Available online at http://jurnal.ahmar.id/index.php/asci

Journal of Applied Science, Engineering, Technology, and Education

ISSN 2685-0591 (Online)

Journal of Applied Science, Engineering, Technology, and Education Vol. 1 No. 2 (2019)
https://doi.org/10.35877/454RI.asci1245

An Improved Mathematical Model Applying Practicable Algorithms

Tanveer Ahmad Tarray[a*], Zahoor Ahmad Ganie[b], Baziga Youssuf[c]

[a]*Department of Mathematical Sciences, Islamic University of Science and Technology, India, tanveerstat@gmail.com*
[b]*Department of Electrical Engineering, Islamic University of Science and Technology, India, zahoor.rifi@gmail.com*
[c]*Department of Electrical Engineering, Islamic University of Science and Technology, India, baziga07@gmail.com*

Abstract

In this article, we have considered the problem of estimation of population variance on two occasion successive sampling. A class of estimators of population variance has been proposed and its asymptotic properties have been discussed. The proposed class of estimators is compared with the sample variance estimator when there is no matching from the previous occasion. Numerical illustrations are also given in support of the present study.

Keywords: Successive (rotation) sampling, Variance estimation, bias, mean square error, optimum replacement policy.

1. Introduction

Jessen introduced the idea of using information achieved on the first occasion in improving the estimates of the current occasion [1]. Later, Yates extended Jessen's schemes to the situations where the population mean of a variable is estimated on each one of the h ($>$2) occasions for a rotation sample design [2]. These results were generalized by Patterson [3], Tikkiwal [4]; Eckler [5], and Rao and Grahm [6]. Generally, in successive sampling, our aim is to estimate the current average; the theory developed so for on successive sampling aims at providing the optimum estimate by combining.

(a) a double sampling regression estimate from the matched portion, where the "large" sample is the first sample and the auxiliary variable x is the value of y (study variable) on the first occasion and
(b) a sample mean based on a random sample from the unmatched portion. It is to be mentioned that a large number of estimators for estimating the current average have been studied by various authors.

A large number of estimators that estimate the population mean on the current occasion have been proposed by various authors, however, only limited efforts have been made to estimate the population variance on the current occasion in two occasions successive (rotation) sampling.

[*] Corresponding author.
E-mail address: tanveerstat@gmail.com (Tanveer Ahmad Tarray)

2. Notations

Let U=($U_1, U_2 \ldots U_N$) be the finite population of N (large) units, and which is assumed to remain unchanged over two occasions. Let x (y) be the character under study on the first (second) occasion respectively. A simple random sample (without replacement) of size n units is drawn on the first occasion and a random subsample of size m = nλ units from the sample on the first occasion is retained (matched) for its use on the current (second) occasion, A fresh (unmatched) sample of size u=(n-m)=nθ units is drawn on the second (current) occasion from the remaining (N-n) units of the population by a simple random sampling without replacement method so that the sample size on the current occasion is also n, λ and θ (λ+ θ=1) are the fractions of the matched and fresh samples, respectively, on the current occasion. Now we consider the following notations for their further use.

$S_x^2 = (N-1)^{-1} \sum_{i=1}^{N}(x_i - \overline{X})$: the population mean square (variance) of variable x.

S_y^2 : The population mean square (variance) of the study variable y.

$S_{x(n)}^2$: The sample mean square (variance) of variable x and based on sample of size n on the first occasion.

$S_{x(m)}^2, S_{y(m)}^2$: The sample mean square (variance) of variable x and y of matched sample of size m on the first and the current (second) occasion respectively.

$S_{y(u)}^2$: The sample mean square (variance) of variable y based on the unmatched (fresh) sample of size m on the first and current (second) occasions.

$f = \left(\frac{n}{N}\right)$: The sampling fraction.

3. The suggested class of estimators

To estimate the population variance S_y^2 (population mean square) on the current (second) occasion, the independent estimators are suggested. One is based on sample size u= (nθ) drawn fresh on the current (second) occasion, and which is given by

$$t_{1u} = S_{y(u)}^2 \tag{1}$$

second estimator is a one parameter class of estimators based on the sample of size m=(n λ) common with both occasion and is defined as

$$t_{2m} = S_{y(m)}^2 \exp\left[\frac{S_{x(n)}^2 - S_{x(m)}^2}{S_{x(n)}^2 + (\alpha - 1)S_{x(m)}^2}\right], \tag{2}$$

where α is a positive real constant. Here we note that:

(i) For α = 0, t_{2m} in (2) reduces

$$t_{2m}^{(1)} = S_{y(m)}^2 \exp(1) \tag{3}$$

which is a biased estimator with the mean square error larger than the usual unbiased estimator $S_{y(m)}^2$ utilizing no auxiliary information as the value of ‘e’ is always greater than ‘unity’.

(ii) For α = 1, in (2.3.2), t_{2m} reduces to

$$t_{2m}^{(2)} = S_{y(m)}^2 \exp\left(\frac{S_{x(n)}^2 - S_{x(m)}^2}{S_{x(n)}^2}\right). \tag{4}$$

(iii) For α = 2, t_{2m} in (2) boils down to the estimator

$$t_{2m}^{(3)} = S_{y(m)}^2 \exp\left(\frac{S_{x(n)}^2 - S_{x(m)}^2}{S_{x(n)}^2 + S_{x(m)}^2}\right). \tag{5}$$

Now considering the convex linear combination of the estimates t_{1u} and t_{2m} we have a class of estimators of S_y^2 as

$$t_d = \phi t_{1u} + (1-\phi)t_{2m}, \tag{6}$$

where φ is an unknown constant to be determined such that MSE of t_d is minimum.

4. Bias and Mean Square Error of the suggested estimator t_d

The suggested class of estimators t_d defined in equation (6) is also a biased estimator of S_y^2, So , its bias B(.) and mean square error MSE(.) up to the first order of approximation is derived under large sample approximation with the following transformations .

$$S_{y(m)}^2 = S_y^2(1+e_{0(m)}),\ S_{y(u)}^2 = S_y^2(1+e_{0(u)}),$$
$$S_{x(m)}^2 = S_x^2(1+e_{1(m)}),\ S_{x(n)}^2 = S_x^2(1+e_{1(n)}),$$

such that

$$E\left(e_{0(m)}\right) = E(e_{0(u)}) = E(e_{1(m)}) = E(e_{1(n)}) = 0$$

and to the first degree of approximation,

$$E(e_{0(m)}^2) = (1/m)(\lambda_{40} - 1),$$
$$E(e_{0(m)}e_{1(n)}) = (1/n)(\lambda_{22} - 1)$$

where $\lambda_{sr} = \dfrac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}$, (r,s) being positive integer, and $\mu_{rs} = \dfrac{1}{N}\sum_{z=1}^{N}(y_i - \overline{Y})^r(x_i - \overline{X})^s$

Under the above transformations, the estimator t_{2m} takes the following form:

$$t_{2m} = S_y^2(1+e_{0(m)})\exp\left[\frac{(1+e_{1(n)})S_x^2 - (1+e_{1(m)})S_x^2}{S_x^2(1+e_{1(n)}) + (\alpha-1)(1+e_{1(m)})S_x^2}\right]$$

$$\left(t_{2m} - S_y^2\right) \cong S_y^2\left[e_{0(m)} - \frac{(e_{1(m)} - e_{1(n)})}{\alpha} - \frac{(e_{0(m)}e_{1(m)} - e_{0(m)}e_{1(n)})}{\alpha} + \frac{\left\{(2\alpha-1)e_{1(m)}^2 - e_{1(n)}^2 - 2(\alpha-1)e_{1(m)}e_{1(n)}\right\}}{2\alpha^2}\right] \tag{7}$$

Taking expectation of both sides of (7) we get the bias of t_{2m} to the first degree approximation as

$$B(t_{2m}) = S_y^2\left[-\frac{1}{\alpha}\left\{\frac{1}{m}(\lambda_{22}-1) - \frac{1}{n}(\lambda_{22}-1)\right\} + \frac{1}{2\alpha^2}\left\{\frac{(2\alpha-1)}{m}(\lambda_{04}-1) - \frac{1}{n}(\lambda_{04}-1) - 2\frac{(\alpha-1)}{n}(\lambda_{04}-1)\right\}\right]$$

$$=\frac{S_y^2}{2\alpha^2}\left(\frac{1}{m}-\frac{1}{n}\right)(\lambda_{04}-1)[2\alpha(1-c)-1], \qquad (8)$$

where $C=\frac{(\lambda_{22}-1)}{(\lambda_{04}-1)}$.

Squaring both sides of (7) and neglecting terms of e's having power greater than two we have

$$\left(t_{2m}-S_y^2\right)^2=S_y^4\left[e_{0(m)}^2+\frac{(e_{1(m)}-e_{1(n)})^2}{\alpha^2}-\frac{2(e_{0(m)}e_{1(m)}-e_{0(m)}e_{1(m)})}{\alpha}\right]. \qquad (9)$$

Taking expectation of both sides of (8) we get the mean square error (MSE) of t_{2m} to the first degree of approximation as

$$MSE(t_{2m})=\left(\frac{s_y^4}{m}\right)\left[(\lambda_{40}-1)+\left(1-\frac{n}{m}\right)\left\{\frac{1}{\alpha^2}(\lambda_{04}-1)-\frac{2(\lambda_{22}-1)}{\alpha}\right\}\right]$$

$$=\left(\frac{S_y^4}{m}\right)\left[(\lambda_{40}-1)+\theta\frac{(\lambda_{04}-1)}{\alpha^2}(1-2\alpha C)\right],$$

where $\theta=\left(1-\frac{m}{n}\right)=(1-\lambda)\text{and}\ \lambda=\frac{m}{n}$.

Differentiating MSE(t_{2m}) with respect to α and equating to zero, we have

$$\frac{\partial MSE(t_{2m})}{\partial\alpha}=\left(\frac{S_y^4}{m}\right)\left[0+\theta(\lambda_{04}-1)\left\{-\frac{2}{\alpha^3}+\frac{2C}{\alpha^2}\right\}\right]=0 \qquad (10)$$

$$\Rightarrow \qquad -\frac{2}{\alpha^3}+\frac{2C}{\alpha^2}=0$$

$$\Rightarrow \qquad -1+\alpha C=0$$

$$\Rightarrow \qquad \alpha=\frac{1}{C}=\alpha_0(\text{say}) \qquad (11)$$

which is optimum value of is α that minimizes the MSE of t_{2m}.

$$MSE(t_{2m})_{opt}=\left(\frac{S_y^4}{m}\right)\left[(\lambda_{04}-1)-\theta(\lambda_{04}-1)C^2\right] \qquad (12)$$

Now we shale the following theorems.

5. Empirical Study

y: Number of households in 1971,
x: Number of households in 1961,

For this population we obtained:

$$\lambda_{04}=16.8930,\ \lambda_{40}=14.6872,\ \ \lambda_{22}=6.6714,\ C=0.356823$$

Population II: [Source: Cochran (1977), p. 325]

y: Number of persons per block,
x: Number of rooms per block,

For this population we obtained:

$$\lambda_{04} = 2.2387,\ \lambda_{40} = 2.3523,\ \ \lambda_{22} = 1.5432,\ C = 0.438524$$

Table 1. Optimum value of θ and percent relative efficiency of t_d or t_d^* with respect to $S^2_{y(n)}$.

Population	I	II
θ_0	0.520	0.524
PRE	103.40	104.78

We have computed the percent relative efficiencies of the proposed estimator t_d or t_d^* (under optimum condition) with respect to $S^2_{y(n)}$. Table 1 exhibits that the proposed estimator t_d or t_d^* is more efficient than the usual unbiased estimator $S^2_{y(n)}$ with substantial gain in efficiency. Hence, the proposed estimator t_d or t_d^* is to be preferred in practice.

References

[1] Jessen RJ (1942): Statistical investigation of a survey for obtaining farm facts. Iowa Agri. Expt. Statist. Res. Bull. 304.

[2] Yates F (1949): *Sampling methods for censuses and Surveys*. Charles Griffin and Co., London.

[3] Patterson HD (1950): Sampling on successive occasions with partial replacement of units. Jour. Roy. Statist. Soc. B(12), 214-255.

[4] Tikkiwal BD (1951): *Theory of successive sampling*. Thesis for Diploma, ICAR, New Delhi.

[5] Eckler AR (1955): *Rotation sampling*. Ann. Math. Statist. 26, 664-685.

[6] Rao JNK and Grahm JE (1964): Rotation design for sampling on repeated occasions. Jour. Ameri. Statist. Assoc. 59, 492-509.