A CERTAIN STUDY ON ESTIMATION OF POPULATION PARAMETERS IN SURVEY SAMPLING

ABSTRACT OF THESIS

SUBMITTED TO
BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
(A CENTRAL UNIVERSITY)
LUCKNOW

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
LUCKNOW
प्रज्ञा शील करुणा
ESTABLISHED 1996

FOR THE AWARD OF DEGREE OF

Doctor of Philosophy

IN
STATISTICS

Submitted by

Diksha Arya
Enrolment No.: 280/17

Under the Supervision of
Dr. Subhash Kumar Yadav
Associate Professor

DEPARTMENT OF STATISTICS
SCHOOL OF PHYSICAL & DECISION SCIENCE
BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY
(A CENTRAL UNIVERSITY) (NAAC: A++ Accredited)
VIDYA VIHAR, RAEBARELI ROAD, LUCKNOW-226025
UTTAR PRADESH, INDIA

2024

ABSTRACT

Chapter 1 offers a concise overview of sampling theory and underscores the significance of auxiliary information in estimating population parameters. It examines different estimators for population mean and variance proposed by researchers. The summary of the first chapter is given in the following paragraphs.

Information can be acquired in two main ways: comprehensively, where every detail is considered, or in summary, where the essence of the overall details is provided. In specific terms, the former is referred to as Complete Enumeration, while the latter is termed as Sampling. Sampling, an inherent aspect of human life, involves actions such as inspecting a handful of rice to assess the quality of an entire sack, or meeting a few individuals from a country to understand the nature of its people, as well as blood group testing. Since a sample represents only a portion of the whole, its extrapolation can lead to errors known as Sampling Errors. Statisticians aim to minimize these errors by devising suitable sampling schemes or by creating efficient estimators of the parameters, or both. In our thesis, we have developed efficient estimators for population parameters. The simplest sampling method is the Simple Random Sampling (SRS) Scheme, where each unit of the population has an equal chance of being selected for the sample, either with replacement (WR) or without replacement (WOR). In our thesis, we have focused on the SRSWOR Scheme.

Sampling proves to be a cost-effective approach for making valid inferences about population parameters, such as mean, variance, skewness, and kurtosis. We have particularly studied the estimation of the population mean and population variance. If we denote the characteristic under study as Y. Utilizing auxiliary information, which is

independent of the characteristic under study but holds some level of association with it, enhances parameter estimation efficiency. For instance, in analyzing annual household expenditure, household size can be employed as auxiliary information.

The significance and evolution of sample surveys have been extensively discussed by various experts, including Yates (1949), Mahalanobis (1952), Hansen *et al.* (1953a, 1953b), Yates (1960), Raj (1968), Des Raj (1972), Cassel *et al.* (1977), Cochran (1977), Singh and Chaudhary (1986), and Das (1988). They have eloquently highlighted the advantages of sample surveys, such as cost-efficiency, quicker results, enhanced result accuracy, broader applicability, and universal adaptability.

Ratio estimators are among the most commonly employed methods for estimating population mean or total, especially when there exists a positive correlation between the auxiliary and study variable. Conversely, when the correlation is negative, the product estimator is utilized. Despite the higher precision of the regression estimator compared to the ratio estimator, in large-scale surveys, opting for the ratio estimator is often preferable due to its simplicity.

In cases of positive correlation, Cochran (1940 and 1942) devised a ratio method for estimating the population mean or total. In cases characterized by a negative correlation between auxiliary and study variable, researchers like Robson (1957) and Murthy (1964) suggested the use of product estimators for estimating the population mean. Watson (1937) introduced a straightforward linear regression estimator, which he used to estimate the average leaf area by considering the weight of leaves as an auxiliary variable.

The utilization of auxiliary information can occur at various stages, including the design stage, sampling stage, or estimation stage. Our current study specifically focuses on the application of auxiliary information at the estimation stage. Numerous estimators have been developed over time to enhance efficiency using auxiliary variables. Parameters of the auxiliary variables, such as mean, median, coefficient of variation, etc., are often leveraged to increase the efficiency of the estimators. Many such estimators have been discussed in the existing literature.

Goodman (1958) improved ratio estimators to ensure unbiasedness in standard simple random sampling. Searls (1964) used the population coefficient of variation for specific variable analysis, while Sisodia and Dwivedi (1981) applied it to auxiliary variables. Srivastava (1967) introduced a generalized ratio-type estimator incorporating auxiliary variables for population mean estimation. Reddy (1974) proposed ratio-type estimators utilizing auxiliary variables for mean estimation. Chakrabarty (1979) developed efficient ratio estimators for mean estimation under optimal conditions. Sahai (1980) suggested two-parameter ratio-type and product-type estimators for finite population mean estimation. Bahl and Tuteja (1991) introduced exponential ratio-type and product-type estimators incorporating positively and negatively correlated auxiliary variables. Upadhyaya and Singh (1999a, 1999b) used a transformed auxiliary variable to estimate the population mean and introduced various ratio and product type estimators, accounting for coefficients of variation and kurtosis. Kadilar and Cingi (2003) analyzed the chain ratio-type estimator, revealing its superior efficiency over the traditional ratio estimator under specific conditions.

Singh *et al.* (2004) proposed a modified ratio estimator incorporating the coefficient of kurtosis for an auxiliary variable to improve efficiency. They derived first-order approximations for bias and MSE, comparing it with other estimators. Singh and Tailor (2005), Tailor and Sharma (2009) developed ratio-cum-product estimators. Yan and Tian (2010) introduced ratio-type estimators based on skewness, showing improved efficiency under certain conditions compared to traditional ones. Al-Omari (2012) proposed two enhanced ratio estimators for population mean estimation, demonstrating the establishment of quartiles of an auxiliary variable when its mean is known.

Jeelani *et al.* (2013) proposed an enhanced ratio-type estimator using skewness and quartile difference of an auxiliary variable to improve estimation accuracy. Swain (2014) presented an alternative ratio-type exponential estimator, comparing it with Bahl and Tuteja's exponential ratio-type estimator and the classical ratio estimator using large sample approximations to assess bias and MSE. Jerjuddin and Kishun (2016) introduced an enhanced ratio estimator for population mean estimation using simple random sampling without replacement (SRSWOR). Ijaz and Ali (2018) developed an efficient estimator for determining the mean of a finite population using simple random sampling. Yadav et al. (2019) focused on estimating the average peppermint production in Barabanki district, India, suggesting various estimators for the population mean using field area as an auxiliary variable. Baghel and Yadav (2020) proposed a method for population mean estimation using an auxiliary variable within simple random sampling. They introduced a new category of estimators, calculating bias and MSE up to the first order of approximation. Yadav and Baghel (2021) introduced an innovative class of estimators utilizing known auxiliary variable information to enhance population mean estimation in simple random sampling. Aggarwal *et al.* (2022) worked on improving the estimation of $\bar{Y}$ in SRS by combining two already existing ratio estimators. Ali *et al.* (2023) introduced a class of estimators for estimating population mean using known auxiliary parameters. Yadav *et al.* (2023) utilized known population median of study variable for elevated estimation of population mean. Gupta *et al.*(2024) improved the estimation of the population mean by integrating additional auxiliary parameters. Yadav *et al.* (2024) proposed an adroit family of estimators of population mean using known auxiliary parameters. Yadav *et al.* (2024) introduced new optimal class of searls estimators of population mean using powered auxiliary parameters.

Certainly, numerous survey statisticians have contributed significantly to the advancement of variance estimation methods by incorporating additional information pertaining to auxiliary attributes. Liu (1974) pioneered the estimation of population variance by incorporating auxiliary information and formulated a set of unbiased estimators. Singh and Pandey (1983) constructed ratio and product type estimators for variance estimation. Isaki (1983) further addressed the problem of variance estimation and proposed various estimators. Srivastava and Jhajj (1995) utilized auxiliary information regarding the third moment of the auxiliary variable to create a category of estimators for both population variance and mean. Singh *et al.* (1995), Agarwal and Sthapit (1995), Singh *et al.* (1996), Jhajj *et al.* (2005), Gupta and Misra (2006), Kadilar and Cingi (2006a, 2006b, 2007), Shabbir and Gupta (2010), Misra and Yadav (2010), Misra *et al.* (2012) have also contributed to this field. Pal and Singh (2018) recommended enhanced estimators for finite population variance, demonstrating their superiority over conventional and other existing estimators. Singh *et al.* (2018) introduced an improved generalized class of estimators based on SRSWOR. Sanaullah *et al.* (2022) proposed an exponential type ratio-cum-product estimator for population variance. Muhammed *et al.* (2022) and Oluwapelumi *et al.* (2022) also proposed estimators for population variance estimation.

These researchers have made significant contributions to the advancement of diverse techniques and methodologies for variance estimation, addressing various survey scenarios and needs.

The present thesis "**A Certain Study on Estimation of Population Parameters in Survey Sampling**" is confined in the use of auxiliary variable at the estimation stage. Inspired by many authors we have proposed estimators in different situations. The sampling scheme is taken as to be the Simple Random Sampling without Replacement. Our interest lies in developing ratio type estimators for better estimation.

Chapter 2, introduces enhanced estimation of population mean $(\bar{Y})$ using known information on an auxiliary variable (X). The suggested estimator's bias and mean squared error (MSE) are calculated up to the first order of approximation. The minimum MSE and the Bias values are achieved by optimising the characterizing scalar. The MSE of the recommended estimator has also been compared both conceptually and empirically with the MSEs of current estimators. Real and Simulated data sets are adopted to verify the theoretical prerequisites for the suggested estimator's greater efficiency over competing estimators.

Chapter 3 aims to discuss the improvement of the estimation of the finite population mean of the study variable by employing an auxiliary variable in a simple random sampling scheme. We propose a new general ratio type estimator consisting of both exponential and logarithm functions. The sampling characteristics, bias and mean squared error (MSE) of the proposed estimator are evaluated up to first order approximation. The optimum values for the characterizing constants are observed. For the optimum values of the characterizing scalars, the minimum value of the MSE of the proposed estimator is also obtained. The proposed estimator is theoretically compared to the competing population mean estimators. The efficiency conditions for the proposed estimator to outperform the competing estimators are also obtained. Real and simulated data sets are used to validate the theoretical requirements for the proposed estimator's inherent superiority over rival estimators.

Chapter 4 introduces an enhanced family of exponential-log ratio type class of estimators for estimating the population mean $(\bar{Y})$ for the main variable (Y) utilizing certain known parameters of population for the auxiliary variable (X). The bias and mean squared error (MSE) of the proposed class of estimators are computed up to the first order of approximation. The lowest MSE and bias values are obtained by optimizing the characterizing scalar. In addition, the recommended estimator's conceptual and empirical MSEs are compared to the MSEs of the existing rival estimators. Real and simulated data sets are used to validate the theoretical requirements for the recommended estimator's superiority over rival estimators.

In chapter 5, the information on auxiliary variables has been shown to be beneficial in improving the precision of study variable estimates. An improved exponential ratio estimator is suggested and verified using auxiliary variable information for population variance under simple random sampling. The proposed estimator's bias and mean squared error (MSE) equations were derived up to first order of approximation. Using real data sets and simulation study, the performance assessment of the proposed estimator is considered and compared to the competing estimators. The empirical and simulation findings demonstrate that the suggested estimator outperforms the competing estimator.

REFERENCES

Aggarwal, S., Soni, S.S. and Chaubey, R.S. (2022). Estimation of the population mean by developing a new estimator. *Bulletin of Pure and Applied Sciences Section – E- Mathematics & Statistics,* 41E(1), 11-15.

Baghel, S. and Yadav, S.K. (2020). Restructured Class of Estimators for Population Mean Using an Auxiliary Variable under Simple Random Sampling Scheme. *Journal of Applied Mathematics, Statistics and Informatics,* 16(1), 61-75.

Bahl S. and Tuteja R.K. (1991). Ratio and product type exponential estimator, *Journal of Information and Optimization Sciences*. 12(1), 159-164.

Chakrabarty, R.P. (1979). Some ratio estimators. *Journal of the Indian Society of Agricultural Statistics*, 31(1), p. 49-57.

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

Cochran, W.G. (1942). Sampling theory when the sampling units are of unequal size. *Journal of the American Statistical Association*, 37(218), 199-212.

Des, R. (1972). *The Design of Sampling Surveys.* McGraw-Hill, New York.

Goodman, L. A. and Hartley, H.O. (1958). The precision of unbiased ratio-type estimators. *Journal of the American Statistical Association*, 53(282), p. 491-508.

Gupta, R.K. and Misra, S. (2006). Estimation of population variance using ratio type estimator. *Indian Journal of Mathematics and Mathematical Sciences*, 2(2), 169-176.

Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78, 117-123.

Jeelani, M. I., Maqbool, S. and Mir, S. A. (2013). Modified ratio estimators of population mean using linear combination of coefficient of skewness and quartile deviation. *International Journal of Modern Mathematical Sciences,* 6, 174-183.

Jerajuddin, M. and Kishun, J. (2016). Modified ratio estimators for population mean using size of the sample selected from population. *IJSRSET*, 2(2), P. 10-16.

Jhajj, H.S., Sharma, M.K. and Grover, L.K. (2005). An efficient class of chain estimators of population variance under sub sampling scheme. *Journal of the Japan Statistical Society*, 35, 273-286.

Kadilar, C. and Cingi, H. (2006a). An improvement in estimating the population mean by using the correlation co-efficient. *Hacettepe Journal of Mathematics and Statistics*, 35(1), 103-109.

Kadilar, C. and Cingi, H. (2006b). Improvement in estimating the population mean in simple random sampling. *Applied Mathematics Letters*, 19(1), 75-79.

Kadilar, C. and Cingi, H. (2007). Improvement in variance estimation in simple random sampling. *Communication in Statistics-Theory and Methods*, 36(11), 2075-2081.

Kadilar, C., Cingi, H. (2003). A study on the chain ratio type estimator. *Hacettepe Journal of Mathematics and Statistics*, Vol. 32, p. 105-108.

Liu, T.P. (1974). A general unbiased estimator for the variance of a finite population. *Sankhya,* 36, 23-32.

Mahalanobis, P.C. (1952). Some aspects of the design of sample survey. *Sankhya*, 12, 1-7.

Misra, S., and Yadav, S.K. (2010). Ratio type estimator of population variance using qualitative auxiliary information. *International Transactions in Mathematical Sciences and Computer*, 3(2), 313-322.

Muhammad, I., Zakari, Y., Audu, A. (2022). Generalized Estimators for Finite Population Variance Using Measurable and Affordable Auxiliary Character. *Asian Research Journal of Mathematics* 18(1): 14-30, Article no.ARJOM.72746

Murthy, M.N. (1964). Product Method of Estimation. *Sankhya*, 26(A), 69-74.

Oluwapelumi, A., Hasaan, A. A., Bolanle, O. O. and Tawakalitu O. (2022). An improved generalized estimators for finite population variance of a study variable based on auxiliary information. *International Journal of Statistics and Applied Mathematics* 7(4): 271-280.

Robson, D.S. (1957). Application of multivariate polykays to the theory of Unbiased Ratio type Estimation. *Journal of the American Statistical Association*, 52, 511-522.

Sahai, A. and Ray, S.K. (1980). An efficient estimator using auxiliary information. *Metrika*, 27, 271-275.

Searls, D.T. (1964). The utilization of known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association*, 59, 1225-1226.

Singh, R.K. and Pandey, S.K. (1983). A transformed generalized product type estimator. *Pure and Applied Mathematics Science*, 18, 67-73.

Singh, R.K., Zaidi, S.M.H. and Rizvi, S.A.H. (1995). Estimation of finite population variance using auxiliary information. *Journal of Statistical Studies*, 15,17-28.

Singh, H. P. and Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of an auxiliary variable. *Statistica, anno LXV*, 3, 301-313.

Singh, H.P., Tailor, R., Tailor, R., Kakran, M.S. (2004). An improved estimator of the population mean using power transformation. *Journal of the Indian Society of Agricultural Statistics*, 58(2), 223-230.

Tailor, R. and Sharma, B.K. (2009). A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis. *Statistics in Transition-New Series*, 10, 15-24.

Upadhyaya, L. N., and Singh, H. P. (1999a). An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathematical Journal,* 19, 14-17.

Upadhyaya, L.N. and Singh, H.P. (1999b). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal*, 41(5), 627-636.

Watson, D.J. (1937). The estimation of leaf areas, *The Journal of Agricultural Science*, 27, 474-483.

Yadav, S.K., Dixit, M.K, Dungana, H.N and Mishra, S.S. (2019). Improved Estimators for Estimating Average Yield Using Auxiliary Variable. *International Journal of Management Sciences,* 4(5), 1228-1238.

Yadav, S.K. and Baghel, S. (2021). A Redefined Class of Ratio Estimators using Auxiliary Information for the Estimation of Population Mean under Simple Random Sampling Scheme. *Revista Investigation Operational*, 42(2), 279-293.

Yan, Z., Tian B. (2010). Ratio Methods to the Mean Estimating using Coefficient of Skewness of Auxiliary Variable. *ICICA 2010, Part II, CCIS*, 106, 103-110.