Applications and Applied Mathematics: An International Journal (AAM)

Volume 10 | Issue 1 | Article 7

6-2015

Exponential Type Product Estimator for Finite Population Mean with Information on Auxiliary Attribute

Monika Saini
Manipal University Jaipur

Ashish Kumar
Manipal University Jaipur

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Vol. 10, Issue 1 (June 2015), pp. 106 - 113

Applications and Applied Mathematics:
An International Journal
(AAM)

Exponential Type Product Estimator for Finite Population Mean with Information on Auxiliary Attribute

Monika Saini and Ashish Kumar
Department of Mathematics and Statistics
Manipal University Jaipur
Jaipur (Rajasthan) - 303007
drmnksaini4@gmail.com

Received: November 4, 2014; Accepted: March 10, 2015

Abstract

The main objective of the present study is to develop a new modified unbiased exponential type product estimator for the estimation of the population mean. The proposed estimator possesses the characteristic of a bi-serial negative correlation between the study variable and its auxiliary attribute. Efficiency comparison has been carried out between the proposed estimator and the existing estimators theoretically and numerically.

Keywords: Finite population; Simple random sampling**;** Product estimator, Auxiliary attribute; Efficiency

MSC 2010 No.: 62D05, 62D99

1. Introduction

In the existing literature of sampling theory, auxiliary information is generally used to improve the efficiency of estimators. Many estimators such as ratio, product, difference and regression estimators are good examples in this reference. When the study and auxiliary variables are positively correlated then the ratio estimator is used in this context. On the other hand, if both variables are negatively correlated then the product estimator is used to estimate the population parameter. The first attempt was made by Cochran (1940) to investigate the problem of estimation of the population mean when auxiliary variables are present and he proposed the usual ratio estimator of population mean. Robson (1957) and Murthy (1964) worked independently on the usual product estimator of the population mean. Recent developments in the ratio and product methods of estimation along with their variety of modified forms by Singh et al. (2010) proposed a ratio-cum-dual to ratio estimator for the estimation of the finite population mean of the study variable. Yadav (2011), Pandey et al. (2011), Shukla et al. (2012), Onyeka (2012) etc. have proposed many estimators utilizing auxiliary information.

There are many real life circumstances where auxiliary information is qualitative in nature, that is auxiliary information is available in the form of an attribute, which is highly correlated with a study variable, for example the sex and height of a 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. (2006)]. In such situations, taking the advantage of the point bi-serial correlation between the study variable y and the auxiliary attributes the estimators of the population parameter of interest can be constructed by using prior knowledge of the population parameter of auxiliary attribute.

The main objective of this study is to suggest a modified unbiased exponential type product estimator for estimating the population mean of the variable under study when the study variable and its auxiliary attribute are negatively correlated. The outlines of the present paper are as follows: in section 2, The sampling procedure and notations are given for constructing various estimators existing as well as proposed. The existing estimators and their properties are discussed in section 3. In section 4, the suggested estimator and its properties are developed. Efficiency comparison between proposed and existing estimators is carried out in section 5. Section 6 is devoted to a numerical study. In section 7, concluding remarks are given.

2. Sampling Procedure and Notations

Consider a finite population which consists of N identifiable units $\Delta_i \, (1 \le i \le N)$. Suppose that there is a complete dichotomy in the population with respect to the presence or absence of an the attribute, say Ψ, and it is assumed that the attribute Ψ takes only two values 0 and 1 according as Ψ_i assumes value 1 when the i^{th} unit of the population possesses attribute, otherwise assumes the value zero.

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 Ψ_i denote the observations on the variable y and Ψ respectively for i^{th} unit (i=1, 2, ..., N). According to the above sampling scheme, we define symbolically. Let

$$\bar{y} = \sum_{i}^{n} y_i / n \quad \text{and} \quad p = \sum_{i}^{n} \Psi_i / n$$

be the sample means of variable of interest y and auxiliary attribute Ψ and

$$\bar{Y} = \sum_{i}^{N} y_i / N \quad \text{and} \quad P = \sum_{i}^{N} \Psi_i / N$$

be the corresponding population means, where

$$P=\sum_{i=1}^{N}\Psi_i/N \text{ and } p=\sum_{i=1}^{n}\Psi_i/n$$

denote the proportion of units in the population and sample respectively possessing attribute Ψ. We take the situation when the mean of the auxiliary attribute (P) is known. Let

$$s_y^2=\sum_{i=1}^{n}(y_i-\bar{y})^2/n-1 \text{ and } s_\Psi^2=\sum_{i=1}^{n}(\Psi_i-p)^2/n-1$$

be the sample variance and

$$S_y^2=\sum_{i=1}^{N}(y_i-\bar{Y})^2/N-1 \text{ and } S_\Psi^2=\sum_{i=1}^{N}(\Psi_i-P)^2/N-1$$

be the corresponding population variance. Let

$$C_y=S_y/\bar{Y} \text{ and } C_\Psi=S_\Psi/P.$$

Finally let

$$\rho_{y\upsilon}=S_{y\Psi}/S_y\,S_\Psi$$

be the point bi-serial correlation coefficient between y and Ψ.

In order to determine the characteristic of the proposed estimators and existing estimators considered here, we define the following terms,

$$\Omega_y=(\bar{y}-\bar{Y})/\bar{Y} \text{ and } \Omega_\Psi=(p-P)/P$$

such that

$$E\,[\Omega_i]=0, \text{ for } (i=y,\Psi),$$

$$E(\Omega_y^2)=\Theta\,S_y^2/\overline{Y^2}\ ,\ E(\Omega_\Psi^2)=\Theta\,S_\Psi^2/P^2 \text{ and}$$

$$E(\Omega_y\,\Omega_\Psi)=\Theta\ \rho_{y\Psi}\,S_y/\bar{Y}\,S_\Psi/P\ ,$$

where

$$\Theta=(1/n)-(1/N).$$

3. Existing Product Estimator

In this section, existing product estimators are considered for the estimation of population mean.

3.1. Exponential Type Product Estimator [Bahl and Tuteja (1991)]

Below Bahl and Tuteja (1991) suggested the exponential type product estimator by using auxiliary attribute:

$$\bar{y}_{(BTP)} = \bar{y}\exp\left(\frac{p-P}{p+P}\right). \tag{3.1}$$

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

$$MSE(\bar{y}_{BTP}) = \Theta\ \bar{Y}^2\left[C_y^2 + \frac{1}{4}C_\Psi^2 + \rho_{y\Psi}C_y\ \ C_\Psi\right]. \tag{3.2}$$

3.2. Product Estimator [Naik and Gupta (1996)]

Naik and Gupta (1996) suggested the product estimator when the same attribute is available is as follows:

$$\bar{y}_{NGP} = \left(\frac{\bar{y}}{P}\right)p\,. \tag{3.3}$$

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

$$MSE(\bar{y}_{NGP}) = \Theta\ \left[S_y^2 + R^2S_\Psi^2 + 2R\rho_{y\Psi}S_y\,S_\Psi\right]. \tag{3.4}$$

4. Suggested Estimator and their Properties

Under the same sampling design, we propose the modified exponential type product estimator for the estimation of the population mean as:

$$\bar{y}_{MEPE} = [\,\bar{y} - k\ (t-1)\,], \tag{4.1}$$

where k is any constant and

$$t = \exp\left[\frac{NP-np}{N-n} - P\right].$$

To obtain the unbiasedness and Mean square error (MSE) of $\bar{y}_{MEPE}$, we expend the equation (4.1) in terms of $\Omega's$

$$(\bar{y}_{MEPE}) = \bar{Y}\left[1+\Omega_y\right] - k\left[\{1+\frac{NP-nP(1+\Omega_\Psi)}{N-n} - P + \ldots\} - 1\right].$$

Expanding the right hand side of equation (4.1) up to the first order of approximation in terms of $\Omega's$ we will have:

$$(\bar{y}_{MEPE}) = \bar{Y}\left[1+\Omega_y + \ldots\right] - k\left[\{1+\frac{NP-nP(1+\Omega_\Psi + \ldots)}{N-n} - P + \ldots\} - 1\right].$$

Rewriting $\bar{y}_{MEPE}$ up to the first order of approximation, we have

$$(\bar{y}_{MEPE}) = \bar{Y}\left[1+\Omega_y\right] + k\left[\frac{nP\Omega_\Psi}{N-n}\right]. \quad (4.2)$$

Taking expectation on both sides of equation (4.2), we can easily prove that $E(\bar{y}_{MEPE}) = \bar{Y}$, i.e., $\bar{y}_{MEPE}$ is an unbiased estimator of the population mean $\bar{Y}$.

Now variance of equation (4.2) can be obtained as

$$Var(\bar{y}_{MEPE}) = E\left[\overline{(y}_{MEPE}) - E((\bar{y}_{MEPE}))\right]^2,$$

$$Var(\bar{y}_{MEPE}) = [\bar{Y}\left[1+\Omega_y + \ldots\right] - k\left[\{1+\frac{NP-nP(1+\Omega_\Psi + \ldots)}{N-n} - P + \ldots\} - 1\right] - \bar{Y}].$$

Up to the first order of approximation, the variance of $\bar{y}_{MEPE}$ can be written as:

$$Var(\bar{y}_{MEPE}) = \left\{\bar{Y}\Omega_y - \left[\frac{NPk - nPk(1+\Omega_\Psi)}{N-n} - Pk\right]\right\}$$

$$= E\left[\bar{Y}\Omega_y + \frac{n}{N-n}kP\Omega_\Psi\right]^2 \qquad \text{(neglecting the higher order terms)}$$

$$= \bar{Y}^2 E(\Omega_y^2) + k^2P^2(\frac{n}{N-n})^2 E(\Omega_\Psi^2) + \frac{2n}{N-n}k\bar{Y}P\,E(\Omega_y\Omega_\Psi).$$

After simplification, variance of $\bar{y}_{MEPE}$ will be:

$$Var(\bar{y}_{MEPE}) = \Theta\left[S_y^2 + k^2(\frac{n}{N-n})^2 S_\Psi^2 + \frac{2n}{N-n}k\,\rho_{y\Psi}S_yS_\Psi\right]. \quad (4.3)$$

Theorem: 4.1.

The estimator $\bar{y}_{MEPE}$ is an unbiased estimator of population mean $\bar{Y}$, up to the first order of approximation and its variance

$$Var(\bar{y}_{MEPE}) = \Theta\left[S_y^2 + k^2\left(\frac{n}{N-n}\right)^2 S_\Psi^2 + \frac{2n}{N-n} k\,\rho_{y\Psi} S_y S_\Psi\right].$$

5. Efficiency of Comparison

Now we compare the suggested estimator defined in section (3) with existing estimators defined in section (2). We derive the following condition in which proposed estimators are better than the existing estimators:

(i) Proposed Estimator vs. Exponential Type Product Estimator [Bahl and Tuteja (1991)].

Condition (i): From equation (3.2) and (4.3)

$$MSE(\bar{y}_{BTP}) - Var(\bar{y}_{MEPE})$$

$$= [-\rho_{y\Psi} R S_y S_\Psi + \frac{1}{4} R^2 S_\Psi^2 - k^2\left(\frac{n}{N-n}\right)^2 S_\Psi^2 - \frac{2n}{N-n} k\rho_{y\Psi} S_y S_\Psi] > 0,$$

$$\text{if } \rho_{y\Psi} < -\frac{S_\Psi}{S_y}\left[\left(\frac{kn}{N-n}\right)^2 - \frac{R^2}{4} \Big/ \frac{2kn}{N-n} - R\right].$$

(ii) Proposed Estimator vs. Naik & Gupta (1996) Ratio Estimator.

Condition (ii): From equation (3.4) and (4.3)

$$MSE(\bar{y}_{NGP}) - Var(\bar{y}_{MEPE})$$

$$= \left[R^2 S_\Psi^2 - 2R\rho_{y\Psi} S_y S_\Psi - k^2\left(\frac{n}{N-n}\right)^2 S_\Psi^2 - \frac{2n}{N-n} k\rho_{y\Psi} S_y S_\Psi\right] > 0$$

$$\text{if } \rho_{y\Psi} < -\frac{S_\Psi}{2S_y}\left(\frac{kn}{N-n} + R\right).$$

From conditions (i) and (ii), we arrive at the following theorems:

Theorem 5.1.

The estimator $\bar{y}_{MEPE}$ is more efficient than $\bar{y}_{BTP}$, if

$$\rho_{y\Psi} < -\frac{S_\Psi}{S_y}\left[(\frac{kn}{N-n})^2 - \frac{R^2}{4} \Big/ \frac{2kn}{N-n} - R \right].$$

Theorem: 5.2.

The estimator $\bar{y}_{MEPE}$ is more efficient than $\bar{y}_{NGP}$, if

$$\rho_{y\Psi} < -\frac{S_\Psi}{2S_y}\left(\frac{kn}{N-n} + R \right).$$

6. Numerical Study

Now we compare the performance of the suggested and existing estimators considered here by using the data sets as previously used by Shabbir and Gupta (2010).

Population: (Source: Sukhatme and Sukhatme (1970), pp. 256).

y = Number of villages in the circles.

Ψ = A circle consisting more than five villages.

Table 1. Values of Parameters

$N=89$	$\bar{Y}=3.36$	$P=0.124$	$\rho_{y\Psi}=0.766$	$n=23$	$C_y=0.604$	$C_\Psi=2.19$

Table 2. Percent Relative Efficiency of Proposed Estimator vs. Existing Product Estimators

k	**Percent Relative Efficiency w.r.t** $\bar{y}_{BTP}$		**Percent Relative Efficiency w.r.t** $\bar{y}_{NGP}$	
	$\bar{y}_{BTP}$	$\bar{y}_{MEPE}$	$\bar{y}_{NGP}$	$\bar{y}_{MEPE}$
–4	100	499	100	946
–2	100	462	100	876
0	100	425	100	808
2	100	391	100	742
4	100	358	100	680

6. Conclusion

In the present study, a new modified unbiased exponential type product estimator and its characteristics are obtained. Theoretically, we obtain the conditions for which the proposed estimator is more efficient than the exponential type product estimator (Bahl and Tuteja (1991)) and product estimator (Naik and Gupta (1996)) always. We support the theoretical results by a numerical study with the help of information used by Shabbir and Gupta (2010). The results of numerical study reveals that proposed estimator is always better than existing estimators proposed by Bahl & Tuteja (1991) and Naik and Gupta (1996).

REFERENCES

Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimators, *Information and Optimization Sciences*, Vol. 12 (1), pp. 159–163.

Cochran, W.G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, *The journal of agricultural science*, Vol. 30, pp. 262-275.

Jhajj, H.S., Sharma, M.K. and Grover, L.K. (2006). A family of estimators of population means using information on auxiliary attributes, *Pakistan journal of Statistics*, Vol. 22(1), pp. 43-50.

Murthy, M.N. (1964): Product method of estimation, *Sankhya A*, Vol.26, pp. 69-74.

Naik, V.D. and Gupta, P.C. (1996). A note on estimation of mean with known population proportion of an auxiliary character, *Jour. Ind. Soc. Agri. Stat.*, Vol. 48(2), pp. 151-158.

Onyeka, A. C. (2012). Estimation of population mean in post stratified sampling using known value of some population parameter(s), *Statistics In Transition-new series*, Vol. 13, No. 1, pp. 65-78.

Pandey, H., Yadav, S.K. and Shukla, A.K. (2011). An Improved General Class of Estimators Estimating Population Mean using Auxiliary Information, *International Journal of Statistics and Systems*, Vol. 6(1), pp. 1–7.

Robson, D.S. (1957). Application of multivariate polykays to the theory of unbiased ratio type estimators, *Jour. Amer. Stat. Assoc*., Vol. 52, pp. 511-522.

Shukla, D., Pathak, S. and Thakur, N.S. (2012). Estimation of population mean using two auxiliary sources in sample survey, Statistics In Transition-new series, March 2012, Vol. 13, No. 1, pp. 21-36.

Sukhatme, P.V., and Sukhatme, B.V. (1970). Sampling Theory of Surveys with Applications, *Lowa State University Press, Ames*, USA.

Shabbir, J. and Gupta, S. (2010). Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes, *Hecettepe Journal of Mathematics and Statistics*, Vol. 39(1), pp. 121-129.

Singh, H.P., Tailor, R., and Tailor, R. (2010). On ratio and product methods with certain known population parameters of auxiliary variable in sample surveys, *Statistics and operations research trans.*, Vol. 34, pp. 157-180.

Yadav, S.K. (2011), Efficient Estimators for Population Variance using Auxiliary Information, *Global Journal of Mathematical Sciences: Theory and Practical*. Vol. 3(4), pp. 369-376.