RT&A, No 1 (82)
Volume 20, March 2025

OPTIMIZATION OF RESOURCE ALLOCATION USING INTEGER PROGRAMMING OF IMPROVED RATIO ESTIMATOR UNDER STRATIFIED RANDOM SAMPLING

Bhatt Ravi Jitendrakumar[1], Monika Saini[2*], Ashish Kumar[3], Yashpal Singh Raghav[4]

•

[1,2,3]Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur, Rajasthan, India.
[4]Department of Mathematics, Jazan University, Jazan, Saudi Arabia.
Emails: [1]ravibhatt0202@gmail.com, [2*]drmnksaini4@gmail.com, [3]ashishbarak2020@gmail.com, [4]yraghav@jazanu.edu.sa

Abstract

This paper provides a case study that illustrates how integer programming may be used to optimize resource allocation. With the known population median of the study variable acting as auxiliary data, an exponential ratio estimator is shown for estimating the finite population mean under stratified random sampling. The objective is to minimize a cost function within specific bounds. Using integer programming techniques and the Lagrange multiplier approach, we transform the proposed problem into an optimization problem with a linear cost function. This allows us to propose an optimal way for minimizing total costs while maintaining desired accuracy levels. We found that the suggested estimator performed better than methods involving stratified random sampling. Additionally, a numerical example is given to verify the theoretical conclusions for real-world applications. We go over how the problem was formulated, how to use LINGO software to solve it, and the results. It is advised to choose the estimator with the lowest MSE in real-world stratified random sampling situations. The strategy shows significant cost savings and efficient use of resources. The effectiveness of the recommended approach is demonstrated by testing the methodology on both simulated and real-world datasets.

Keywords: linear cost function, integer programming, optimization, resource allocation, lingo software, cost minimization

1. Introduction

The problem of effectively estimating the mean of a study variable in the presence of auxiliary information using different sample procedures has been attempted several times in the literature on sampling theory. The problem of creating effective estimators has been thoroughly researched by a number of authors. Regression estimators, products, and ratios are common examples. Stratified random sampling is the suggested sample design for collecting data from a variety of populations due to its low cost and high efficiency. Allocating resources optimally is essential for increasing productivity and cutting expenses in operations research and management science. Because stratified random sampling can yield estimates that are more accurate than those obtained from plain random sampling, it is a widely used technique in statistical surveys. In order to maximize estimate precision within budgetary limits, sample sizes must be distributed among different strata. Conventional methods, like Cochran [1] suggested, make use of continuous optimization techniques, which might not be useful when sample sizes have to be integers. In order to determine the best integer solutions for sample size allocation in stratified random sampling, this work investigates the application of Lagrange multipliers and integer programming. Numerous studies have been conducted on the use of simple random sampling [1, 2, 5].

In order to increase estimate precision, a number of scholars have concentrated on maximizing sample size allocation using auxiliary information [2, 5]. Cochran [4] has discussed a number of sampling strategies, including stratified sampling, systematic sampling, simple random sampling, and others. In the topic of survey sampling, Cochran's work is essential since it offers thorough instructions on various methods. In order to increase the efficiency of population parameter estimation, Bahl and Tuteja [6] presents ratio and product-type exponential estimators. Under some circumstances, the suggested techniques perform better in basic random sampling than conventional estimators. The application of optimization theory to large-scale systems is covered in [3], with a focus on computational and mathematical methods for complex system optimization. Neyman [7] contrasted two techniques: purposive selection, which is a non-probabilistic approach, and stratified sampling, which is a probabilistic approach. In order to guarantee representative samples, author suggested stratified sampling. The optimization problem has been expanded to include linear cost functions in more recent research [8, 10]. By adding integer restrictions to the optimization issue, this work expands on these foundations and offers a more useful solution for real-world scenarios. Shi et al. [9] examines methods based on optimization, fusing theoretical underpinnings with real-world applications. In order to determine the best integer solutions for sample size allocation in stratified random sampling, [10] investigates the application of Lagrange multipliers and integer programming. In stratified sampling, [11] suggest a technique for calculating the interquartile range under a nonlinear cost function. Their method guarantees accurate and economical estimations for all stratified populations. While the method for creating effective stratum borders in stratified sampling while taking survey expenses into consideration is developed in [12]. The technique lowers the overall cost of the survey while improving sampling efficiency. Recently In stratified sampling, the study [14] suggests the best method for determining the population mean under a linear cost function. Comparing the results to current estimators, they show increased cost-effectiveness and accuracy. In [15], a linear cost function is used to present an efficient and cost-effective estimator for the population mean in stratified sampling. Superior efficiency is demonstrated by the approach, which has been confirmed using real-world data. In order to minimize a cost function under predetermined limits, a resource allocation issue is studied using integer programming techniques. We employ LINGO software to determine the best option and show that this strategy works.

2. Material and Methods

The methodology and optimization strategies employed in this work to create and assess an enhanced median based ratio estimator in stratified random sampling under cost functions are described in this part. The integer programming technique and langrage's multiplier technique were used to solve the optimization issue. Furthermore, the suggested estimator's mathematical characteristics, such as its bias and mean squared error (MSE), are calculated and contrasted with those of other estimators.

I. Study Design

- The study variable (Y) and auxiliary variable (X) are used to split the population into four strata. In order to guarantee that the sample sizes are integer values optimized using integer programming and langrage's multiplier technique, a stratified random sampling design is utilized. The suggested optimization method is validated using the real-world dataset, which is derived from census data. Under the restriction of decreasing the overall survey cost while preserving precision, the ideal sample sizes for each stratum are determined.

Four strata are given in the population, one for each research variable (Y) and auxiliary variable (X).

II. Problem Formulation

- The optimization problem is formulated as follows:
- Minimize the objective function:

$$\text{Minimize } \sum_{I=1}^{4} \frac{c_i}{n_i} \tag{1}$$

Subject to the constraints:

$$c_1 = 2,\ c_2 = 3,\ c_3 = 4,\ c_4 = 5$$
$$c_0 = 500$$
$$2 \leq n_h \leq N_h$$
$$h = 1, 2, 3, 4.$$

The suggested approach was used to ascertain the ideal sample sizes using actual data from [`https://censusindia.gov.in/census.website/data/census-tables`]. The findings suggest that when compared to conventional techniques, the integer programming and langrage's methodology produces a more economical use of resources.

3. Solution Techniques

In this instance, a real population from the literature [13] is used to compare the effectiveness of the suggested median-based estimator by [13] with existing estimators. The number of households and the square kilometers of villages and cities, which provide information on study variables and auxiliary variables, respectively, are significant features.

The Neyman allocation is then used to divide the population into four non-crossover strata, and a numerical depiction is finished.

$$n_h = n \frac{N_h S_h}{\sum_{h=1}^{k} N_h S_h}$$

where $i = 1, 2, p$.

Table 1: *Data statistics (source: [13])*

Population ($N = 645$; $h = 4$)											
H	N_h	n_h	$\bar{Y}_h$	M_h	C_{yh}^2	C_{ymh}	C_{mh}^2	S_{yh}	S_{ymh}	λ_h	θ_h
1	237	4.13025	116.236	116.81	0.31485	0.20065	0.14554	65.2218	2724.33	0.2379	1.37869
2	164	5.78153	307.603	292.295	0.18397	0.14238	0.30406	131.936	12801	0.16687	0.46825
3	90	16.8718	547.444	548.77	1.64244	2.49501	3.84895	701.592	749552	0.04816	0.64823
4	154	68.2164	757.1	727.165	4.79469	6.20317	8.78042	1657.81	3415068	0.00817	0.70648

Table 2: *MSE values of different estimators*

Estimators	MSE
$\mu_{0(st)}$ Stratified	50064.21813
$\mu_{p(st)}$ Bahl and Tuteja 1991	298413.7926
$\mu_{pe(st)}$ Bahl and Tuteja 1991	156446.8056
$\mu_{1(st)}$ Kadilar and Cingi 2004	73914.17572
$\mu_{2(st)}$ Kadilar and Cingi 2004	73610.17851
$\mu_{3(st)}$ Kadilar and Cingi 2004	73581.24647
$\mu_{4(st)}$ Kadilar and Cingi 2004	73764.64679
$\mu_{5(st)}$ Kadilar and Cingi 2004	73585.78191
$\mu_{6(st)}$ Kadilar and Cingi 2004	73794.12042
$\mu_{7(st)}$ Kadilar and Cingi 2004	73599.63542
$\mu_{8(st)}$ Kadilar and Cingi 2004	73841.96289
$\mu_{9(st)}$ Kadilar and Cingi 2004	73572.84621
$\mu_{10(st)}$ Kadilar and Cingi 2004	73824.07952
$\mu_{11(st)}$ Kadilar and Cingi 2004	73291.58656
$\mu_{12(st)}$ Kadilar and Cingi 2004	46271.34602
$\mu_{subr(st)}$ Subramani 2016	17357.5585
$\mu_{CR(st)}$ Cochran estimator 1940	8660.837079
$\mu_{**(st)}$ Yadav 2019	6020.730985
$\mu_{prop(st)}$ Estimator	4267.075487

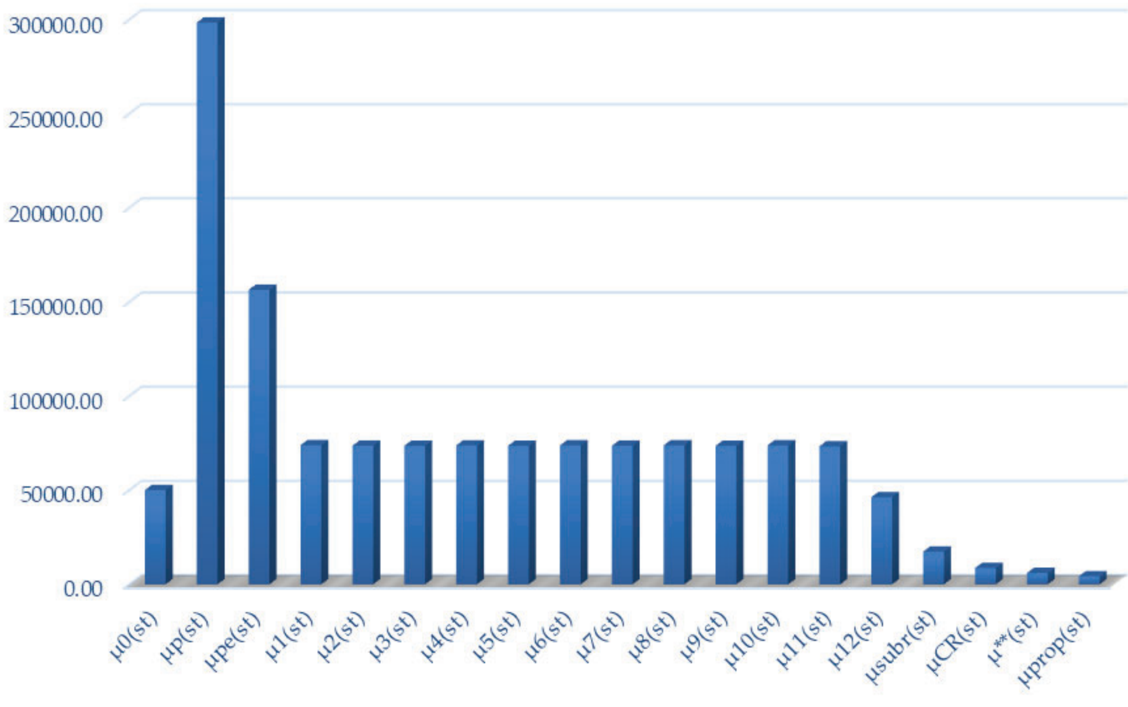

Figure 1: *Standard MSE*

Table 3: *PRE of different estimators*

Estimators	PRE
$\mu_{0(st)}$ Stratified	100
$\mu_{CR(st)}$ Cochran estimator 1940	578.0529
$\mu_{p(st)}$ Bahl and Tuteja 1991	16.77678
$\mu_{pe(st)}$ Bahl and Tuteja 1991	32.00079
$\mu_{1(st)}$ Kadilar and Cingi 2004	67.7329
$\mu_{2(st)}$ Kadilar and Cingi 2004	68.01263
$\mu_{3(st)}$ Kadilar and Cingi 2004	68.03937
$\mu_{4(st)}$ Kadilar and Cingi 2004	67.87021
$\mu_{5(st)}$ Kadilar and Cingi 2004	68.03518
$\mu_{6(st)}$ Kadilar and Cingi 2004	67.8431
$\mu_{7(st)}$ Kadilar and Cingi 2004	68.02237
$\mu_{8(st)}$ Kadilar and Cingi 2004	67.79914
$\mu_{9(st)}$ Kadilar and Cingi 2004	68.04714
$\mu_{10(st)}$ Kadilar and Cingi 2004	67.81557
$\mu_{11(st)}$ Kadilar and Cingi 2004	68.30827
$\mu_{12(st)}$ Kadilar and Cingi 2004	108.197
$\mu_{subr(st)}$ Subramani 2016	288.4289
$\mu_{**(st)}$ Yadav 2019	831.5306
$\mu_{prop(st)}$ Estimator	1173.268

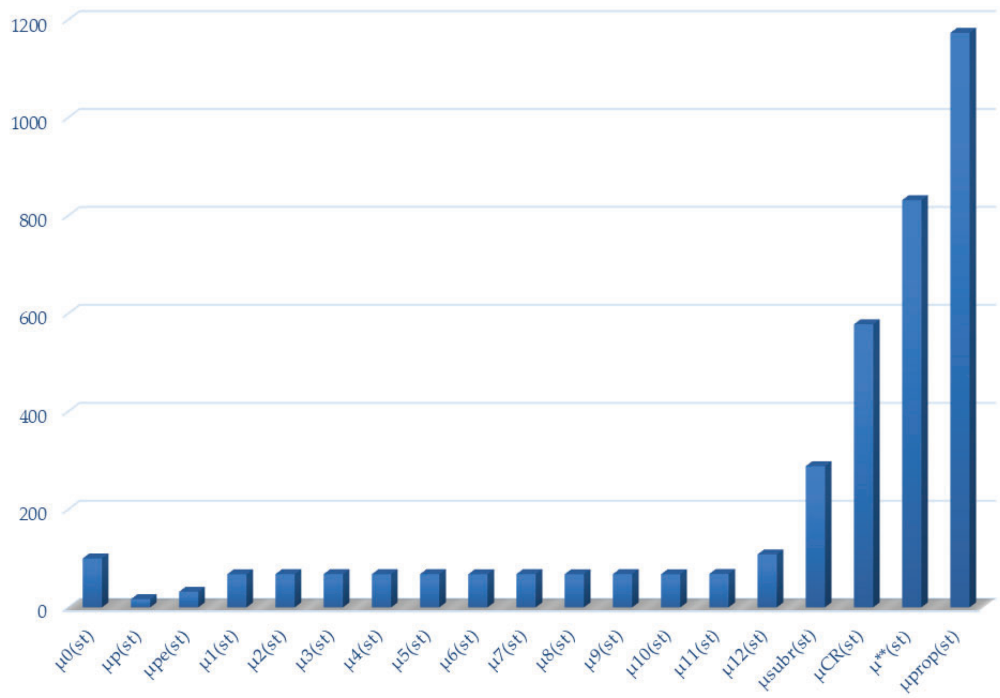

Figure 2: *Standard PRE*

The comparison of the proposed estimator with existing estimators utilizing stratified random sampling, Tables 2 and 3 unequivocally demonstrate that the proposed estimator has the greatest PRE and the lowest MSE value and their graphs were also given as Figure 1 and 2.

4. Cost Function

The main factor that influences of the number of samples across strata is survey expenditure. [8] introduced linear cost and fixed total cost C_0 of the survey as a linear function of n_h; $h = 1, 2, \ldots, L$.

$$C_0 = \sum_{h=1}^{L} c_h n_h \tag{2}$$

where c_h denotes the cost per unit of measuring each characteristic in the hth stratum.; $h = 1, 2, \ldots, L$. In this instance, our goal is to determine the fixed linear cost function's least mean square error. Thus, the optimization issue for the proposed estimator in [9] may be described as follows:

$$\begin{aligned}
&\text{Minimize } MSE(t_{pr(st)}) \\
&\text{subject to } \sum_{h=1}^{L} c_h n_h \leq C_0 \\
&\qquad\qquad 2 \leq n_h \leq N_h \\
&\text{and} \qquad n_h \text{ are integers;} \quad h = 1, 2, \ldots, L.
\end{aligned}$$

Using the cost function, the mean square error will now be

$$\widehat{\mu}_{P(st)_{\min}} = \sum_{h=1}^{L} \bar{Y}_h^2 \left(\frac{1 - f_h}{n_h} \right) [C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}]. \tag{3}$$

Integer Programing and Lagrange's Multiplier Technique

Integer Programing:

With a constant linear cost function and actual data, we get the least mean square error. This allows the optimization issue to be stated as follows:

$$\text{Minimize } \frac{507.3364037}{n_1} + \frac{10707.94895}{n_2} + \frac{6113.182684}{n_3} + \frac{131645.0146}{n_4}$$

Subject to

$$\begin{aligned}
&\sum_{h=1}^{L} c_h n_h \leq C_0 \\
&c_1 n_1 + c_2 n_2 + c_3 n_3 + c_4 n_4 \leq C_0 \\
&2n_1 + 3n_2 + 4n_3 + 5n_4 \leq 500
\end{aligned}$$

Bounds on variables:

$$\begin{aligned}
&2 \leq n_h \leq N_h \\
&\textit{and } n_h \textit{ are integers;} \ \ h = 1, 2, 3, 4 \\
&2 \leq n_1 \leq 237, \ \ 2 \leq n_2 \leq 164 \\
&2 \leq n_3 \leq 90, \ \ 2 \leq n_4 \leq 154
\end{aligned}$$

The Lagrange multiplier method produces an optimality criterion in some applications. Additionally, the conditions are suitable to set a minimum or maximum. Therefore, the most optimal n value may be found using the Lagrange multiplier method.

The Lagrange function is so defined as:

$$L(x, \lambda) = f(x) - \lambda g(x),$$

where L = Lagrangian, λ = Lagrange multiplier, $f(x)$ = Function, x = integer.
Now

$$L(n_h, \lambda) = MSE + \lambda \left(\sum_{h=1}^{L} C_h n_h - C_0 \right)$$

$$L = \sum_{h=1}^{L} \bar{Y}_h^2 \left(\frac{1-f_h}{n_h} \right) [C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}] + \lambda \left(\sum_{h=1}^{L} C_h n_h - C_0 \right). \quad (4)$$

Now let us partially differentiate the above equation (4) with respect to n_h, we get

$$\frac{dL}{dn_h} = 0$$

$$\frac{d \left(\sum_{h=1}^{L} \bar{Y}_h^2 \left(\frac{1-f_h}{n_h} \right) [C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}] + \lambda \left(\sum_{h=1}^{L} C_h n_h - C_0 \right) \right)}{dn_h} = 0$$

Then

$$n_h = \sqrt{\frac{\bar{Y}_h^2 (1-f_h)(C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}})}{\lambda C_h}}.$$

Again, differentiate the equation (4) with respect to λ, we get

$$\frac{dL}{d\lambda} = 0$$

$$\frac{d \left(\sum_{h=1}^{L} \bar{Y}_h^2 \left(\frac{1-f_h}{n_h} \right) [C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}] + \lambda \left(\sum_{h=1}^{L} C_h n_h - C_0 \right) \right)}{d\lambda} = 0$$

Using the value of equation (4) after differentiating above equation, we get

$$\sqrt{\lambda} = \frac{\sqrt{\bar{Y}_h^2 (1-f_h)(C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}) C_h}}{C_0}. \quad (5)$$

Now putting the value of equation (5) in equation (4) to find out the value of n_h, we get

$$n_h = \frac{C_0 \sqrt{\bar{Y}_h^2 (1-f_h)(C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}})}}{\sqrt{(\bar{Y}_h^2 (1-f_h)(C_{y_h}^2 + {\theta_h}^2 C_{mh}^2 - 2\theta_h C_{y_{mh}})) C_h^2}}$$

$$n_h = \frac{C_0}{C_h}.$$

5. Empirical Study with Cost Function

In this part, we prove the efficiency of the proposed estimator using the real data set. The actual population as reported by the Indian census conducted in Lucknow, Uttar Pradesh, is taken into account in the data set (`https://censusindia.gov.in/census.website/data/census-tables`). The data $N = 645$, $h = 4$, which were used to apply the recommended estimator, contain information on the number of households and the area in square kilometers of certain cities and villages, respectively. These details provide information on the auxiliary variable and the variable under investigation. The population is then split up into four distinct, non-overlapping strata. Integer programming and Lagrange multiplier approaches have been used in numerical illustration. A reference to the data summary may be found in Table 1. When variables in an optimization problem have to handle integer values, the problem is known as integer programming. If all of the functions are linear, then an integer linear programming problem can be considered. Now, using real data and a fixed linear cost function, we can calculate the least mean square error. Next, the following is a description of the optimization scenario:

Problem Formulation of Proposed Estimator
Objective Function

$$\widehat{\mu}_{P(st)_{\min}} = \sum_{h=1}^{k} \bar{Y}_h^2 \delta_h [C_{y_h}^2 + {\theta_h}^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}]$$

Limited population factor will be ignored,

$$\widehat{\mu}_{P(st)_{\min}} = \sum_{h=1}^{k} \bar{Y}_h^2 \frac{1}{n_h} [C_{y_h}^2 + \theta_h^2 C_{mh}^2 - 2\theta_h C_{y_{mh}}].$$

The objective is to minimize the cost function defined as:

$$\text{Minimize } \frac{507.3364037}{n_1} + \frac{10707.94895}{n_2} + \frac{6113.182684}{n_3} + \frac{131645.0146}{n_4}$$

Subject to

$$\sum_{h=1}^{L} c_h n_h \leq C_0$$
$$c_1 n_1 + c_2 n_2 + c_3 n_3 + c_4 n_4 \leq C_0$$
$$2n_1 + 3n_2 + 4n_3 + 5n_4 \leq 500$$

Bounds on variables:

$$2 \leq n_h \leq N_h$$
$$\textit{and } n_h \textit{ are integers; } h = 1, 2, 3, 4$$
$$2 \leq n_1 \leq 237, \;\; 2 \leq n_2 \leq 164$$
$$2 \leq n_3 \leq 90, \;\; 2 \leq n_4 \leq 154$$

We apply integer programming techniques along with the Lagrange multiplier approach to solve this optimization issue. To determine the best integer values for the sample sizes, the LINGO program is used. Integer variables are used in the model formulation to represent resource allocations, together with an objective function to minimize costs and restrictions to guarantee workable solutions. The variables' ideal values were determined to be n_1, n_2, n_3, and n_4.

These numbers show effective resource allocation by minimizing the cost function while meeting all restrictions.

Table 4: *Optimized MSE and PRE of different estimators using integer programming*

Population $(N, h) = (645, 4)$							
Estimators	n_1	n_2	n_3	n_4	n	*MSE*	*PRE*
$\mu_{0(st)}$ Stratified	5	9	37	63	114	37811.71301	100
$\mu_{p(st)}$ Bahl and Tuteja 1991	4	6	26	74	110	254647.4931	14.84864922
$\mu_{pe(st)}$ Bahl and Tuteja 1991	5	6	28	72	111	132124.2009	28.6183097
$\mu_{1(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62734.46491	60.27263173
$\mu_{2(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62470.09829	60.52769892
$\mu_{3(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62444.95167	60.55207345
$\mu_{4(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62606.61364	60.3957167
$\mu_{5(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62439.33502	60.55752034
$\mu_{6(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62632.11269	60.37112814
$\mu_{7(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62460.7517	60.53675624
$\mu_{8(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62675.90102	60.32895002
$\mu_{9(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62437.71155	60.55909492
$\mu_{11(st)}$ Kadilar and Cingi 2004	4	6	26	74	110	62233.04837	60.75825305
$\mu_{12(st)}$ Kadilar and Cingi 2004	4	5	28	73	110	38419.51093	98.41799671
$\mu_{subr(st)}$ Subramani 2016	4	16	36	60	116	12234.51456	309.0577304
$\mu_{CR(st)}$ Cochran estimator 1940	2	3	18	83	106	7361.991722	513.6071112
$\mu_{**(st)}$ Yadav19	5	20	15	74	114	4412.062967	857.0075561
$\mu_{prop(st)}$ Estimator	7	26	17	68	118	2779.875899	1360.194282

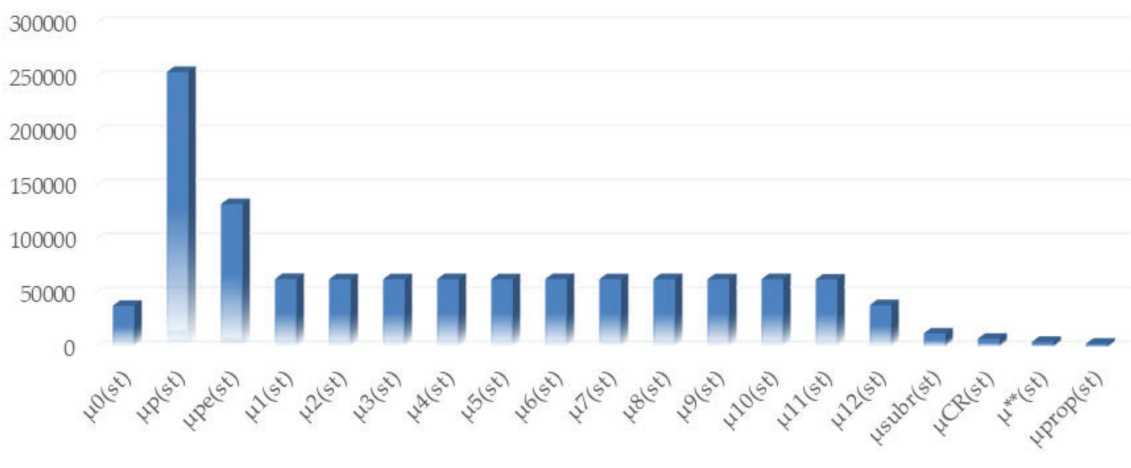

Figure 3: *Optimized MSE integer programming*

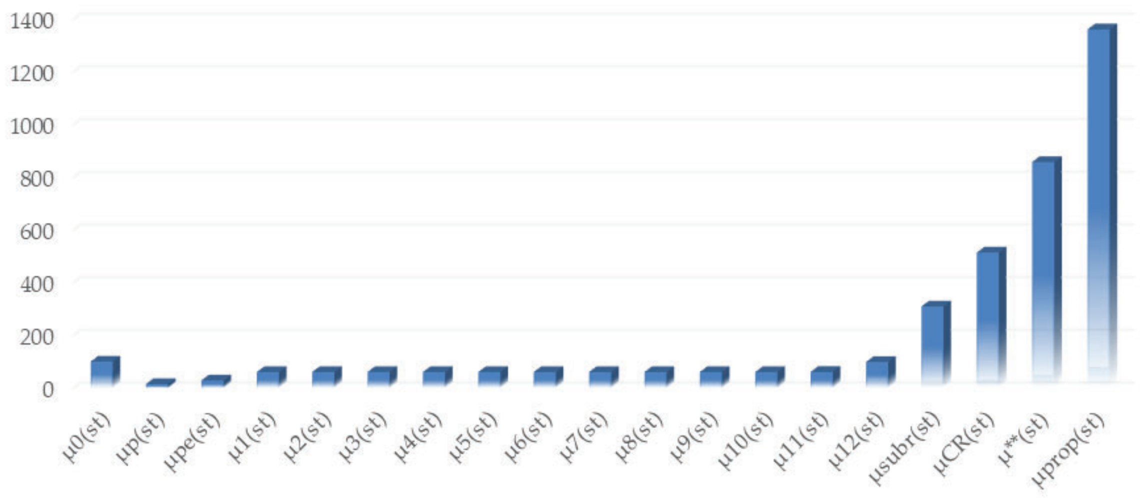

Figure 4: *Optimized PRE integer programming*

In Table 4 the optimized MSE and PRE using integer programing technique is give along with their graphs in Figure 3 and 4.

Table 5: *MSE and PRE comparison of different estimators (standard vs integer)*

Estimators	*MSE*	*Optimized MSE*	*PRE*	*Optimized PRE*
$\mu_{0(st)}$ Stratified	50064.21813	37811.71301	100	100
$\mu_{p(st)}$ Bahl and Tuteja 1991	298413.7926	254647.4931	16.77678	14.84864922
$\mu_{pe(st)}$ Bahl and Tuteja 1991	156446.8056	132124.2009	32.00079	28.6183097
$\mu_{1(st)}$ Kadilar and Cingi 2004	73914.17572	62734.46491	67.7329	60.27263173
$\mu_{2(st)}$ Kadilar and Cingi 2004	73610.17851	62470.09829	68.01263	60.52769892
$\mu_{3(st)}$ Kadilar and Cingi 2004	73581.24647	62444.95167	68.03937	60.55207345
$\mu_{4(st)}$ Kadilar and Cingi 2004	73764.64679	62606.61364	67.87021	60.3957167
$\mu_{5(st)}$ Kadilar and Cingi 2004	73585.78191	62439.33502	68.03518	60.55752034
$\mu_{6(st)}$ Kadilar and Cingi 2004	73794.12042	62632.11269	67.8431	60.37112814
$\mu_{7(st)}$ Kadilar and Cingi 2004	73599.63542	62460.7517	68.02237	60.53675624
$\mu_{8(st)}$ Kadilar and Cingi 2004	73841.96289	62675.90102	67.79914	60.32895002
$\mu_{9(st)}$ Kadilar and Cingi 2004	73572.84621	62437.71155	68.04714	60.55909492
$\mu_{10(st)}$ Kadilar and Cingi 2004	73824.07952	62653.90433	67.81557	60.35013047
$\mu_{11(st)}$ Kadilar and Cingi 2004	73291.58656	62233.04837	68.30827	60.75825305
$\mu_{12(st)}$ Kadilar and Cingi 2004	46271.34602	38419.51093	108.197	98.41799671
$\mu_{subr(st)}$ Subramani 2016	17357.5585	12234.51456	288.4289	309.0577304
$\mu_{CR(st)}$ Cochran estimator 1940	8660.837079	7361.991722	578.0529	513.6071112
$\mu_{**(st)}$ Yadav19	6020.730985	4412.062967	831.5306	857.0075561
$\mu_{prop(st)}$ Estimator	4267.075487	2779.875899	1173.268	1360.194282

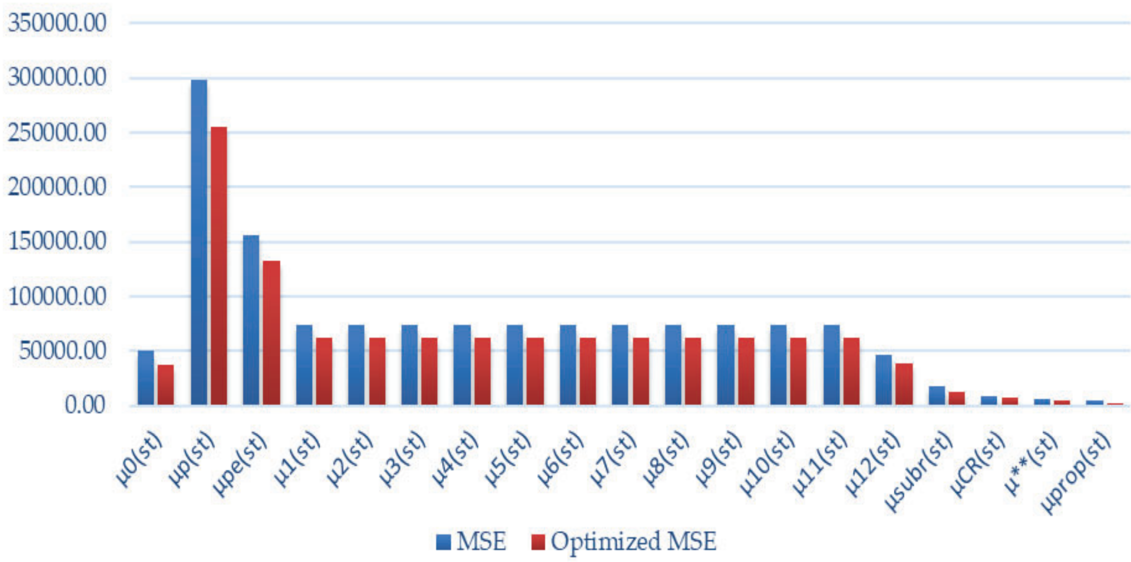

Figure 5: *MSE Comparison (standard vs integer)*

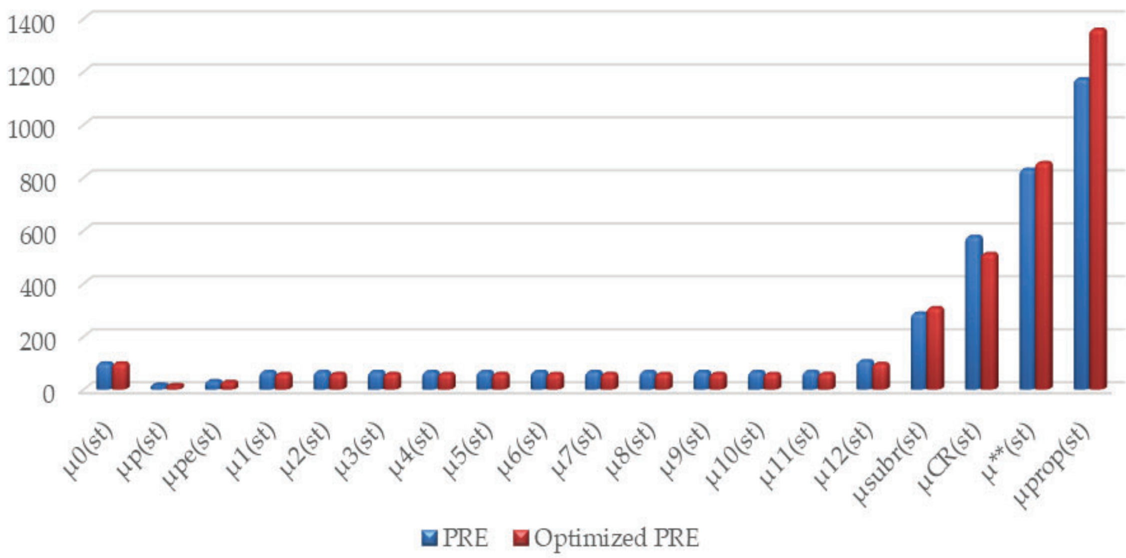

Figure 6: *PRE Comparison (standard vs integer)*

In Table 5 the comparison of MSE and PRE of existing estimator and proposed estimator using integer programming technique is given with their graphs as Figure 5 and 6.

Table 6: *Optimized MSE and PRE of different estimators using lagrange's multiplier technique*

Population $(N, h) = (645, 4)$							
Estimators	n_1	n_2	n_3	n_4	n	*MSE*	*PRE*
$\mu_{0(st)}$ Stratified	4	61	24	43	131	37800.05	100
$\mu_{p(st)}$ Bahl and Tuteja 1991	4	6	25	74	110	254555.8	14.8
$\mu_{pe(st)}$ Bahl and Tuteja 1991	4	6	27	73	111	132051	28.6
$\mu_{1(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.7	110	62716.0	60.3
$\mu_{2(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.7	110	62452.1	60.5
$\mu_{3(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.8	110	62427.6	60.6
$\mu_{4(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.7	110	62588.4	60.4
$\mu_{5(st)}$ Kadilar and Cingi 2004	4.3	5.5	26.3	73.9	110	62418.1	60.6
$\mu_{6(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.8	110	62614.9	60.4
$\mu_{7(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.7	110	62442.7	60.5
$\mu_{8(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.7	110	62658.1	60.3
$\mu_{9(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.8	110	62420.4	60.6
$\mu_{10(st)}$ Kadilar and Cingi 2004	4.4	5.7	26.3	73.8	110	62635.5	60.3
$\mu_{11(st)}$ Kadilar and Cingi 2004	4.3	5.7	26.2	73.9	110	62220.1	60.8
$\mu_{12(st)}$ Kadilar and Cingi 2004	3.9	5.7	29.2	71.6	110	38373.7	98.5
$\mu_{subr(st)}$ Subramani 2016	4.0	16.6	35.1	60.4	116	12230.5	309.1
$\mu_{CR(st)}$ Cochran estimator 1940	2.0	3.4	18.0	82.8	106	7360.1	513.6
$\mu_{**(st)}$ Yadav19	5.4	20.5	14.6	73.9	114	4410.8	857.0
$\mu_{prop(st)}$ Estimator	6.8	25.3	16.6	68.8	118	2778.996	1360.2

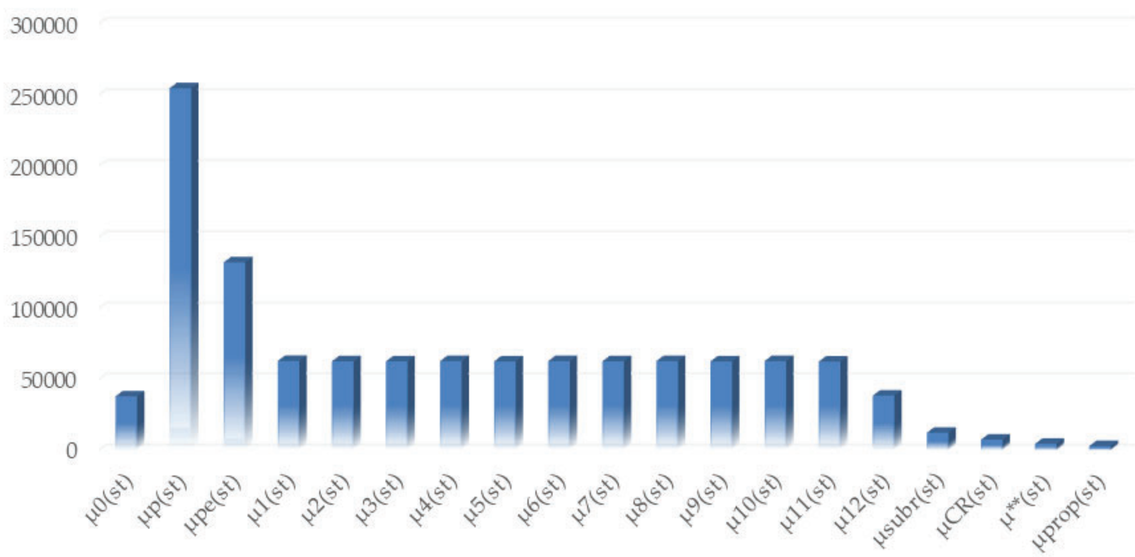

Figure 7: *Optimized MSE lagrange's multiplier*

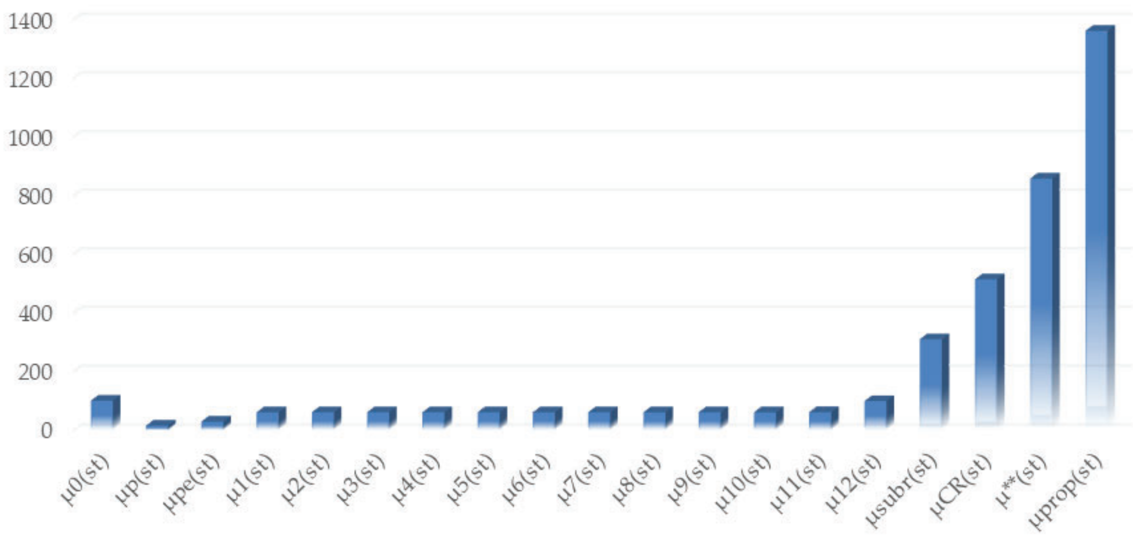

Figure 8: *Optimized PRE lagrange's multiplier*

Similarly in Table 6 the optimized MSE and PRE using language's multiplier technique is give along with their graphs as Figure 7 and 8.

Table 7: *MSE and PRE comparison of different estimators (standard vs lagrange's)*

Estimators	*MSE*	*Optimized MSE*	*PRE*	*Optimized PRE*
$\mu_{0(st)}$ Stratified	50064.21813	37800.05	100	100
$\mu_{p(st)}$ Bahl and Tuteja 1991	298413.7926	254555.8	16.8	14.8
$\mu_{pe(st)}$ Bahl and Tuteja 1991	156446.8056	132050.8	32.0	28.6
$\mu_{1(st)}$ Kadilar and Cingi 2004	73914.17572	62715.97	67.7	60.3
$\mu_{2(st)}$ Kadilar and Cingi 2004	73610.17851	62452.11	68.0	60.5
$\mu_{3(st)}$ Kadilar and Cingi 2004	73581.24647	62427.61	68.0	60.6
$\mu_{4(st)}$ Kadilar and Cingi 2004	73764.64679	62588.4	67.9	60.4
$\mu_{5(st)}$ Kadilar and Cingi 2004	73585.78191	62418.13	68.0	60.6
$\mu_{6(st)}$ Kadilar and Cingi 2004	73794.12042	62614.93	67.8	60.4
$\mu_{7(st)}$ Kadilar and Cingi 2004	73599.63542	62442.74	68.0	60.5
$\mu_{8(st)}$ Kadilar and Cingi 2004	73841.96289	62658.06	67.8	60.3
$\mu_{9(st)}$ Kadilar and Cingi 2004	73572.84621	62420.37	68.0	60.6
$\mu_{10(st)}$ Kadilar and Cingi 2004	73824.07952	62635.47	67.8	60.3
$\mu_{11(st)}$ Kadilar and Cingi 2004	73291.58656	62220.05	68.3	60.8
$\mu_{12(st)}$ Kadilar and Cingi 2004	46271.34602	38373.67	108.2	98.5
$\mu_{subr(st)}$ Subramani 2016	17357.5585	12230.45	288.4	309.1
$\mu_{CR(st)}$ Cochran estimator 1940	8660.837079	7360.138	578.1	513.6
$\mu_{**(st)}$ Yadav19	6020.730985	4410.764	831.5	857.0
$\mu_{prop(st)}$ Estimator	4267.075487	2778.996	1173.3	1360.2

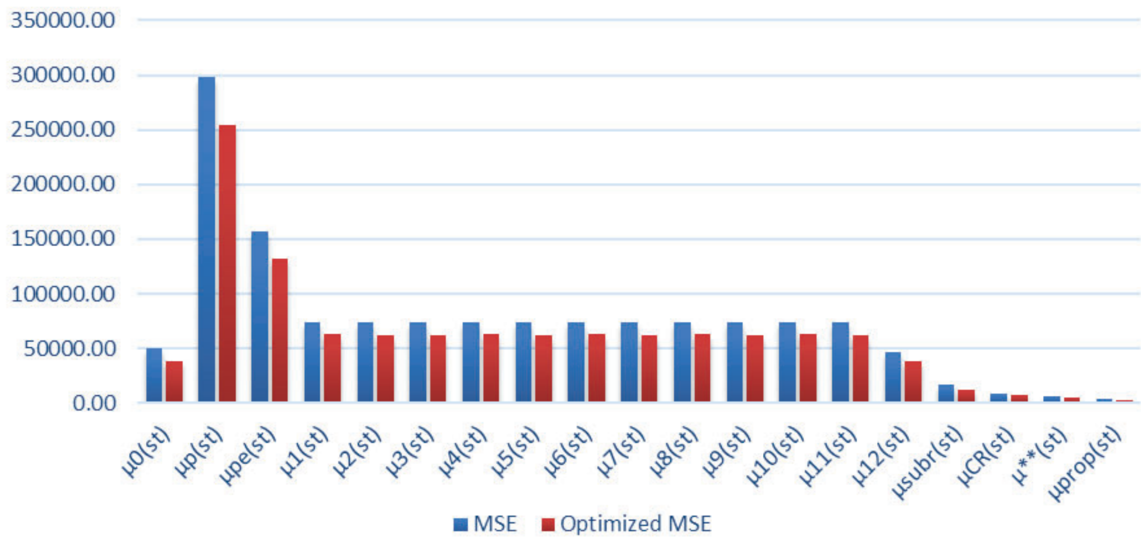

Figure 9: *MSE Comparison (standard vs lagrange's)*

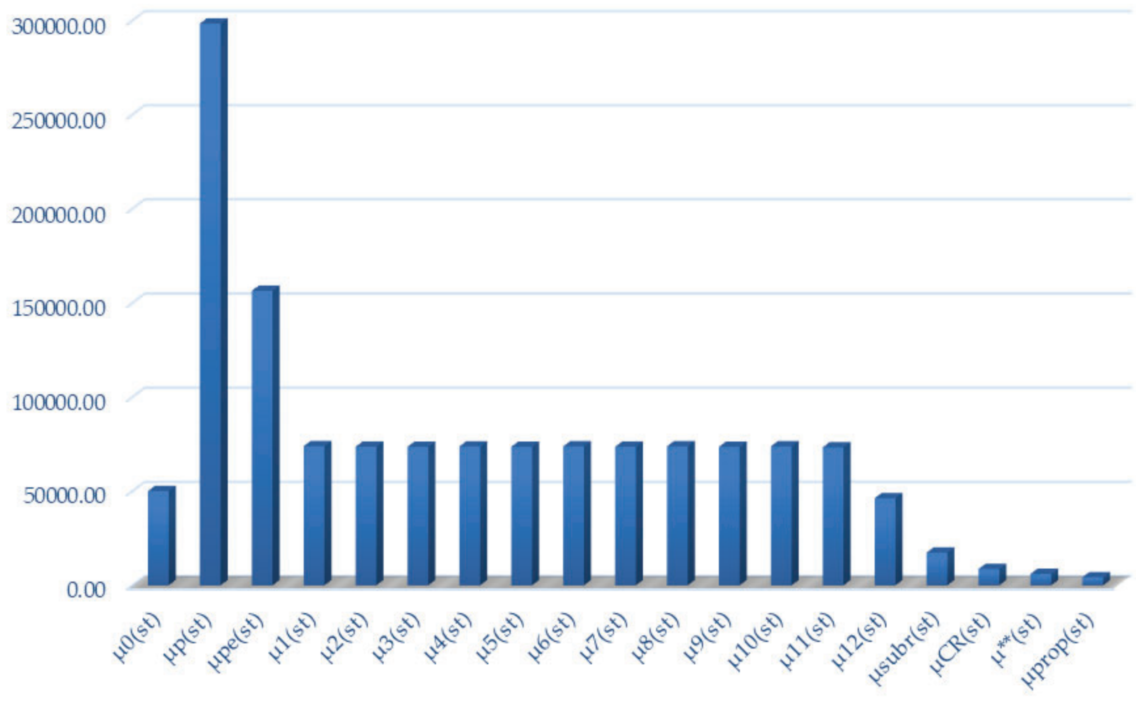

Figure 10: *PRE Comparison (standard vs lagrange's)*

In Table 7 the comparison of MSE and PRE of existing estimator and proposed estimator using langrage's multiplier technique is given with their graphs as Figure 9 and 10.

6. Discussion and Conclusion

In this study, we optimized a new median-based ratio estimator for restricted population means estimation under stratified random sampling. Up to the first level of approximation, bias and MSE formulas are created for the suggested estimators. The suggested estimator was compared theoretically to existing estimators. We determined the conditions in which the suggested estimator performs better than the traditional estimators. We compare the performance of the proposed estimator quantitatively, considering a real population. The suggested estimator consistently performs better than the existing estimators under stratified random sampling with cost function, both theoretically and numerically. Considering these results, we advise future research to employ the proposed estimator for effective population mean estimation when supplementary data is available. The results indicate a significant reduction in costs through optimal resource allocation. The integer programming and langrage's approach ensures that solutions are both feasible and practical. This methodology can be applied to similar problems in various industries for improved operational efficiency. The problem was successfully solved with the help of LINGO software, which offered a workable solution with either minimizing cost or maximizing precision. To further improve resource allocation tactics, future study might investigate more intricate models and other optimization methodologies.

References

[1] Cochran, W.G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, 30(2):262–275.

[2] Murthy, M.N. (1964). Product method of estimation. *Sankhyā: The Indian Journal of Statistics*, Series A, 69–74.

[3] Lasdon, L.S. (1970). Optimization theory for large systems. Macmillan Company, London, England, 1(9):7.

[4] Cochran, W.G. (1977). Sampling techniques, John Wiley & Sons.

[5] Srivastava, S.K. (1980). A class of estimators using auxiliary information for estimating finite population variance. *Sankhya*, C, 42:87–96.

[6] Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimators. *Information and Optimization Sciences*, 12:159-163.

[7] Neyman, J. (1992). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. In *Breakthroughs in Statistics: Methodology and Distribution*, New York, NY: Springer New York, 123–150.

[8] Khan, M.G., Khan, E.A. and Ahsan, M.J. (2008). Optimum allocation in multivariate stratified sampling in presence of non-response. *Journal of the Indian Society of Agricultural Statistics*, 62(1):42–48.

[9] Shi, Y., Tian, Y.J., Kou, G., Peng, Y. and Li, J.P. (2011). Optimization based data mining: Theory and applications. Springer, Berlin.

[10] Mradula, Yadav, S.K., Varshney, R. and Dube, M. (2021). Efficient estimation of population mean under stratified random sampling with linear cost function. *Communications in Statistics-Simulation and Computation*, 50(12):4364–4387.

[11] Shabbir, J. and Ahmed, A. (2022). Estimation of interquartile range in stratified sampling under non-linear cost function. *Communications in Statistics - Simulation and Computation*, 51(4):1891-1898.

[12] Reddy, K.G. and Khan, M.G.M. (2023). Constructing efficient strata boundaries in stratified sampling using survey cost. Heliyon, 6:e21407.

[13] Saini, M., Jitendrakumar, B.R. and Kumar, A. (2024). Improved ratio estimator under simple and stratified random sampling. *Life Cycle Reliability and Safety Engineering*, 13(2):181–187.

[14] Yadav, S.K., Verma, M.K. and Varshney, R. (2024). Optimal strategy for elevated estimation of population mean in stratified random sampling under linear cost function. Annals of Data Science.

[15] Zaagana, A.A., Verma, M.K., Mahnashi, A.M., Yadav, S.K., Ahmadini, A.A.H., Meetei, M.Z. and Varshney, R. (2024). An effective and economic estimation of population mean in stratified random sampling using a linear cost function. Heliyon, 10(10):e31291.