A New Modified Exponential Type Ratio Estimator with Auxiliary Attribute and Simple Random Sampling

Monika Saini and Ashish Kumar

Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur (Rajasthan) - 303007

Abstract: In the present study, a new modified exponential type ratio estimator is proposed for the estimation of population mean using auxiliary attribute and analyze their properties. A comparative study, theoretically as well as numerically, of the proposed estimator has been made with some existing estimators.

Key words: Finite population · Ratio estimator · Auxiliary attribute · Efficiency

INTRODUCTION

The estimation of the population mean is an unrelenting issue in sampling theory and several efforts have been made to improve the precision of an estimator of unknown population parameter of interest when study variable y is highly correlated with the auxiliary variable x. There are many situations when auxiliary information is qualitative in nature that is auxiliary information is available in the form of an attribute, which is highly correlated with study variable e.g. - sex and height of the person, amount of milk produced and a particular breed of the cow, amount of yield of wheat crop and a particular variety of wheat etc. (see Jhajj et.al. [1]). In such situations, taking the advantage of point bi-serial correlation between the study variable y and the auxiliary attributes the estimators of population parameter of interest can be constructed by using prior knowledge of the population parameter of auxiliary attribute.

Now consider a finite population which consists of N identifiable units $U_i (1 \leq i \leq N)$. Assume that a sample of size n drawn by using simple random sampling without replacement (SRSWOR) from a population of size N. Let $y_i$ and $u_i$ denote the observations on the variable y and $u$ respectively for $i^{th}$ unit ($i=1,2, N$). Suppose there is a complete dichotomy in the population with respect to the presence or absence of an attribute, say $u$ and it is assumed that attribute $u$ takes only two values 0 and 1 according as $u_i$ If $i^{th}$ unit of the population possesses attribute $u = 0$ otherwise.

Let $\sum_{i=1}^{N} y_i/N$ and $\sum_{i=1}^{n} u_i/n$ denote the proportion of units in the population and sample respectively possessing attribute $u$. Let $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ and $\bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i$ be the sample means of variable of interest y and auxiliary attribute $u$ and $\bar{y}_p = \frac{1}{N} \sum_{i=1}^{N} y_i$ be the corresponding population means. We take the situation when the mean of the auxiliary attribute (P) is known. Let $s^2_y = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$ and $s^2_u = \frac{1}{n-1} \sum_{i=1}^{n} (u_i - \bar{u})^2$ be the sample variance and $s^2_{yp} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y}_p)^2$ be the corresponding population variance. Let $c_y = s_y/\bar{y}$ and $c_u = s_u/\bar{u}$. Finally let $\rho_{yu} = s_{yu}/s_y s_u$ be the point bi-serial correlation coefficient between y and u.

In order to determine the characteristic of the proposed estimators and existing estimators considered here, we define the following terms,
The aim of present paper is to propose modified exponential type ratio estimator for estimating the population mean of the variable under study, which make use of information regarding the population proportion of certain attribute. The expressions for variance have been obtained. A comparative study, theoretically, as well as numerically, of the proposed estimator has been made with some existing estimators.

**Some Existing Estimators:** In this section, we consider the several existing estimators which are used for the estimation of population mean.

**Traditional Sample Mean Estimator** ($\bar{y}_{TE}$): The variance of the sample mean $\bar{y}$, the usual unbiased estimator with no use of auxiliary attribute is given by

$$Var(\bar{y}_{TE}) = \Omega S^2_y$$  \hspace{1cm} (1)

**Bahl and Tuteja [2] Exponential Estimators** ($\bar{y}_{BTEE}$): For the estimation of population mean $\bar{y}$ of the study variable $y$, assuming the knowledge of the population proportion $P$, Bahl and Tuteja suggested the exponential type ratio estimator are as follows:

$$\bar{y}_{BTEE} = \bar{y} \exp \left( \frac{P - p}{P + p} \right)$$  \hspace{1cm} (2)

The mean square error (MSE) of $\bar{y}_{BTEE}$, is given by

$$MSE(\bar{y}_{BTEE}) = \Omega \bar{y}^2 \left[ C_y^2 + \frac{1}{4} C_y^2 - P_y C_y C_v \right]$$  \hspace{1cm} (3)

**Naik and Gupta [3] Ratio Estimator** ($\bar{y}_{NGRE}$): Naik and Gupta defined ratio estimator of population when the prior information of the population proportion of units, possessing the same attribute is available. Naik and Gupta proposed following estimator:

$$\bar{y}_{NGRE} = \frac{\bar{y}}{p}$$  \hspace{1cm} (4)

The mean square error (MSE) of $\bar{y}_{NGRE}$ up to the first order of approximation are given by

$$MSE(\bar{y}_{NGRE}) = \Omega \left[ S_y^2 + R^2 S_v^2 - 2 R P_y S_y S_v \right]$$  \hspace{1cm} (5)

**Proposed Estimators** ($\bar{y}_{METE}$) and its Properties: Under the same sampling design, we propose the modified exponential type ratio estimator for estimation of population mean as:

$$\bar{y}_{METE} = \bar{y} + a \left[ E(\Delta_y) - 1 \right]$$  \hspace{1cm} (6)

Whereas is any constant and $E(\Delta_y) = \exp \left[ \frac{p - n P}{N - n} \right]$

To obtain the unbiasedness and variance of $\bar{y}_{METE}$ up to the first order of approximation, we expand the equation (6) in terms of $\Delta_y$.

Expanding the right hand side of equation (6) up to the first order of approximation in terms of $\Delta_y$.

$$\bar{y}_{METE} = \bar{y} + a \left[ 1 + P - \frac{n P - n}{N - n} \right]$$  \hspace{1cm} (7)

Taking expectation on both sides of equation (7), it is clear that $E(\bar{y}_{METE}) = \bar{y}$

I.e. $\bar{y}_{METE}$ is an unbiased estimator of the population mean $\bar{y}$ up to the first order of approximation.

Now variance of equation (6) can be obtained as

$$Var(\bar{y}_{METE}) = \frac{\bar{y}^2}{p} \left[ E(\Delta_y) + a^2 P^2 \frac{n}{N - n} E(\Delta_y^2) - \frac{2 a n}{N - n} \bar{y} P \Delta_y E(\Delta_y^2) \right]$$
After simplification variance of \( \bar{y}_{METE} \) is

\[
\text{Var}(\bar{y}_{METE}) = \Omega \left[ S_y^2 + a^2 \left( \frac{n}{N-n} \right)^2 S_0^2 - \frac{2n}{N-n} a \rho_y \mu S_y S_0 \right]
\]

(8)

**Theorem 1:** Up to the first order of approximation, the estimator \( \bar{y}_{METE} \) is an unbiased estimator of population mean \( \bar{y} \) and its variance is

\[
\text{Var}(\bar{y}_{METE}) = \Omega \left[ S_y^2 + a^2 \left( \frac{n}{N-n} \right)^2 S_0^2 - \frac{2n}{N-n} a \rho_y \mu S_y S_0 \right]
\]

**Efficiency of Estimators:** Now we compare the proposed estimators \( \bar{y}_{METE} \) with existing estimators defined in section 2. We derive the following condition in which proposed estimator is better than the existing estimators:

**Proposed Estimator vs. Traditional Sample Mean Estimator \( \bar{y}_{TE} \)**

**Observation (A):** From equation (1) and (8)

\[
\text{Var}(\bar{y}_{TE}) - \text{Var}(\bar{y}_{METE}) = -a^2 \left( \frac{n}{N-n} \right)^2 S_0^2 + \frac{2n}{N-n} a \rho_y \mu S_y S_0 \]

if \( \rho_y > \frac{a S_y}{2 S_0} \left( \frac{n}{N-n} \right) \)

**Proposed Estimator vs. Bahl and Tuteja [2] exponential Estimators \( \bar{y}_{BTEE} \)**

**Observation (B):** From equation (3) and (8)

\[
\text{MSE}(\bar{y}_{BTEE}) - \text{Var}(\bar{y}_{METE}) = -a^2 \left( \frac{n}{N-n} \right)^2 S_0^2 + \frac{2n}{N-n} a \rho_y \mu S_y S_0 \]

if \( \rho_y > \frac{S_y}{2 S_0} \left( \frac{n}{N-n} \right) \)

**Proposed Estimator vs. Naik and Gupta [3] Ratio Estimator \( \bar{y}_{NGRE} \)**

**Observation (C):** From equation (5) and (8)

\[
\text{MSE}(\bar{y}_{NGRE}) - \text{Var}(\bar{y}_{METE}) = R^2 S_0^2 - 2 R \rho_y \mu S_y S_0 - a^2 \left( \frac{n}{N-n} \right)^2 S_0^2
\]

if \( \rho_y > \frac{S_y}{2 S_0} \left( \frac{n}{N-n} \right) \)

We arrive at the following theorem from observation (A), (B) and (C):

**Theorem 2:** The estimator \( \bar{y}_{METE} \) is more efficient than \( \bar{y}_{TE} \)

if \( \rho_y \mu > \frac{a S_y}{2 S_0} \left( \frac{n}{N-n} \right) \).

**Theorem 3:** The estimator \( \bar{y}_{METE} \) is more efficient than \( \bar{y}_{BTEE} \)

if \( \rho_y > \frac{S_y}{2 S_0} \left( \frac{n}{N-n} \right) \)

**Theorem 4:** The estimator \( \bar{y}_{METE} \) is more efficient than \( \bar{y}_{NGRE} \)

if \( \rho_y > \frac{S_y}{2 S_0} \left( \frac{n}{N-n} \right) \).

**Numerical Study:** In this section we compare the performance of various estimators considered here using the data sets as previously used by Shabbir and Gupta [4].
Population: (Source: Sukhatme and Sukhatme [5], pp. 256).

The variables of interest are: $y$ is the number of villages in the circle and $v$ represents a circle consisting more than five villages.

$$N = 89, \bar{y} = 3.36, P = 0.124, \rho_{y,v} = 0.766$$

$$n = 23 \quad C_y = 0.604 \quad C_v = 2.19$$

**CONCLUSION**

We have developed a new exponential type ratio estimator and obtained the characteristics for the proposed estimator. Theoretically, we have demonstrated that proposed estimator are always more efficient than the traditional sample mean estimator, Bahl and Tuteja [2] exponential type ratio estimator and ratio estimator proposed by Naik and Gupta [3]. In addition, we support this theoretical result numerically using the data used by Shabbir and Gupta [4].

**REFERENCES**