International Journal of Chemical Studies

Chetan Patel
Assistant Professor, Department of Agriculture Mandsaur University, Mandsaur, Madhya Pradesh, India

Vinit Jaiswal
Assistant Professor, Department of Agriculture Mandsaur University, Mandsaur, Madhya Pradesh, India

RA Sharma
Director Agriculture, Department of Agriculture Mandsaur University, Mandsaur, Madhya Pradesh, India

KS Kushwaha
Retired Professor, Department of Mathematics and Statistics, College of Agriculture, JNKVV, Jabalpur, Madhya Pradesh, India

Use of jack – Knife technique and unbiased class of product type estimators in circular systematic sampling (C.S.S.) scheme

Chetan Patel, Vinit Jaiswal, RA Sharma and KS Kushwaha

Abstract
A general class of product type estimators in C.S.S. scheme is proposed to estimate the population mean of a response variable y. Jack – knife technique pioneered by quenouille (1949, 1956) is applied to make the class unbiased. The explicit expression for sampling variance of the class T_{pu} is derived to first order (n-1) of approximation. Minimum Variance Unbiased Estimator (M.V.U.E.) in the class T_{pu} is investigated. An empirical illustration is provided to examine the applicability of the results derived.

Keywords: Product estimator, Jack–knife technique, Circular System sampling, MVU estimator.

Introduction
The classical product estimator under linear systematic sampling scheme was proposed by Shukla (1971) [6] and its properties were studied. In general, the product estimator is biased, A weighted class of product type estimator was proposed and made exactly unbiased by Kushwaha and Singh (1989) [2] using Jack – knife technique.
A serious demerit of linear systematic sampling scheme is that it is not estimate the sampling variance of the estimator under study but it can be estimated unbiasedly with the use of interpenetrating systematic sampling technique with independent random start.
A sampling modification of linear systematic (L.S.S.) makes it possible to ensure a fixed sample size n and to make the sample mean unbiased for population mean, even in class if N Kn. This sampling scheme is known as 'circular systematic sampling' (CSS) scheme.
Murthy (1977) [3] and Sukhatame (1970) [8] have suggested to used CSS in situation when N = KN, where K is positive integer.

The main steps involved in selecting a sample using CSS are as follow:
1. Select the random number 'r' from 1 to N and name it as 'random start'.
2. Choose some integer value of K = N/n or take integer nearest to N/n and name it as skip or sampling span.
3. Select all units in the sample with serial numbers given on.

r+ jk if (r+jk) < N, {j = 0,1,2,3,…………(n-1)}, 1< r < N
r+ jk if (r+jk) > N, {j = 0,1,2,3,…………(n-1)}, 1< r < N

Sudhakar (1978) pointed out that the use of skip or span of sampling as an integer nearest to N/n in C.S.S. does not draw a sample of desired size. Sudhakar (1978) has also mentioned that if we take the span of sampling as nearest integer < N/n, we do not encounter the above cited difficulty although this depends upon n.
Jack–knife technique has been successfully employed in several estimation and testing problems. Here we have utilized Jake–knife technique to get rid of bias in usual product type estimators.
The expression for the sampling variance of the proposed class of estimators is derived. An empirical illustration is provided to examine the performance of the derived estimators with respect to efficiency point or views over other estimators existing in literature.

Corresponding Author:
Chetan Patel
Assistant Professor, Department of Agriculture Mandsaur University, Mandsaur, Madhya Pradesh, India

Material and methods

Let the population be consists of N units U = (U₁, U₂,....UN) labeled from (1 to N). Let N ≠ nk where k is an integer nearest to N/n. We select one sample at random and measure both the variables (Y, X) for each and every unit included in the sample. Let $(Y_{r+jk}, X_{r+jk}, j = 0, 1\ldots\ldots n-1$ and r= 1, 2,......N) be the sample units.

The circular systematic sample means (y, x) are defined as:

$$\bar{y} = \frac{1}{n}\sum_{j=0}^{n-1} y_{r+jk} \qquad j = 0, 1,\ldots\ldots(n-1)$$

$$\bar{x} = \frac{1}{n}\sum_{j=0}^{n-1} x_{r+jk} \qquad k = \text{sampling span}$$

(2.1)

Sample means $(\bar{y},\bar{x})$ are respectively unbiased estimators for the population means $(\bar{Y}, \bar{X})$. The X of covariate X is assumed to be known in prior.

The usual product estimator $\bar{Y}_p$ for $\bar{Y}$ based on a circular systematic sample of size n is defined as: -

$$\bar{Y}_p = \frac{\bar{y}}{\bar{x}}\bar{X} \qquad\qquad (2.2)$$

To reduce the bias of $\bar{y}_p$, we take n = gm and split the sample into g sub samples each of size m = n/g in a systematic manner as this avoids the need for selecting the samples in the form of sub sample of smaller size m, thereby retaining the efficiency generally obtained by taking a larger circular systematic sample of size n.

Let $(\bar{y}_t, \bar{x}_t,$ t = 1, 2,.....g) be unbiased estimators for $(\bar{y},\bar{x})$ based on circular systematic sub – sample each of size m units. Let us define complimentary sub – sample means $(\bar{y}'_t, \bar{x}'_t$, t = 1, 2,.....g) as:

$$\bar{y}'_t = \frac{(n\bar{y} - m\bar{y}_t)}{m'}$$

$$\bar{x}'_t = \frac{(n\bar{x} - m\bar{x}_t)}{m'}$$

(2.3)

Which are the sample means based on subsample, each of size m = (n-m) units obtaining by omitting t^{th} (t=1, 2,………,g) subsample from the circular systematic sample of size n. The estimators $(\bar{y}'_t, \bar{x}'_t)$ are also unbiased for $(\bar{Y},\bar{X})$ respectively. With this background, we define another product estimator reported as

$$\bar{y}'_{pt} = \frac{\bar{y}'_t}{\bar{x}'_t}\bar{X}$$

With its Jack-Knife version written as

$$\bar{y}'_p{}_\blacksquare = \frac{1}{g}\sum_{t=1}^{n} \bar{y}'_{pt}$$

(2.4)

The expression for bias of $(y_p , \bar{y}'_p)$ to the terms of order 0(n-1), are respectively made available as

$$B_1(\bar{y}_p) = \frac{\bar{y}}{n}\{1 + (n-1)\rho_w\}k^*c_x^2$$

$$B_1(\bar{y}'_p{}_\blacksquare) = \frac{\bar{Y}}{n}\{g + (n-g)\rho_w\}k^*c_x^2$$

(2.5)

Where $v(\bar{y})_{css} = \frac{1}{N}\sum_{r=1}^{N}(\bar{y}_r - \bar{Y})^2$

$$= \frac{1}{N}\sum_{r-1}^{N}(\frac{1}{n}\sum_{j=0}^{n-1} y_{r+jk} - \bar{Y})^2$$

$$= \frac{\sigma_y^2}{n}\{1 + (n-1)\rho_{yw}\}$$

And

$$v(\bar{x})_{css} = \frac{\sigma_x^2}{n}\{1 + (n-1)\rho_{xw}\}$$

By definition, the intra class correlation within the sample for the same character is define as

$$\rho_{yy} = \rho_{y\omega} = \frac{E(y_{r+jk} - \bar{Y})(y_{r+j'k} - \bar{Y})}{E(y_{r+jk} - \bar{Y})^2}$$

$$\rho_{y\omega} = \frac{\sum_{r=1}^{N}\sum_{j=0}^{n-1}\sum_{j<j}(y_{r+jk} - \bar{Y})}{E(y_{r+jk} - \bar{Y})^2}$$

$$\rho_{yy} = \rho_{y\omega} = \frac{E(y_{r+jk} - \bar{Y})(y_{r+jk} - \bar{x})}{\sqrt{E(y_{r+jk} - \bar{Y})^2 E(y_{r+jk} - \bar{x})^2}}$$

$$\text{COV}(\bar{y}, \bar{x}) = \frac{\rho_{yx}\sigma_y\sigma_x}{n}\sqrt{(1+(n-1)\rho_{y\omega})(1+(n-1)\rho_{x\omega})}$$

$$c_z = \frac{s_z}{\bar{z}} \quad, (Z = y. x)$$

$$k^* = \rho\frac{c_y}{c_x} \quad, \quad \rho = \rho_{yx}$$

Here ρ_{yx} is the population correlation coefficient and ρ_ω is the intra class correlation coefficient for both the variables (y,x) and has been assumed to be same for the both the variates (See Murthy, 1977, pp-374-375), (c_y, c_x) are the C. V's for (y, x) respectively.

Now, taking the linear combination of $\bar{y}$, $\bar{y}_p$ and $\bar{y}'_p$ we propose a weighted class of estimator Tp written as:

$$T_p = \omega_1\bar{y} + \omega_2\bar{y}'_{p\blacksquare} + \omega_{3\bar{y}_p} \quad, \quad \Sigma_{i=1}^3\omega'_{i=1} \tag{2.6}$$

Where (i=1, 2, 3,) are suitably chosen weights attached to different estimators. Now we have the following theorem-

Theorem (2.1):-the weighted class T_p of the estimator proposed in (2.6) would be unbiased if:

$hw_z + w_3 = 0$

For $h = \frac{g+(n-g)\rho_w}{(1+(n-1)\rho_w)}$ \tag{2.7}

Which can be proved easily.

If we take $\omega_1 = \omega, \omega_2 = \omega = \lambda$ and (1-ω- λ) the unbiasedness condition (2.7) reduced to:

$\lambda = -\frac{(1-\omega)}{(h-1)}$

Where (ω, λ) are constants to be chosen suitably. Thus, we obtain a general class T_p reduced to exactly unbiased product type estimators in C.S.S. scheme written as:

$$T_{PU} = w\bar{y} - \left(\frac{1-w}{h-1}\right)\bar{y}'_p + \left(\frac{1-w}{h-1}\right)h\bar{y}_p \tag{2.8}$$

Properties of class T_p

The expression for sampling variance of the proposed class T_p in (2.6) cal be written as

$$V(T_P)\ \omega_1^2 v(\bar{y}) + \omega_2^2 v(\bar{y}_{p\blacksquare}) + \omega_3^2 v(\bar{y}_p)\ 2\omega_1\omega_2 cov(\bar{y}, \bar{y}_{p\blacksquare}) + 2\omega_2\omega_3(\bar{y}'_p, \bar{y}_p)2\omega_3\omega_1(\bar{y}_p, \bar{y}) \tag{3.1}$$

Following the standard technique reported by Sukhatame *et al.* (1970, p.p. 162-64) [8] under large sample approximation to the terms or under o (n-1), the variance and covariance expressions for the various estimators in (3.1) cal be derived easily and cited in the following lemma (3.1):

$$v(\bar{y}) = \frac{\bar{Y}^2}{n}\{1+(n-1)\rho_\omega\}c_y^2$$
$$v(\bar{y}'_p) = (\bar{y}_p) = cov(\bar{y}'_p, \bar{y}_p)$$
$$= \frac{\bar{Y}^2}{n}\{1+(n-1)\rho_\omega\}[c_y^2 + (1+2k^*)c_x^2]$$
$$\text{Cov}(\bar{y}, \bar{y}'_{p\blacksquare}) = \text{Cov}(\bar{y}, \bar{y}'_p)$$
$$= \frac{\bar{y}^2}{n}\{1+(n-1)\rho_\omega\}(C_x^2 + k^* + C_X^2) \tag{3.2}$$

Putting the results (3.2) and $\omega_1 = \omega, \omega_2 = \lambda, \omega_3 = (1-\omega-\lambda)$ in simplifing it, the v(T$_{pu}$) can be obtained as:

$$V(T_{pu}) = \omega^2 V(\bar{y}) + (1-\omega)^2 V(\bar{y}_p) + 2\omega(1-\omega)cov(\bar{y}, \bar{y}_p)$$
$$= \frac{\bar{Y}^2}{n}\{1+(n-1)\rho_\omega\}[c_y^2 + (1-\omega)\{(1-\omega) + 2k^*\}c_x^2] \tag{3.3}$$

The V(T$_{pu}$) in (3.3) will be minimum for :-

$$\omega = (1+k^*) = \omega_0 \ (say) \tag{3.4}$$

Thus, minimum V(T$_{pu}$) is given as :-

$$\text{MinV}(T_{pu}) = \frac{\bar{y}^2}{n}\{1+(n-1)\rho_\omega\}(C_x^2 - k^{*2}C_X^2) \tag{3.5}$$

Which is equivalent to the approximate variance of usual biased linear regression estimator in C.S.S. scheme written as:

$$\bar{y}_{lr} = \bar{y} + b_{yx}(\bar{X} - \bar{x}) \tag{3.6}$$

Substituting $\omega_1 = \omega^* = (1+k^*), \omega_2 = \lambda^* = k^*/(h-1)$ and $\omega_3 = (1-\omega^*-\lambda^*) = -k^*h/(h-1)$ in (2.6) we obtain as optimum estimator T_p written as:

$$T_{PO} = (1+k^*)\bar{y} + \left(\frac{k^*}{h-1}\right)\bar{y}'_{p*} - \left(\frac{k^*h}{h-1}\right)\bar{y}_p \tag{3.7}$$

With variance given in (3.5)

It is to be pointed out that the class Tpu in (2.8) would be more efficient than the unbiased sample estimator y and the product estimator y defined under C.S.S. scheme according if either

$1 < \omega < (1+2k^*)$

\tag{3.8}

Or $(1+2k^*) < \omega < 1$

And either $0 < \omega < 2(1+k^*)$

\tag{3.9}

$2(1+k^*) < \omega < 0$

Remark 3.1:- The variance of any estimator obtained from the class T_{pu} (2.8) can be derived by substituting the appropriate value of constant in V (T_{PU}) in (3.3).

Numerical illustation

To see the performance of the class T_{pu}, we consider the data on y: the pounds of steam used monthly and X: the average atmospheric temperature in Fahrenheit. A machine point out (25 values) have been taken from Draper and smith (1966, p.p. 615-16). The summarized data are complied as follows.

N=25 $C_y^2 = 0.0299$ $\rho_{yx} = -0.8486$
N=10 $C_x^2 = 0.1077$ $k^* = -0.4450$
$\bar{y} = 9.424$ $S_{xy} = -23.79$ $\rho_\omega = -0.218$
$\bar{x} = 52.6$ k=2.5 ≈ 3(sampling span)

The interclass correlation coefficient ρ_ω has been worked out as:

$$\rho_\omega = 1 - \frac{n}{(n-1)}\frac{\sigma_\omega^2}{\sigma^2}$$

Where $(\sigma^2, \sigma_\omega^2)$ are the population and within sample variance respectively defined under CSS scheme as:

$$\sigma^2 = \frac{1}{N}\sum_{r=1}^{N}\sum_{j=0}^{n-1}(y_{r+jk} - \bar{y})^2$$

$$\sigma_\omega^2 = \frac{1}{N}\sum_{r=1}^{N}\sigma_{r\omega}^2 = \frac{1}{N}\sum_{r=1}^{N}\frac{1}{n}\sum_{j=0}^{n-1}(y_{r+jk} - \bar{y}_n)^2$$

We have worked out the value of

$$V_1(.) = \frac{V(T_{PU})}{\frac{\bar{y}^2}{n}\{1+(n-1)\rho_\omega\}} = \left[C_y^2 + (1+\omega)\{(1-\omega) + 2k^*\}C_x^2\right]$$

And P.R.E. $(., \bar{Y}) = \frac{V(\bar{y})}{V(T_{PU})} \times 100$, worked out and are complied in the table (4.1).

Table (4.1): value of $V_1(.)$ and P.R.E. $(., \bar{Y})$

Values of ω	Estimators	$V_1(.)$	P.R.E. $(., \bar{Y})$
………	$\bar{y}$	29.9X10⁻³	100.00
$\omega_o = 0.555$	T_{PUO}	8.6×10⁻³	347.67
0.950	T_{PU}	25.4×10⁻³	117.72
$(1+2k^*)<\omega<1.0$ }	T_{PU}	$<V_1(\bar{y})$	>100.00
$-0.1100<\omega 1.0$			
-0.075	T_{PU}	$<V_1(\bar{y})$	>100.00
………	$\bar{y}_p$ or $\bar{y}'_{p0}$	41.74×10⁻³	>100.00
-0.25	T_{PU}	78.3×10⁻³	38.18
$0<\omega<2(1+k^*)$	T_{PU}	$<V(\bar{y})$	>71.31
Or }			
$0<\omega<1.11$			
0.75	T_{PU}	12.6×10⁻³	237.31
1.05	T_{PU}	24.8×10⁻³	120.05

Table (4.1) reveals that P.R.E. (T_{puo}) is maximum among the estimators considered in the study which indicates that T_{puo} is the most efficient (optimum) estimator in the class Tpu In practice. One may substitute the estimated values of variances and covariance's in order to obtain a "near optimum values of ω". For the choice of ω in the interval (-0.1100< ω <1.0), the class Tpu is always more efficient than sample mean $\bar{y}$. It is also evident that the class Tpu in interval (0< ω <1.11), is more efficient than usual product estimator $\bar{y}_p$ as well as it jack-knife version $\bar{y}'_p$.

References

1. Draper NR, Smith H. Applied regression analysis. Wiley series in probabilities and mathematical statistics, New York, 1966.
2. Kushwaha KS, Singh HP. class of almost unbiased ratio and product type estimator in systematic sampling. J and Soc. Agric statistics. 1989; 41(2):193-205.
3. Murthy MN. Sampling theory and methods. Statistical pub soc Calcutta India, 1977.
4. Quenouille MH. Notes in bias in estimation. Biometric. 1956; 43:353-360
5. Quenouille MH. Approximation test of correlation in time series. J Roy Stat Soc. 1949; 11:68,84
6. Shukla ND. Systematic Sampling and product method of estimation. Proceeding of all India Seminar on Demography and Stats. B.H.U. Varanasi, India, 1971.
7. Sudhakar K. A note on circular systematic sampling design. Somhra and J. Stats: 1978; 40(C1)723-73, Calcutta
8. Sukhatame PV, Sukhatame BV. Sampling Theory of Surveys with Applications, Iowa State Unive Press Ames, Iowa USA, 1970.