MEMORY-TYPE MEAN ESTIMATORS FOR SURVEYS WITH NON-RESPONSE: A REGRESSION-IMPUTATION AND EWMA APPROACH.

ABSTRACT

Sampling survey mostly faces the challenge of missing information or non-response, Missing data naturally occurs in sample surveys when a few sampling units refuse to respond or are unable to participate in the survey. In time based survey, this could happen due to relocation or system malfunctioning after some period. This missing information can bring about a substantial amount of bias, make the handling and analysis of the data more arduous and create reductions in efficiency of the estimator. Surmounting this challenge bring about the development of various imputation-based mean estimation methods. Some of these, ratio-type regression estimators have been devised to compute population parameters using only current sample data. However, recent researches has changed this approach by integrating both past and current sample information through the application of exponentially weighted moving averages (EWMA). This new invented methodology has given rise to the creation of memory-type estimators tailored for surveys conducted over time. In this paper, we present regression- imputation and EWMA based memory-type mean estimators in the presence of non-response. For the performance assessment between existing and proposed estimators, two real-life time-scaled data sets are considered. In the first data set proposed estimators (Tp1 and Tp2) have the minimum MSE and higher PRE and also in the second data set, results show that the proposed estimators (estimators (Tp1 and Tp2) have the minimum MSE and higher PRE when compared with the existing estimators at all levels of smoothing parameter. Therefore the proposed estimator will perform better in estimating the EWMA-based mean in the presence non response for time based survey and longitudinal survey. Most especially in quality control data that are collected over time. The current work is limited to simple random sampling which assumes homogeneity of the population units. Further work can be done on stratify random sampling which assumes heterogeneity of the population units.

Keywords: Auxiliary information, missing information; imputation methods; EWMA, Mean Square Error, Percentage Relative Efficiency.

UNDER PEER REVIEW

1.0 INTRODUCTION

Sample surveys are valuable tools for researches in different field of endeavor of which quality control is among. It provides insights into the quality of products and processes that would be difficult to obtain otherwise. Incorporating the use of sample surveys, organizations can identify areas for improvement, monitor process performance, evaluate the effectiveness of quality control initiatives, and predict future quality problems. Therefore the role of sampling in the field of quality control cannot be over emphasized [1]. Majority of the estimators available in the existing literature rely solely on the present sample. However, by incorporating data from both current and previous samples aids in the improvement of the efficiency of the estimator significantly. This is particularly crucial in scenarios where surveys are carried out at regular intervals, like monthly or annual surveys, hence the introduction exponentially weighted moving average (EWMA)[2]

Recently, several authors have introduced concepts of estimators in EWMA statistics using current and past data. Some of these authors includes;[3] , [4], [5] ,[6], [7], [1],. . However, the EWMA based estimators proposed by above mentioned authors are based on the assumptions that complete sample information is available which is not always true in real life situation due to non-response or missing observation(s).

REVIEW

Non-Response or Missing data is a common and serious problem in survey sampling. Missing data naturally occurs in sample surveys when a few sampling units refuse to respond or are unable to participate in the survey. The problem of non-response make the estimators so obtained from the responding units alone to be biased and less efficient even if the sample size is increased to make up for nonresponse, this may reduce only their variances but not the biases. The occurrence of missing data within a dataset can be attributed to various factors. These include human or machine errors, unanswered questions during surveys, equipment malfunctions, latent variables, and misjudgments. By employing the imputation approach, we can investigate the root causes behind the presence of missing data in the dataset. This entails identifying whether the missing-ness is inherent to the fundamental observations or originates from the broader population that serves as the source of the sample. Rubin's [15]. The work primarily categorized its concepts into three distinct mechanisms based on whether a particular missing value is linked to fundamental observations. These mechanisms are referred to as Missing Completely at Random (MCAR), Missing at Random (MAR), and Not Missing at Random (NMAR). Prior to employing a suitable statistical methodology to analyze a dataset with missing values, it is crucial to ascertain which of the aforementioned conditions is at play within the dataset. A dataset is labeled as MCAR when the probability of data being missing is not connected to any observable or unobservable measurements present in the data. Conversely, a dataset is termed as MAR when the probability of missing values is tied to observed measurements but not to unobserved ones. Lastly, a dataset is assigned to the NMAR category when the probability of missing data is influenced by both observable and unobservable measurements [16].

This challenge of missing information brings about the development of various imputation-based mean estimation methods to address this concern. Among these is ratio-type regression estimators which has been devised to compute population parameters using only current sample data [6],. Recent researches has revolutionized the approach of missing data in estimator by integrating both past and current sample information through the application of exponentially weighted moving averages (EWMA). This methodology has given rise to the creation of memory-type estimators tailored for surveys conducted over time such as EWMA-based imputed traditional mean estimators proposed by [8],

2.0 Methodology

2.1 Existing estimators in the literature

The EWMA statistic is defined as in (1.1).

$$Z_t = \lambda \bar{x}_t + (1-\lambda)Z_{t-1} \tag{1.1}$$

where $0 < \lambda \leq 1$, $Z_{t-1} =$ EWMA of the previous period , $\bar{x} =$ The mean of the current sample of the study variable, $\lambda =$ denotes the smoothing constant, a weight assigned to the observations.

A higher value of λ indicates greater emphasis on recent data, while less importance is given to pastvalues. Conversely, a smaller λ value gives more weight to historical data and less to the latest observations. The starting value, Z_0 is determined by taking the expected mean or average of the prior sample. [9], in his work titled "Memory Type Ratio and Product Estimators for Population Mean for Time-based Surveys" proposed memory type ratio and product estimator as defined in (1.2) and (1.3)

$$\hat{\hat{\bar{Y}}}_{rmi} = \frac{Z_t}{Q_t}\mu_x \tag{1.2}$$

$$\hat{\hat{\bar{Y}}}_{pmi} = \frac{Z_t}{\mu_x}Q_t \tag{1.3}$$

where $Z_t = \lambda \bar{y} + (1-\lambda)Z_{t-1}$ and $Q_t = \lambda \bar{x} + (1-\lambda)Q_{t-1}$

The mean square errors of (1.2) and (1.3) are given as in (1.4) and (1.5) respectfully

$$MSE(\hat{\hat{\bar{Y}}}_{rmi}) = \theta \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2t}\right)\left[C_y^2 + C_x^2 - 2\rho C_y C_x\right] \tag{1.4}$$

UNDER PEER REVIEW

$$MSE(\hat{\bar{Y}}_{pmi}) = \theta \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2t}\right)\left[C_y^2 + C_x^2 + 2\rho C_y C_x\right] \quad (1.5)$$

Where $\theta = \frac{1}{n} - \frac{1}{N}$

The limiting form, that is as $t \to 0$, the variance in (1.4) and (1.5)are given by (1.6) and (1.7) respectively

$$MSE(\hat{\bar{Y}}_{rmi}) = \theta \frac{\lambda}{2-\lambda}\left[C_y^2 + C_x^2 - 2\rho C_y C_x\right] \quad (1.6)$$

$$MSE(\hat{\bar{Y}}_{pmi}) = \theta \frac{\lambda}{2-\lambda}\left[C_y^2 + C_x^2 + 2\rho C_y C_x\right] \quad (1.7)$$

[10], in their work titled "improving efficiencies of ratio-and product-type estimators for estimating population mean for time based survey" suggested ratio and product type estimators of the population mean using the known coefficient of the variation C_x of the auxiliary variable as defined in (1.8) and (1.9) respectively

$$T_{R4} = Z_t\left(\frac{\bar{X} + C_x}{Q_t + C_x}\right) \quad (1.8)$$

$$T_{P4} = Z_t\left(\frac{Q_t + C_x}{\bar{X} + C_x}\right) \quad (1.9)$$

The biases and mean square errors of (1.8) and (1.9) are given as in (1.10) (1.11), (1.12) and (1.13) respectfully

$$Bias(T_{R4}) = \theta \bar{Y} \frac{(\lambda_1 \lambda_2)^2}{(\lambda_2 - \lambda_1)^2} \Theta_t \alpha \left[C_y^2 + \alpha^2 C_x^2 - 2\alpha\rho C_y C_x\right] \quad (1.10)$$

$$Bias(T_{P4}) = \theta \bar{Y} \frac{(\lambda_1 \lambda_2)^2}{(\lambda_2 - \lambda_1)^2} \Theta_t \alpha \rho C_y C_x$$

(2.14)

UNDER PEER REVIEW

$$MSE(T_{R4}) = \theta\bar{Y}^2 \frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta_t\left[C_y^2+\alpha^2C_x^2-2\alpha\rho C_yC_x\right] \quad (1.11)$$

$$MSE(T_{P4}) = \theta\bar{Y}^2 \frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta_t\left[C_y^2+\alpha^2C_x^2+2\alpha\rho C_yC_x\right] \quad (1.12)$$

Where

$$\Theta_t = \frac{(1-\lambda_1)^2(1-(1-\lambda_1)^{2t})}{1-(1-\lambda_1)^2} + \frac{(1-\lambda_2)^2(1-(1-\lambda_2)^{2t})}{1-(1-\lambda_2)^2} - \frac{2(1-\lambda_1)(1-\lambda_2)\left\{1-(1-\lambda_1)^t(1-\lambda_2)^t\right\}}{1-(1-\lambda_1)(1-\lambda_2)}$$

The limiting form, that is as $t\to 0$, the biases and MSEs in (2.13) (2.14), (2.15) and (2.16) respectfully are given by((2.17) ((2.18), ((2.19) and ((2.20) respectfully

$$Bias(T_{R4}) = \theta\bar{Y}\frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta\alpha\left[C_y^2+\alpha^2C_x^2-2\alpha\rho C_yC_x\right] \quad ((1.13)$$

$$Bias(T_{P4}) = \theta\bar{Y}\frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta\alpha\rho C_yC_x \quad ((1.14)$$

$$MSE(T_{R4}) = \theta\bar{Y}^2\frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta\left[C_y^2+\alpha^2C_x^2-2\alpha\rho C_yC_x\right] \quad ((1.15)$$

$$MSE(T_{P4}) = \theta\bar{Y}^2\frac{(\lambda_1\lambda_2)^2}{(\lambda_2-\lambda_1)^2}\Theta\left[C_y^2+\alpha^2C_x^2+2\alpha\rho C_yC_x\right] \quad ((1.16)$$

Where $\Theta = \frac{(1-\lambda_1)^2}{1-(1-\lambda_1)^2} + \frac{(1-\lambda_2)^2}{1-(1-\lambda_2)^2} - \frac{2(1-\lambda_1)(1-\lambda_2)}{1-(1-\lambda_1)(1-\lambda_2)}$

Other authors that worked on memory-type estimators are [11], and [12],

[8], proposed memory- type estimators for a time based survey in the presence of non-response using compromised imputation method as in (1.17), (1.18) and (1.19)

UNDER PEER REVIEW

$$T_1 = \frac{Z_{t_r} + b\left(\bar{X} - Q_t\right)}{Q_t}\bar{X} \tag{1.17}$$

$$T_2 = \frac{Z_{t_r} + b\left(\bar{X} - Q_{t_r}\right)}{Q_{t_r}}\bar{X} \tag{1.18}$$

$$T_3 = \frac{Z_{t_r} + b\left(\bar{x} - Q_{t_r}\right)}{Q_{t_r}}\bar{x} \tag{1.19}$$

The MSE of the estimators $T_1, T_2\ and\ T_3$ are as in (1.20), (1.21) and (1.22)

$$MSE\left(T_1\right) = \left(1-\left(1-\lambda\right)^{2t}\right)\left(\theta_2 S_y^2 + \theta_1\left(R^2 S_x^2 - \rho S_y^2\right)\right) \tag{1.20}$$

$$MSE\left(T_2\right) = \theta_2\left(1-\left(1-\lambda\right)^{2t}\right)\left(S_y^2\left(1-\rho^2\right) + R^2 S_x^2\right) \tag{1.21}$$

$$MSE\left(T_3\right) = \left(1-\left(1-\lambda\right)^{2t}\right)\left(\theta_2 S_y^2 + \theta_3\left(R^2 S_x^2 - \rho^2 S_y^2\right)\right) \tag{1.22}$$

where $\theta_1 = \frac{\lambda}{2-\lambda}\left(\frac{1}{n} - \frac{1}{N}\right)$, $\theta_2 = \frac{\lambda}{2-\lambda}\left(\frac{1}{r} - \frac{1}{N}\right)$ and $\theta_3 = \frac{\lambda}{2-\lambda}\left(\frac{1}{r} - \frac{1}{n}\right)$

The limiting form, that is, as $t \to 0$, MSEs in (1.20), (1.21) and (1.22)are given by (1.23), (1.24) and (1.25) respectively

$$MSE\left(T_1\right) = \theta_2 S_y^2 + \theta_1\left(R^2 S_x^2 - \rho S_y^2\right) \tag{1.23}$$

$$MSE\left(T_2\right) = \theta_2\left(S_y^2\left(1-\rho^2\right) + R^2 S_x^2\right) \tag{1.24}$$

$$MSE\left(T_3\right) = \theta_2 S_y^2 + \theta_3\left(R^2 S_x^2 - \rho^2 S_y^2\right) \tag{1.25}$$

2.2 Proposed Estimators

However the estimators proposed by [8], are less efficient when the correlation between the study and auxiliary variables is weak and the estimators are based on compromised imputation approach in which the responses of the respondents are not directly used.

This study proposed modified classes of estimators with the following modification

UNDER PEER REVIEW

i. Modified estimators using regression-type imputation schemes and incorporating additional auxiliary variable to obtain new classes of estimators that are robust against weak correlations
ii. The modified estimators were designed based on the regression-type imputation which utilizes the real responses of the respondents.

Motivated by imputation schemes proposed by [13], and [14], the following regression-type estimators for EWMA statistics were proposed.

$$T_{p_1} = \frac{r}{n} Z_{tr} + \left(1 - \frac{r}{n}\right) \frac{Z_{tr} + \hat{\beta}\left(\bar{X} - Q_t\right)}{\left(\pi_1 Q_t + \pi_2\right)} \left(\pi_1 \bar{X} + \pi_2\right) \tag{2.1}$$

$$T_{p_2} = \frac{r}{n} Z_{tr} + \left(1 - \frac{r}{n}\right) \frac{Z_{tr} + \hat{\beta}\left(\bar{X} - Q_{tr}\right)}{\left(\pi_1 Q_{tr} + \pi_2\right)} \left(\pi_1 \bar{X} + \pi_2\right) \tag{2.2}$$

The proposed estimator T_{p_1} was formulated for the aforementioned cases of Missing Completely at Random (MCAR) and T_{p_2} was formulated for the cases of Missing at Random (MAR)

2.3 Properties (Bias and MSE)

We define sampling relative error of EWMA statistics as follows

$$e_0 = \frac{Z_{tr} - \bar{Y}}{\bar{Y}};\; e_1 = \frac{Q_{tr} - \bar{X}}{\bar{X}},\; e_2 = \frac{Q_t - \bar{X}}{\bar{X}},\; e_3 = \frac{\bar{x} - \bar{X}}{\bar{X}} \tag{2.4}$$

The expectations of the error terms for large sample at time t are

$$\begin{aligned}
&E(e_0) = E(e_1) = E(e_2) = E(e_3) = 0\\
&E\left(e_0^2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_2 C_y^2,\; E\left(e_1^2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_2 C_x^2,\\
&E\left(e_2^2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_1 C_x^2,\; E\left(e_3^2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_4 C_x^2\\
&E\left(e_0 e_1\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_2 C_{yx},\; E\left(e_0 e_2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_1 C_{yx},\\
&E\left(e_0 e_3\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_4 C_{yx} E\left(e_1 e_2\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_1 C_X^2,\\
&E\left(e_1 e_3\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_4 C_X^2,\; E\left(e_2 e_3\right) = \left(1 - (1-\lambda)^{2t}\right)\theta_4 C_X^2
\end{aligned} \tag{2.5}$$

As time t tends to zero ($t \to 0$), the expectations becomes as is (2.6)

UNDER PEER REVIEW

$$E(e_0)=E(e_1)=E(e_2)=E(e_3)=0$$
$$E(e_0^2)=\theta_2 C_y^2,\ E(e_1^2)=\theta_2 C_x^2,\ E(e_2^2)=\theta_1 C_x^2,\ E(e_3^2)=\theta_4 C_x^2$$
$$E(e_0e_1)=\theta_2 C_{yx},\ E(e_0e_2)=\theta_1 C_{yx},\ E(e_0e_3)=\theta_4 C_{yx} \quad (2.6)$$
$$E(e_1e_2)=\theta_1 C_X^2,\ E(e_1e_3)=\theta_4 C_X^2,\ E(e_2e_3)=\theta_4 C_X^2$$

where
$$\theta_1=\frac{\lambda}{\lambda-2}\left(\frac{1}{n}-\frac{1}{N}\right),\ \theta_2=\frac{\lambda}{\lambda-2}\left(\frac{1}{r}-\frac{1}{N}\right),$$
$$\theta_3=\frac{\lambda}{\lambda-2}\left(\frac{1}{n}-\frac{1}{n}\right),\ \theta_4=\left(\frac{1}{n}-\frac{1}{N}\right)$$

The estimators $T_{p_i}\ (i=1,2,3)$ can be expressed using errors terms then simplified to the first order approximation using Taylor's series as in (2.7)

$$T_{p_1}=\frac{r}{n}(1+e_0)\bar{Y}+\left(1-\frac{r}{n}\right)\frac{(1+e_0)\bar{Y}+\hat{\beta}(1+e_2)\bar{X}}{\left(\pi_1(1+e_2)\bar{X}+\pi_2\right)}\left(\pi_1\bar{X}+\pi_2\right) \quad (2.7)$$

Simplify (2.7) to first order approximation to obtain (2.8)

$$T_{p_1}=\bar{Y}+\bar{Y}e_0+\left(1-\frac{r}{n}\right)\left(\left(\bar{Y}H+\hat{\beta}\bar{X}\right)e_2+\left(\bar{Y}H^2+\hat{\beta}\bar{X}H\right)e_2^2-\left(\bar{Y}H\right)e_0e_2\right) \quad (2.8)$$

Subtract $\bar{Y}$ from both side of (2.8) and take the expectation, the bias of the proposed estimator is obtained as in (2.9) below

$$Bias\left(T_{p_1}\right)=\left(1-\frac{r}{n}\right)\theta_1 H\left(\left(\bar{Y}H+\hat{\beta}\bar{X}\right)C_x^2-\bar{Y}\,C_{yx}\right) \quad (2.9)$$

To obtain the MSE T_{p_1}, Expectation of the square of (2.9) is taking as in (2.10)

UNDER PEER REVIEW

$$E\left(T_{p_1}-\bar{Y}\right)^2 = E\begin{pmatrix} \bar{Y}e_0 + e_2\left(\bar{Y}H - \frac{r}{n}\bar{Y}H + \hat{\beta}\bar{X} - \frac{r}{n}\hat{\beta}\bar{X}\right) - e_0 e_2\left(\bar{Y}H - \frac{r}{n}\bar{Y}H\right) \\ + e_2^2\left(\bar{Y}H^2 - \frac{r}{n}\bar{Y}H^2 + \hat{\beta}\bar{X}H - \frac{r}{n}\hat{\beta}\bar{X}H\right) \end{pmatrix}^2 \tag{2.10}$$

Simplify (2.10) to first order approximation,

$$MSE\left(T_{p_1}\right) = Y\theta_2 S_y^2 + \left(1-\frac{r}{n}\right)\left(RH+\hat{\beta}\right)^2 S_X^2\theta_1 - 2\theta_1\left(1-\frac{r}{n}\right)\left(RH+\hat{\beta}\right)\rho S_Y S_X \tag{2.11}$$

For estimator T_{p_2} ,

$$T_{p_2} = \frac{r}{n}\left(1+e_0\right)\bar{Y} + \left(1-\frac{r}{n}\right)\frac{\left(1+e_0\right)\bar{Y} + \hat{\beta}\left(1+e_1\right)\bar{X}}{\left(\pi_1\left(1+e_1\right)\bar{X}+\pi_2\right)}\left(\pi_1\bar{X}+\pi_2\right) \tag{2.12}$$

Simplify (2.12) to first order approximation to obtain (2.13)

$$T_{p_2} = \bar{Y} + \bar{Y}e_0 + \left(1-\frac{r}{n}\right)\left(\left(\bar{Y}H + \hat{\beta}\bar{X}\right)e_1 + \left(\bar{Y}H^2 + \hat{\beta}\bar{X}H\right)e_1^2 - \left(\bar{Y}H\right)e_0 e_1\right) \tag{2.13}$$

Subtract $\bar{Y}$ from both side of (2.13) and take the expectation to obtain the bias as in (2.14)

$$Bias\left(T_{p_1}\right) = \left(1-\frac{r}{n}\right)\theta_2 H\left(\left(\bar{Y}H + \hat{\beta}\bar{X}\right)C_x^2 - \bar{Y}\,C_{yx}\right) \tag{2.14}$$

To obtain the MSE T_{p_2} , Expectation of the square of (2.13) is taking as in (2.15)

$$E\left(T_{p_1}-\bar{Y}\right)^2 = E\left(\bar{Y}e_0 + \left(1-\frac{r}{n}\right)\left(\left(\bar{Y}H + \hat{\beta}\bar{X}\right)e_1 + \left(\bar{Y}H^2 + \hat{\beta}\bar{X}H\right)e_1^2 - \left(\bar{Y}H\right)e_0 e_1\right)\right)^2 \tag{2.15}$$

Simplify (2.15) to first order approximation gives;

UNDER PEER REVIEW

$$MSE\left(T_{p_1}\right)=\theta_2 S_y^2+\left(1-\frac{r}{n}\right)\left(RH+\hat{\beta}\right)^2 S_X^2\theta_2-2\theta_2\left(1-\frac{r}{n}\right)\left(RH+\hat{\beta}\right)\rho S_Y S_X \qquad (2.16)$$

3. Results and Discussion

This section illustrate numerical analysis to show the performance of proposed estimators $T_{p_i}\ \left(i=1,2,3\right)$ with respect to existing estimators of Alomair and Shahzad (2023) using two (2) existing datasets.

Dataset 1: Murthy (1967) [17]

$\mathrm{N}=284;\mathrm{n}=35;\mathrm{r}=25;Y\mathrm{bar}=29.36;X\mathrm{bar}=245.088;C\mathrm{y}=1.76;C\mathrm{x}=2.43;$

$\mathrm{Sy}=Y\mathrm{bar}*C\mathrm{y};S\mathrm{x}=X\mathrm{bar}*C\mathrm{x};\mathrm{rho}=0.961;\ \mathrm{kt}=88.88;\mathrm{sk}=8.77;$

Dataset 2: Cochran (1977) [18]

$\mathrm{N}=80;\mathrm{n}=25;\mathrm{r}=20;Y\mathrm{bar}=5182.638;X\mathrm{bar}=285.125;C\mathrm{y}=0.3542;C\mathrm{x}=0.9485;$

$\mathrm{Sy}=Y\mathrm{bar}*C\mathrm{y};S\mathrm{x}=X\mathrm{bar}*C\mathrm{x};\mathrm{rho}=0.914;\ \mathrm{kt}=3.5360;\mathrm{sk}=1.2680;$

The Percentage Relative Efficiency (PRE) is given by:

$$PRE\left(estimators\right)=\frac{MSE\left(t_0\right)}{MSE\left(estimator\right)}\times 100$$

Table 1: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023
Data Set1 at $\lambda=0.2$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	10.82272	100	t_0	10.82272	100
T1	18.12664	59.70616	T1	18.12664	59.70616
T2	21.45887	50.43471	T2	21.45887	50.43471
T3	14.15495	76.45891	T3	14.15495	76.45891
Proposed			Proposed		
$T_{P_{11}}$	4.592535	235.6589	$T_{P_{21}}$	1.750171	618.3808
$T_{P_{12}}$	6.385788	169.4813	$T_{P_{22}}$	4.361549	248.1393
T_{P3}	4.609453	234.794	T_{P23}	1.774806	609.7971
$T_{P_{14}}$	5.130033	210.9678	$T_{P_{24}}$	2.532888	427.2877
$T_{P_{15}}$	4.653055	232.5938	$T_{P_{25}}$	1.838301	588.7348

UNDER PEER REVIEW

$T_{P_{16}}$	4.592564	235.6574	$T_{P_{26}}$	1.750212	618.3661
$T_{P_{17}}$	4.593577	235.6054	$T_{P_{27}}$	1.751688	617.845
$T_{P_{18}}$	4.592638	235.6536	$T_{P_{28}}$	1.75032	618.3278
$T_{P_{19}}$	5.737601	188.6279	$T_{P_{29}}$	3.417643	316.672
$T_{P_{110}}$	4.834336	223.8719	$T_{P_{210}}$	2.102286	514.807
$T_{P_{111}}$	5.737601	188.6279	$T_{P_{211}}$	3.417643	316.672
$T_{P_{112}}$	4.638911	233.303	$T_{P_{212}}$	1.817705	595.4057
$T_{P_{113}}$	4.592726	235.6491	$T_{P_{213}}$	1.750449	618.2825
$T_{P_{114}}$	4.593224	235.6236	$T_{P_{214}}$	1.751174	618.0264
$T_{P_{115}}$	5.01839	215.6611	$T_{P_{215}}$	2.370312	456.5947
$T_{P_{116}}$	4.59447	235.5597	$T_{P_{216}}$	1.752988	617.387
$T_{P_{117}}$	4.662335	232.1309	$T_{P_{217}}$	1.851814	584.4386

Table 2: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023

Data Set1 at $\lambda = 0.5$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	32.46815	100	t_0	32.46815	100
T1	54.37991	59.70616	T1	54.37991	59.70616
T2	64.3766	50.43471	T2	64.3766	50.43471
T3	42.46484	76.45891	T3	42.46484	76.45891
Proposed			Proposed		
$T_{P_{11}}$	13.77761	235.6589	$T_{P_{21}}$	5.250512	618.3808
$T_{P_{12}}$	19.15736	169.4813	$T_{P_{22}}$	13.08465	248.1393
T_{P3}	13.82836	234.794	T_{P23}	5.324419	609.7971
$T_{P_{14}}$	15.3901	210.9678	$T_{P_{24}}$	7.598664	427.2877
$T_{P_{15}}$	13.95916	232.5938	$T_{P_{25}}$	5.514903	588.7348
$T_{P_{16}}$	13.77769	235.6574	$T_{P_{26}}$	5.250636	618.3661
$T_{P_{17}}$	13.78073	235.6054	$T_{P_{27}}$	5.255065	617.845
$T_{P_{18}}$	13.77791	235.6536	$T_{P_{28}}$	5.250961	618.3278
$T_{P_{19}}$	17.2128	188.6279	$T_{P_{29}}$	10.25293	316.672
$T_{P_{110}}$	14.50301	223.8719	$T_{P_{210}}$	6.306859	514.807
$T_{P_{111}}$	17.2128	188.6279	$T_{P_{211}}$	10.25293	316.672
$T_{P_{112}}$	13.91673	233.303	$T_{P_{212}}$	5.453115	595.4057
$T_{P_{113}}$	13.77818	235.6491	$T_{P_{213}}$	5.251346	618.2825

UNDER PEER REVIEW

$T_{P_{114}}$	13.77967	235.6236	$T_{P_{214}}$	5.253522	618.0264
$T_{P_{115}}$	15.05517	215.6611	$T_{P_{215}}$	7.110935	456.5947
$T_{P_{116}}$	13.78341	235.5597	$T_{P_{216}}$	5.258963	617.387
$T_{P_{117}}$	13.987	232.1309	$T_{P_{217}}$	5.555443	584.4386

Table 3: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023 Data Set1 at $\lambda = 0.7$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	52.44856	100	t_0	52.44856	100
T1	87.84447	59.70616	T1	87.84447	59.70616
T2	103.993	50.43471	T2	103.993	50.43471
T3	68.59705	76.45891	T3	68.59705	76.45891
Proposed			Proposed		
$T_{P_{11}}$	22.25613	235.6589	$T_{P_{21}}$	8.481596	618.3808
$T_{P_{12}}$	30.94651	169.4813	$T_{P_{22}}$	21.13674	248.1393
T_{P3}	22.33812	234.794	T_{P23}	8.600985	609.7971
$T_{P_{14}}$	24.86093	210.9678	$T_{P_{24}}$	12.27476	427.2877
$T_{P_{15}}$	22.54942	232.5938	$T_{P_{25}}$	8.90869	588.7348
$T_{P_{16}}$	22.25627	235.6574	$T_{P_{26}}$	8.481797	618.3661
$T_{P_{17}}$	22.26118	235.6054	$T_{P_{27}}$	8.488951	617.845
$T_{P_{18}}$	22.25663	235.6536	$T_{P_{28}}$	8.482322	618.3278
$T_{P_{19}}$	27.8053	188.6279	$T_{P_{29}}$	16.56243	316.672
$T_{P_{110}}$	23.42793	223.8719	$T_{P_{210}}$	10.188	514.807
$T_{P_{111}}$	27.8053	188.6279	$T_{P_{211}}$	16.56243	316.672
$T_{P_{112}}$	22.48088	233.303	$T_{P_{212}}$	8.808878	595.4057
$T_{P_{113}}$	22.25706	235.6491	$T_{P_{213}}$	8.482943	618.2825
$T_{P_{114}}$	22.25947	235.6236	$T_{P_{214}}$	8.486459	618.0264
$T_{P_{115}}$	24.31989	215.6611	$T_{P_{215}}$	11.4869	456.5947
$T_{P_{116}}$	22.26551	235.5597	$T_{P_{216}}$	8.495249	617.387
$T_{P_{117}}$	22.59439	232.1309	$T_{P_{217}}$	8.974177	584.4386

UNDER PEER REVIEW

Table 4: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023 Data Set2 at $\lambda = 0.2$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	14040.66	100	t_0	14040.66	100
T1	79274.76	17.71139	T1	79274.76	17.71139
T2	102996.2	13.63221	T2	102996.2	13.63221
T3	37762.15	37.18184	T3	37762.15	37.18184
Proposed			Proposed		
$T_{P_{11}}$	5833.065	240.7082	$T_{P_{21}}$	2848.484	492.9171
$T_{P_{12}}$	7583.183	185.1553	$T_{P_{22}}$	5235.009	268.2071
T_{P3}	5840.251	240.412	T_{P23}	2858.283	491.2272
$T_{P_{14}}$	5859.938	239.6043	$T_{P_{24}}$	2885.128	486.6564
$T_{P_{15}}$	5842.675	240.3122	$T_{P_{25}}$	2861.589	490.6597
$T_{P_{16}}$	5833.091	240.7071	$T_{P_{26}}$	2848.52	492.9108
$T_{P_{17}}$	5833.164	240.7041	$T_{P_{27}}$	2848.619	492.8937
$T_{P_{18}}$	5833.1	240.7067	$T_{P_{28}}$	2848.532	492.9087
$T_{P_{19}}$	7650.211	183.533	$T_{P_{29}}$	5326.41	263.6046
$T_{P_{110}}$	5861.403	239.5444	$T_{P_{210}}$	2887.127	486.3196
$T_{P_{111}}$	7650.211	183.533	$T_{P_{211}}$	5326.41	263.6046
$T_{P_{112}}$	6421.528	218.6499	$T_{P_{212}}$	3650.933	384.5774
$T_{P_{113}}$	5835.095	240.6244	$T_{P_{213}}$	2851.253	492.4384
$T_{P_{114}}$	5835.78	240.5962	$T_{P_{214}}$	2852.186	492.2773
$T_{P_{115}}$	7297.187	192.412	$T_{P_{215}}$	4845.013	289.7962
$T_{P_{116}}$	5838.731	240.4746	$T_{P_{216}}$	2856.21	491.5837
$T_{P_{117}}$	5854.24	239.8375	$T_{P_{217}}$	2877.358	487.9706

Table 5: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023
Data Set2 at $\lambda = 0.5$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	42121.99	100	t_0	42121.99	100
T1	237824.3	17.71139	T1	237824.3	17.71139
T2	308988.7	13.63221	T2	308988.7	13.63221
T3	13286.5	37.18184	T3	13286.5	37.18184
Proposed			Proposed		
$T_{P_{11}}$	17499.19	240.7082	$T_{P_{21}}$	8545.451	492.9171
$T_{P_{12}}$	22749.55	185.1553	$T_{P_{22}}$	15705.03	268.2071
$T_{P_{13}}$	17520.75	240.412	T_{P23}	8574.849	491.2272
$T_{P_{14}}$	17579.81	239.6043	$T_{P_{24}}$	8655.385	486.6564

UNDER PEER REVIEW

$T_{P_{15}}$	17528.03	240.3122	$T_{P_{25}}$	8584.767	490.6597
$T_{P_{16}}$	17499.27	240.7071	$T_{P_{26}}$	8545.56	492.9108
$T_{P_{17}}$	17499.49	240.7041	$T_{P_{27}}$	8545.856	492.8937
$T_{P_{18}}$	17499.3	240.7067	$T_{P_{28}}$	8545.596	492.9087
$T_{P_{19}}$	22950.63	183.533	$T_{P_{29}}$	15979.23	263.6046
$T_{P_{110}}$	17584.21	239.5444	$T_{P_{210}}$	8661.38	486.3196
$T_{P_{111}}$	22950.63	183.533	$T_{P_{211}}$	15979.23	263.6046
$T_{P_{112}}$	19264.58	218.6499	$T_{P_{212}}$	10952.8	384.5774
$T_{P_{113}}$	17505.29	240.6244	$T_{P_{213}}$	8553.758	492.4384
$T_{P_{114}}$	17507.34	240.5962	$T_{P_{214}}$	8556.557	492.2773
$T_{P_{115}}$	21891.56	192.412	$T_{P_{215}}$	14535.04	289.7962
$T_{P_{116}}$	17516.19	240.4746	$T_{P_{216}}$	8568.63	491.5837
$T_{P_{117}}$	17562.72	239.8375	$T_{P_{217}}$	8632.075	487.9706

Table 6: MSEs and PREs of the proposed and existing estimators of alomair and shazhad 2023
Data Set2 at $\lambda = 0.7$

estimator	MSE	PRE	estimator	MSE	PRE
t_0	68043.21	100	t_0	68043.21	100
T1	384177.7	17.71139	T1	384177.7	17.71139
T2	499135.6	13.63221	T2	499135.6	13.63221
T3	183001.2	37.18184	T3	183001.2	37.18184
Proposed			Proposed		
$T_{P_{11}}$	28267.93	240.7082	$T_{P_{21}}$	13804.19	492.9171
$T_{P_{12}}$	36749.27	185.1553	$T_{P_{22}}$	25369.66	268.2071
T_{P3}	28302.76	240.412	T_{P23}	13851.68	491.2272
$T_{P_{14}}$	28398.16	239.6043	$T_{P_{24}}$	13981.78	486.6564
$T_{P_{15}}$	28314.5	240.3122	$T_{P_{25}}$	13867.7	490.6597
$T_{P_{16}}$	28268.06	240.7071	$T_{P_{26}}$	13804.37	492.9108
$T_{P_{17}}$	28268.41	240.7041	$T_{P_{27}}$	13804.84	492.8937
$T_{P_{18}}$	28268.1	240.7067	$T_{P_{28}}$	13804.42	492.9087
$T_{P_{19}}$	37074.1	183.533	$T_{P_{29}}$	25812.6	263.6046
$T_{P_{110}}$	28405.26	239.5444	$T_{P_{210}}$	13991.46	486.3196
$T_{P_{111}}$	37074.1	183.533	$T_{P_{211}}$	25812.6	263.6046
$T_{P_{112}}$	31119.71	218.6499	$T_{P_{212}}$	17692.98	384.5774

UNDER PEER REVIEW

$T_{P_{113}}$	28277.77	240.6244	$T_{P_{213}}$	13817.61	492.4384
$T_{P_{114}}$	28281.09	240.5962	$T_{P_{214}}$	13822.13	492.2773
$T_{P_{115}}$	35363.29	192.412	$T_{P_{215}}$	23479.68	289.7962
$T_{P_{116}}$	28295.39	240.4746	$T_{P_{216}}$	13841.63	491.5837
$T_{P_{117}}$	28370.55	239.8375	$T_{P_{217}}$	13944.12	487.9706

Critical review of Table 1 to table 6 shows that the mean square errors and percentage relative efficiencies of the proposed estimators and the existing related estimators considered in the study using the two datasets revealed that, proposed estimators has minimum MSEs and higher PREs compared to the existing estimators at different levels of smoothing constant λ. It could also be observed that the estimator T_{P2} is highly efficient than other proposed estimator. This shows that the suggested estimators are more efficient than the counterparts existing estimator and have higher chances of producing estimates closer to the true values of means for any population of interest.

4. Conclusions

This study proposed two modified memory-type regression imputation estimators which utilizes additional auxiliary information to obtain robust EWMA statistics that is capable of detecting reliable shift during production process.

The variants of the proposed estimators were derived up to first order of approximation using Taylor series techniques. From the empirical study, the results show that the proposed estimators are better than the existing estimators considered in this study. Hence we recommend that the proposed estimators should be used in practice.

UNDER PEER REVIEW

References

[1] Singh, G.N.; Bhattacharyya, D.; Bandyopadhyay, A.; Khalid, M.(2021). Study of a memory type shrinkage estimator of population mean in quality control process. *IEEE Access*, 9, 161555–161564.

[2] Montgomery, D.C. (1991), *Introduction to Statistical Quality Control*, Second Edition, John Wiley and Sons, New York.

[3] Haq A. A (2013). New hybrid exponentially weighted moving average control chart for monitoring process mean. *Quality and Reliability Engineering International*; **29**(7):1015–1025

[4] Crowder, S.V. (1987), "Run-Length Distributions of EWMA Charts," *Technometrics*, 29, 401-407.

[5] Lucas, J.M. and Saccucci, M.S. (1990), "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," with discussion, *Technometrics*,32, 1-29.

[6] Noor-ul-Amin, M. (2021). Memory type estimators of population mean using exponentially weighted moving averages for time scaled surveys. *Commun. Stat.-Theory Methods*, 50, 2747–2758.

[7] Aslam, I.; Noor-ul-Amin, M.; Hanif, M. and Sharma, P (2022) Memory type ratio and product estimators under ranked-based sampling schemes. *Commun. Stat.-Theory Methods*, 52, 1155–1177.

[8] Alomair,M.A. and Shahzad, U.(2023). Compromised-Imputation and EWMA-Based Memory-Type Mean Estimators Using Quantile Regression. *Symmetry*, 15, 1888.

[9] Noor-ul-Amin, M. (2020) Memory type ratio and product estimators for population mean for time-based surveys. *J. Stat. Comput. Simul*. 90, 3080–3092.

UNDER PEER REVIEW

[10] Chhaparwal, P.; Kumar, S.(2022) Improving Efficiencies of Ratio-and Product-type Estimators for Estimating Population Mean for Time-based Survey. *J. Reliab. Stat. Stud.* , 15, 325–340.

[11] Bhushan, S.; Kumar, A.; Al-Omari, A.I.; Alomani, G.A.(**2023).** Mean Estimation for Time- Based Surveys Using Memory-Type Logarithmic Estimators. *Mathematics*,11, 125.

[12] Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling.*Metrika*, *51*(3), 266-276

[13] Audu, A. and Singh, R.V.K (2021): Exponential-Type Regression Compromised Imputation class of estimators. *Journal of statistics and management system*, 24(6), 1253-1266

[14] Audu, A., Singh, R. and Khare, S. (2023): New Regression –Type Compromised Imputation Class of estimators with Known Parameters of Auxiliary Variable. *Comm. stat.: Simulation and computations*. 52(10), 4789-4801.

[15] Rubin, D.B. (1976): Inference and missing data. *Biometrika*, 63, 581–592. [CrossRef]

[16] Gordon, C.A. (2010): Investigating Statistical Approaches to Handling Missing Data in the Context of the Gateshead Millennium Study, Master's Thesis, University of Glasgow, Glasgow, UK, 2010. 3.

[17] Murthy, M.N. (1967): Sampling theory and methods. Calcutta. India: Statistical Publishing Society

[18] Cochran, W.G. (1977) Sampling techniques 3rd ed. New York: John Wiley and Son

UNDER PEER REVIEW