE(\hat{\xi}_u^*)_{min} = \mu_Y^2 B^* - \frac{1}{\mu_Z^2} \rho_{yz}^2 (f_2 - f_2') C_y^2 \text{ and } MSE(\hat{\xi}_m^*)_{min} = \mu_Y^2 b^* - \frac{\mu_Y^2 (D^* - d^*)^2}{(E^* - e^*)} - \frac{1}{\mu_Z^2} \rho_{yz}^2 (f - f') C_y^2.$$

The minimum MSE of $\hat{\xi}^*$ in (10) is depending on unknown constant ' ϕ ', then we differentiate it w.r.t ' ϕ ' and equating to zero, we obtained the optimum value of ' ϕ ' as

$$\phi_{opt} = \frac{MSE(\hat{\xi}_m^*)_{min}}{MSE(\hat{\xi}_u^*)_{min} + MSE(\hat{\xi}_m^*)_{min}}.$$

After substituting the optimum value of ϕ_{opt} in (10), we get the optimum mean square error of the class of estimators $\hat{\xi}^*$ as

$$MSE(\hat{\xi}^*)_{opt} = \frac{MSE(\hat{\xi}_u^*)_{min} \cdot MSE(\hat{\xi}_m^*)_{min}}{MSE(\hat{\xi}_u^*)_{min} + MSE(\hat{\xi}_m^*)_{min}}.$$

5. Efficiency Comparisons

To check the performance of the proposed class of estimators $\hat{\xi}^*$ in the presence of non-response and measurement errors, we compare it with complete response situations. The percent relative efficiency of the proposed estimators are calculated with respect to complete response estimator τ , which is defined as

$$\tau = \psi \tau_u + (1 - \psi) \tau_m,$$

where $\tau_u = \mu_{yu}$, $\tau_m = \frac{\mu_{ym}}{\mu_{xm}} \mu_{xn}$ and $\psi (0 \leq \psi \leq 1)$ is an unknown scalar to be determined by the minimization of mean square error of the estimator τ .

The minimum mean square error of the estimator τ up to the first order approximations is obtained as

$$MSE(\tau)_{min} = \frac{V(\tau_u) \cdot MSE(\tau_m)}{V(\tau_u) + MSE(\tau_m)},$$

where $V(\tau_u) = \mu_y^2 f_2 C_y^2$, $MSE(\tau_m) = \mu_y^2 [f_1 G^* - f H^*]$, $G^* = (C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x)$ and $H^* = (C_x^2 - 2\rho_{yx} C_y C_x)$.

Thus, the percent relative efficiency E , of the proposed class of estimators $\hat{\xi}^*$ with respect to the estimator τ are given as

$$PRE(E) = \left[\frac{MSE(\tau)_{min}}{MSE(\hat{\xi}^*)_{opt}} \right] \times 100.$$

6. Simulation Study

For numerical analysis, we have simulated the theoretical result using R software. For this we generate an artificial population of size $N = 2000$ from normal distribution by $Z = rnorm(N, 15, 8)$, $Y = 1 + 3*Z + rnorm(N, 0, 1)$ and $X = 1 + 2.5*Z + rnorm(N, 0, 1)$. Further on first occasion, a random sample of 700 units is selected for first phase and then another random sample of 450 units is selected from first phase units. Again on second occasion, a random sample of 550 units for first phase and then we take different sample size from first phase units on second occasion and similarly, we take different sample size for matched portion. Also, there is a measurement error having mean 0 and standard deviation 3 employed. The percent relative efficiency (PREs) of proposed estimators with respect to complete response for different values of m, u, q and k . The results are shown in Tables 1 and 2, respectively.

It is envisaged from Table 1 that for $m > u$, the value of PRE decreases with the increases in the value of ' k '. Also, the PRE decreases with the increases in the size of matched units ' m ' and unmatched units ' u ' with $m > u$. For $m > u$, the proposed estimator ' $\hat{\xi}^*$ ' performs better in case of higher non-response rate. Similar pattern of the performance of the proposed estimator ' $\hat{\xi}^*$ ' in term of having PRE is found for different values of ' q ' and ' k ' in Table 2 as Table 1 under the condition $m < u$. But for $m < u$, the PRE's are less as compared to case of $m > u$.

Table 1 PRE (E) of estimator $\hat{\xi}^*$ with respect to estimator τ , when $m > u$

$q = 0.3$					
m	u	k=2	k=3	k=4	k=5
10	5	137.1248	129.3899	122.5450	116.3239
11	7	131.6873	124.2591	117.6856	111.7110
12	9	128.0486	120.8257	114.4340	108.6239
13	11	125.4397	118.3641	112.1028	106.4103
$q = 0.5$					
m	u	k=2	k=3	k=4	k=5
10	5	137.0650	129.3284	122.4771	116.2279
11	7	131.6216	124.1923	117.6131	111.6119
12	9	127.9778	120.7541	114.3572	108.5214
13	11	125.3641	118.2880	112.0218	106.3042
$q = 0.7$					
m	u	k=2	k=3	k=4	k=5
10	5	137.1110	129.3843	122.5536	116.3236
11	7	131.6713	124.2512	117.6917	111.7085
12	9	127.0309	120.8161	114.4383	108.6195
13	11	125.4205	118.3529	112.1058	106.4042
$q = 0.9$					
m	u	k=2	k=3	k=4	k=5
10	5	137.0227	129.2977	122.4401	116.3315
11	7	131.5842	124.1657	117.5804	111.7145
12	9	127.9433	120.7300	114.3270	108.6241
13	11	125.3315	118.2654	111.9932	106.4076

Table 2 PRE (E) of estimator $\hat{\xi}^*$ with respect to estimator τ , when $m < u$

$q = 0.3$					
m	u	k=2	k=3	k=4	k=5
6	9	114.6629	108.1936	102.4649	97.2825
7	11	113.8522	107.4288	101.7409	96.5937
8	13	113.2726	106.8820	101.2235	96.1010
9	15	112.8352	106.4694	100.8331	95.7290
$q = 0.5$					
m	u	k=2	k=3	k=4	k=5
6	9	114.6545	108.1818	102.4478	97.2449
7	11	113.8378	107.4113	101.7183	96.5507
8	13	113.2525	106.8592	101.7183	96.0528
9	15	112.8097	106.4414	100.8002	95.6757
$q = 0.7$					
m	u	k=2	k=3	k=4	k=5
6	9	114.6620	108.1965	102.4698	97.2939
7	11	113.8496	107.4302	101.7451	96.6034
8	13	113.2683	106.8820	101.2270	96.1090
9	15	112.8294	106.4681	100.8361	95.7354
$q = 0.9$					
m	u	k=2	k=3	k=4	k=5
6	9	114.6380	108.1780	102.4368	97.3039
7	11	113.8213	107.4071	101.7071	96.6124
8	13	113.2359	106.8544	101.1840	96.1172
9	15	112.7928	106.4361	100.7882	95.7428

Further, it is observed from Table 2 that the PRE's are less efficient for $k = 5$ with respect to complete response estimator ' τ '.

7. Conclusion

In the present paper, we have suggested a general class of estimators for estimating the population mean of the study variables in successive sampling when the population mean of auxiliary variable is not known in current occasion. The properties of the proposed estimator ' ξ^* ' has been studied to the first degree of approximation. We have compared the efficiency of the proposed estimator with respect to the complete response estimator. Simulation study also carried out to support the theoretical results, which shows that the proposed estimator ' ξ^* ' is highly efficient for higher response rates. It is also observed that the proposed estimator out performed in terms of PRE when $m > u$ and smaller values of ' q ' as compared to $m < u$. So, we conclude the presence of non-response and measurement errors are heavily affecting estimators and the proposed general class of estimators are recommended for use when non-response and measurement errors occur in successive sampling surveys.

Acknowledgements

We would like to thank the referees for their comments and suggestions on the manuscript which improved the quality of the paper.

References

- Azeem M. On Estimation of population mean in the presence of measurement error and non-response. Lahore: Unpublished Ph.D. thesis. National College of Business Administration and Economics; 2014.
- Azeem M, Hanif M. Joint influence of measurement error and non-response on estimation of population mean. *Commun Stat Theory Methods*. 2017; 46(4): 1679-1693.
- Chaudhary MK, Kumar A. Estimating the population mean in stratified sampling using two phase sampling in the presence of non-response. *World Appl Sci J*. 2015; 33(6): 874882.
- Cochran WG. *Sampling Techniques*. New York: John Wiley, 3rd edition; 1977.
- Fuller WA. Estimation in the presence of measurement error. *Int Stat Rev*. 1995; 63(2): 121-141.
- Hansen MH, Hurwitz WN. The problem of the non-response in sample surveys. *J Am Stat Assoc*. 1946; 41: 517-529.
- Khare BB, Srivastava S. Transformed ratio type estimators for the population mean in the presence of non-response. *Commun Stat Theory Methods*. 1997; 26(7): 1779-1791.
- Kumar S, Singh H, Bhougal S, Gupta R. A class of ratio cum product type estimators under double sampling in the presence of non-response. *Hacettepe J Math Stat*. 2011; 40: 589-599.
- Kumar S, Sharma V. Study on non-response and measurement error using double sampling scheme. *J Stat Appl Probab*. 2020; 7(2): 87-96.
- Kumar S, Trehan M, Singh Joorel JP. A simulation study: estimation of population mean using two auxiliary variables in stratified random sampling. *J Stat Comput Simul*. 2018; 88(1): 1-14.
- Singh HP, Kumar S. A regression approach to the estimation of the finite population mean in the presence of nonresponse. *Aust N Z J Stat*. 2008; 50(4): 395-408.
- Singh HP, Kumar S. A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information. *SORT*. 2009b; 33: 7184.
- Singh HP, Kumar S. Estimation of mean in presence of non-response using two phase sampling scheme. *Stat Pap*. 2010; 50: 559582.
- Singh HP, Karpe N. On the estimation of ratio and product of two population means using supplementary information in presence of measurement errors. *Statistica*. 2009; 69: 27-47.

- Singh GN, Khalid M, Sharma AK. Some efficient classes of estimators of population mean in two-phase successive sampling under random non response. *Commun Stat Theory Methods*. 2017; 46(24): 12194-12209.
- Sharma P, Singh R. A generalized class of estimators for finite population variance in presence of measurement errors. *J Mod Appl Stat Methods*. 2013; 12: 231-241.
- Wang L. A simple adjustment for measurement errors in some dependent variable models. *Stat Probab Lett*. 2002; 58: 427-433.
- Zahid E, Shabbir J. Estimation of population mean in the presence of measurement error and non-response under stratified random sampling. *PLOS ONE*. 2018; 13(2): 1-12.