Oriental Journal of Physical Sciences Vol. 1(1 & 2),21-25 (2016)

On the Estimation of Population Mean Under Systematic Sampling using Auxiliary Attributes

USMAN SHAHZAD

Department of Mathematics and Statistics, PMAS Arid Agriculture University, Rawalpindi. Pakistan.
* Corresponding author E-mail: usman.stat@yahoo.com

(Received: 22 Aug 2016; Accepted: 2 Sept 2016)

ABSTRACT

Naik and Gupta (1996), Singh *et al.,* (2007) and Abd-Elfattah *et al.,* (2010) introduced some estimators for estimating population mean $\bar{Y}$ using available auxiliary attributes under simple random sampling scheme. We adapt these estimators under systematic random sampling scheme using available auxiliary attributes. Further, a new family of estimators is proposed for the estimation of population mean $\bar{Y}$ under systematic random sampling scheme. The properties such as bias and mean square error of the proposed estimators are derived. From numerical illustration it is shown that proposed estimators are more efficient than the reviewed ones.

Key-words: Mean Square Error, Attributes, Study Variable, Systematic Random Sampling.

INTRODUCTION

Systematic random sampling is the simplest type of sampling scheme, requires only one random start. It provides good results in some situations like; forest regions for assessing the volume of the timber etc. For details see Murthy (1967) and Cochran (1977). When supplementary information is available Swain (1964), Shukla (1971), Singh and Solanki (2012), and Singh *et al.,* (2012) have developed some estimators for $\bar{Y}$ using available supplementary information. But none of these have paid their attention towards auxiliary attributes. So in our work, we utilize available auxiliary attributes.

In the theory of survey sampling, supplementary information plays a vital role for increasing the efficiency of population parameters. A number of authors have developed estimators based on auxiliary information. Another way to enhance the efficiency of an estimator is to utilize auxiliary attributes. Naik and Gupta (1996), Singh *et al.,* (2007), Abd-Elfattah *et al.,* (2010), Solanki and Singh (2012) and Koyuncu (2012) introduced various estimators utilizing available auxiliary attributes under simple random sampling. Taking motivation from these we are going to propose a family of estimators under systematic random sampling scheme using available auxiliary attributes.

Preliminaries and Adapted Estimators

Let ℓ be the finite population having units 1 to N. Further, we consider N=nk, where n and k are positive whole numbers. Hence there will be k samples of size n. Let M be random variable having range 1 to k. The systematic random sample is then selected by the following random sequence as

$$\{I_{M,}\ I_{M+k,}\ldots,I_{M+(n-1)k,}\ \}.$$

Let y_{ij} and ϕ_{ij} denote the values of the study variable and auxiliary attribute for (i=1,2,…,k) and (j=1, 2,…,n).

Note that ϕ_{ij} is the binary character so it can take only two possible values i.e.,

ϕ_{ij}=1, if the ith unit of the population possesses attribute ϕ,

ϕ_{ij}=0, otherwise.

Let A= $\sum_{j=1}^{N} \phi_{ij}$ and a= $\sum_{j=1}^{N} \phi_{ij}$ denote the total number of units in the population and sample respectively, possessing an auxiliary attribute ϕ. Hence, the corresponding sample and population proportions are

P=A/N and p_{lss}=a/n .

Similarly, $\bar{Y} = \frac{\sum_{j=1}^{N} y_{ij}}{N}$ and $y_{lss} = \frac{\sum_{j=1}^{N} y_{ij}}{N}$ are the sample and population means of Y. For finding MSE, Let we define $e_o = \frac{\bar{y}_{lss}-\bar{Y}}{\bar{Y}}$ and $e_1 = \frac{p_{lss}-P}{P}$, using these notations, we have

$E(e_o)=E(e_1)=0$,
$E(e_o^2)=(N-1)/Nn)\, C_y^2\, [1+(n-1)\, \rho_y]=y_0$,
$E(e_1^2)=(N-1)/Nn)\, C_P^2\, [1+(n-1)\, \rho_P]=y_1$,
$E(e_o\, e_1)=(N-1)/Nn)\, \rho C_y C_P\, [1+(n-1)\, \rho_y]^{1/2} = [1+(n-1)\, \rho_P]^{1/2} y_{01}$,

where ρ_P is the intra-class correlation of P, ρ_y is the intra-class correlation of Y and ρ is the correlation between P and Y.

The variance of the traditional sample mean is

$V(\bar{Y}_{lss})=\bar{Y}^2 y_0$.

Following Naik and Gupta (1996), we propose the usual ratio and product estimators utilizing available auxiliary attributes under systematic random sampling scheme,

$$\hat{t}_1 = \bar{y}_{lss}\left[\frac{P}{p_{lss}}\right],$$

$$\hat{t}_2 = \bar{y}_{lss}\left[\frac{p_{lss}}{P}\right].$$

The MSEs of $\hat{t}_1$ and $\hat{t}_2$ are

$$MSE(\hat{t}_1) = \bar{Y}^2[\gamma_0 + \gamma_1 - 2\gamma_{01}],$$

$$MSE(\hat{t}_2) = \bar{Y}^2[\gamma_0 + \gamma_1 + 2\gamma_{01}].$$

Motivated by Singh *et al.,* (2007), we propose the exponential ratio and product estimators utilizing available auxiliary attributes under systematic random sampling scheme,

$$\hat{t}_3 = \bar{y}_{lss} exp\left[\frac{P-p_{lss}}{P+p_{lss}}\right],$$

$$\hat{t}_4 = \bar{y}_{lss} exp\left[\frac{p_{lss}-P}{P+p_{lss}}\right].$$

The MSEs of $\hat{t}_3$ and $\hat{t}_4$ are

$$MSE(\hat{t}_4) = \bar{Y}^2\left[\gamma_0 + \frac{1}{4}\gamma_1 + 2\gamma_{01}\right].$$

$$MSE(\hat{t}_4) = \bar{Y}^2\left[\gamma_0 + \frac{1}{4}\gamma_1 + 2\gamma_{01}\right].$$

On the lines of Abd-Elfattah *et al.,* (2010), we propose the family of estimators utilizing available auxiliary attributes under systematic random sampling scheme,

$$\hat{t}_5 = m_1 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}} P + m_2 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss} + \beta_2(\phi)}(P + \beta_2(\phi)) ,$$

$$\hat{t}_6 = m_1 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}} P + m_2 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss} + C_p}(P + C_p) ,$$

$$\hat{t}_7 = m_1 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}} P + m_2 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}\beta_2(\phi) + C_p}(P\beta_2(\phi) + C_p) ,$$

$$\hat{t}_8 = m_1 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}} P + m_2 \frac{[\bar{y}_{lss} + b_\phi\,(P - p_{lss})]}{p_{lss}C_p + \beta_2(\phi)}\left(PC_p + \beta_2(\phi)\right) .$$

The minimum MSE of these estimators $(\hat{t}_5\ to\ \hat{t}_8)$ is equal to the MSE of regression estimator i.e.

MSE $(\hat{t}_{reg})$=$\gamma_o(1-\rho^2)\bar{Y}^2$,

Where $\hat{t}_{reg} = \bar{y}_{lss} + b_\phi\,(P - p_{lss})$.

Solanki and Singh (2013) developed the generalized estimator, given below

$$\hat{t}_\alpha = \bar{y}_{lss} exp\left[\frac{\alpha(P-p_{lss})}{P+p_{lss}}\right],$$

Where by putting (α=0, 1, -1) we get $\hat{t}_o, \hat{t}_3$ and $\hat{t}_4$ respectively.

The Proposed Family of Estimators

Taking motivation from Abd-Elfattah *et al.,* (2010), we propose the following family of estimators utilizing available auxiliary attributes under systematic random sampling scheme,

$$\hat{t}_N = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{ap_{lss} + b}{aP + b}\right] ,$$

where a and b be any known population characteristics or 1.

Let we express $\hat{t}_N$ in terms of e_o and e_1 as follows

$$\hat{t}_N = \left[\mathrm{m}_{1p}\bar{Y}(1+e_o) + \mathrm{m}_{2p}(1+\theta e_1)\right],$$

where θ =aP/aP+b.

Now by simplifying, we get

$$\hat{t}_N - \bar{Y} = \left[\mathrm{m}_{1p}\bar{Y}(1+e_o) + \mathrm{m}_{2p}(1+\theta e_1)\right] - \bar{Y}.$$

Let we take expectation on both sides and get the bias of $\hat{t}_N$ as

$$B(\hat{t}_N) = \left[\mathrm{m}_{1p}\bar{Y} + \mathrm{m}_{2p}\right] - \bar{Y}.$$

Now squaring both sides of $(\hat{t}_N - \bar{Y})$ as

$$(\hat{t}_N - \bar{Y})^2 = \bar{Y}^2 + \mathrm{m}_{1p}^2\bar{Y}^2\{1+\mathrm{e}_0^2\} + \mathrm{m}_{2p}^2\{1+\theta^2\mathrm{e}_1^2\} + 2m_{1p}m_{2p}\bar{Y}\{1+\theta e_o e_1\} - 2m_{1p}\bar{Y}^2 - 2m_{2p}\bar{Y}.$$

The MSE of $\hat{t}_N$ is given by

$$\text{MSE }(\hat{t}_N) = \bar{Y}^2 + \mathrm{m}_{1p}^2\mathrm{A^P} + \mathrm{m}_{2p}^2\mathrm{B^P} + 2\mathrm{m}_{1p}\mathrm{m}_{2p}\mathrm{C^P} - 2\mathrm{m}_{1p}\mathrm{E^P} - \mathrm{m}_{2p}\mathrm{F^P}.$$

Where

$A^p=\bar{Y}^2\{1+\bar{Y}_o\}$,

$B^p=\{1+\theta^2 y_1\}$,

$C^p=\bar{Y}\{1+\theta y_{o1}\}$,

$D^p=\bar{Y}^2$,

$E^p=2\bar{Y}$.

Partially differentiating MSE $\hat{t}_N$ w.r.t m_1p and m_2p and equating to zero, we have the following equations

$m_{1p}A^p + m_{2p}C^p = D^p$,

$m_{1p}C^p + m_{2p}B^p = E^p/2$,

Now by solving matrix inversion method we get the optimum values of m_{1p} and m_{2p} i.e

$$m_{1p}^{opt} = \frac{\mathrm{B^pD^p} - \frac{\mathrm{C^pE^p}}{2}}{\mathrm{A^pB^p} - \mathrm{C^p}^2},$$

and

$$m_{2p}^{opt} = \frac{\frac{\mathrm{A^pE^p}}{2} - \mathrm{C^pD^p}}{\mathrm{A^pB^p} - \mathrm{C^p}^2}.$$

By putting m_{1p}^{opt} and m_{2p}^{opt} in MSE $\hat{t}_N$ and get minimum mean square error of $\hat{t}_N$ i.e

$$\mathrm{MSE_{min}}(\hat{t}_N) = \left[\bar{Y}^2 - \frac{\mathrm{B^pD^p}^2 - \mathrm{C^pD^pE^p} + \frac{\mathrm{A^pE^p}^2}{4}}{\mathrm{A^pB^p} - \mathrm{C^p}^2}\right].$$

Table 1: Some members of proposed class

$\hat{t}_N = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{ap_{lss}+b}{aP+b}\right]$	***a***	***b***
$\hat{t}_{N1} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{\rho p_{lss}+\beta_1(\phi)}{\rho P+\beta_1(\phi)}\right]$	p	$\beta_1(\phi)$
$\hat{t}_{N2} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{\rho p_{lss}+C_P}{\rho P+C_P}\right]$	p	C_P
$\hat{t}_{N3} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{\rho p_{lss}+\beta_2(\phi)}{\rho P+\beta_2(\phi)}\right]$	p	$\beta_2(\phi)$
$\hat{t}_{N4} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{p_{lss}+\beta_2(\phi)}{P+\beta_2(\phi)}\right]$	1	$\beta_2(\phi)$
$\hat{t}_{N5} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{p_{lss}+C_P}{P+C_P}\right]$	1	C_P
$\hat{t}_{N6} = m_{1p}\bar{y}_{lss} + m_{2p}\left[\frac{C_P p_{lss}+\beta_2(\phi)}{C_P P+\beta_2(\phi)}\right]$	C_P	$\beta_2(\phi)$

Efficiency Comparison

In current section, we find the efficiency conditions for the proposed estimators $\hat{t}_N$ by looking at the mean square error of the existing estimators as given below

MSE $(\hat{t}_N)$ < MSE $(\hat{t}_1)$ If

$$\bar{Y}^2[1-y_o-y_1+2y_{o1}]-\left[\frac{B^pD^{p^2}-C^pD^pE^p+\frac{A^pE^{p^2}}{4}}{A^pB^p-C^{p^2}}\right]<0$$

MSE $(\hat{t}_N)$ < MSE $(\hat{t}_2)$ If

$$\bar{Y}^2[1-y_o-y_1+2y_{o1}]-\left[\frac{B^pD^{p^2}-C^pD^pE^p+\frac{A^pE^{p^2}}{4}}{A^pB^p-C^{p^2}}\right]<0$$

MSE $(\hat{t}_N)$ < MSE $(\hat{t}_3)$ If

$$\bar{Y}^2[1-y_o-1/4\,y_1+2y_{o1}]-\left[\frac{B^pD^{p^2}-C^pD^pE^p+\frac{A^pE^{p^2}}{4}}{A^pB^p-C^{p^2}}\right]<0$$

MSE $(\hat{t}_N)$ < MSE $(\hat{t}_4)$ If

$$\bar{Y}^2[1-y_o-1/4\,y_1+2y_{o1}]-\left[\frac{B^pD^{p^2}-C^pD^pE^p+\frac{A^pE^{p^2}}{4}}{A^pB^p-C^{p^2}}\right]<0$$

MSE $(\hat{t}_N)$ < MSE $(\hat{t}_{reg})$ If

$$\bar{Y}^2[1-y_o(1-P^2]-\left[\frac{B^pD^{p^2}-C^pD^pE^p+\frac{A^pE^{p^2}}{4}}{A^pB^p-C^{p^2}}\right]<0$$

From the above mentioned conditions, we can say that proposed estimators are more efficient as compare to adapted estimators.

Numerical Illustration

The performance of proposed and existing estimators examined through two real data sets.

Population 1

Data is taken from Murthy (1967) where Y=Volume of the timber and ϕ =length ≥6. Descriptives of the population are N=176, $\bar{Y}$=282.61, P=0.6022, S_y=155.73, S_p =0.6421, P=0.51, p_p= -0.0016, p_y=)–0.0019, C_p=1.0661, β_2 (ϕ)=29.1110, β_1 (ϕ)=10.1458, n=16 .

Population 2

Data is taken from Murthy (1967) where Y=Volume of the timber and ϕ= Volume >200. Descriptives of the population are N=176, $\bar{Y}$=282.61, P=0.6363, S_y=155.73, S_p=0.4824, $_p$=0.7831, P_p= -0.0023, P_y=–0.0019, C_p=0.7580, β_1 (ϕ)=1.3214, β_2)=0.3214, n=16 .

Table 2: MSEs of Adapted and Proposed Estimators

Est.	Pop 1	Est.	Pop 1	Est.	Pop 2	Est.	Pop 2
$\hat{t}_o$	1463.01	$\hat{t}_{N1}$	3.72	$\hat{t}_o$	1463.01	$\hat{t}_{N1}$	997.27
$\hat{t}_1$	4073.97	$\hat{t}_{N2}$	272.4	$\hat{t}_1$	1071.61	$\hat{t}_{N2}$	375.17
$\hat{t}_2$	9864.57	$\hat{t}_{N3}$	0.45	$\hat{t}_2$	7355.68	$\hat{t}_{N3}$	144.02
$\hat{t}_3$	1391.92	$\hat{t}_{N4}$	1.74	$\hat{t}_3$	579.65	$\hat{t}_{N4}$	223.86
$\hat{t}_4$	4287.22	$\hat{t}_{N5}$	678.21	$\hat{t}_4$	3721.68	$\hat{t}_{N5}$	534.75
$\hat{t}_{reg}$	1082.4	$\hat{t}_{N6}$	1.97	$\hat{t}_{reg}$	565.72	$\hat{t}_{N6}$	135.59

CONCLUSION

We have proposed a class of estimators for $\bar{Y}$ using available auxiliary attributes under systematic random sampling scheme and obtained its bias and minimum MSE equations. All the adapted estimators are compared with proposed estimators using MSE. With the help of these comparisons, efficiency condition has been found where proposed estimators perform much better. The theoretical conditions and numerical illustrations show that proposed estimators are much better. Hence, it is advisable to use the proposed class of estimators.

REFERENCES

1. 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. Agr. Stat*,48(2): 151-158.
2. Singh, R., Chouhan, P., Sawan, N., Smarandache, F. (2007): Ratio-product Type Exponential Estimator for Estimating Finite Population Mean Using Information on

Auxiliary Attribute. Renaissance High Press, USA. pp. 18–32.

3. Abd-Elfattah, A. M., El-Sherpieny, E. A., Mohamed, S. M., & Abdou, O. F. (2010): Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute. *Applied mathematics and computation*, 215(12): 4198-4202.
4. Cochran, W. G. (1977): Sampling techniques, 34rd ed. New York, NY, John Wiley and Sons. Shukla, N. D. (1971): Systematic sampling and product method of estimation. In Proceeding of all India Seminar on Demography and Statistics. BHU, Varanasi, India.
5. Swain, A.K.P.C.(1964): The use of systematic sampling in ratio estimate. J*our. Ind. Stat. Assoc*, 2(213), 160-164.
6. Singh, H. P., and Solanki, R. S. (2012): An efficient class of estimators for the population mean using auxiliary information in systematic sampling. *Journal of Statistical Theory and Practice*, 6(2): 274-285.
7. Singh, R., Malik, S., Chaudhary, M. K., Verma, H. K., & Adewara, A. A. (2012): A general family of ratio-type estimators in systematic sampling.*Journal of Reliability and Statistical Studies*, 5(1): 73-82.
8. Koyuncu, N. (2012): Efficient estimators of population mean using auxiliary attributes. *Applied Mathematics and Computation*, 218(22): 10900-10905.
9. Murthy, M.N. (1967): Murthy, M. N.: Sampling theory and methods. Sampling theory and methods.
10. Solanki, R. S., and Singh, H. P. (2013): Improved estimation of population mean using population proportion of an auxiliary character. *Chilean journal of Statistics*, 4(1): 3-17.