Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v12n3p705

Linear systematic sampling with unequal sampling intervals in the presence of linear trend

By Subramani

Published: 20 November 2019

Electronic Journal of Applied Statistical Analysis
Vol. 12, Issue 03, November 2019, 705-719
DOI: 10.1285/i20705948v12n3p705

Linear systematic sampling with unequal sampling intervals in the presence of linear trend

Jambulingam Subramani

Department of Statistics, Pondicherry University
Puducherry-605014, India

Published: 20 November 2019

The present paper deals with the linear systematic sampling with unequal sampling intervals in the presence of linear trend among the population values. As a result, explicit expressions for the linear systematic sample means with different random starts in a labelled population with linear trend for a pre-assigned fixed sample size n and the population size N together its variance are obtained. The efficiencies of the proposed linear systematic sampling with that of simple random sampling without replacement, linear systematic sampling and diagonal systematic sampling schemes are assessed algebraically and also for certain natural populations. It is observed that the proposed linear systematic sampling performs better than the sampling schemes mentioned above.

keywords: Diagonal Systematic Sampling, Linear Systematic Sampling, Linear Trend, Simple Random Sampling, Trend Free Sampling, Yates Type End Corrections

1 Introduction

Consider a finite population $U = (U_1, U_2, .., U_N)$ with N distinct units and Y is a real variable with value Y_i measured on the population unit U_i,i=1,2,...,N giving a vector of N measurements. Let $Y = (Y_1, Y_2, .., Y_N)$.Let $Y_i = a + ib$,i=1,2,...,N a hypothetical population with a perfect linear trend among the population values. This population is also called as a labelled population, since the value Y_i depends on the label i. In general the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ is an unknown parameter and is to be estimated on

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

the basis of a random sample of size n selected from the finite population U. A random sample of size n is defined as an ordered sequence $S = (u_1, u_2, .., u_n) = (U_{i1}, U_{i2}, .., U_{in})$, $1 \leq i \leq N and 1 \leq l \leq n$. Several sampling schemes, like simple random sampling without replacement (SRSWOR), linear systematic sampling (LSS) schemes are available in the literature for selecting a sample of fixed size n from a finite population of size N. For the labelled population discussed above with N=kn, the LSS is recommended for selecting a sample of fixed size n. Further it is shown algebraically (Cochran, 2007) that the LSS sample mean is better than the SRSWOR sample mean in estimating the finite population mean in the presence of linear trend. Further improvement on LSS sample mean can be achieved by introducing changes in the estimator itself like Yates type end corrections (Yates, 1948). Subramani (2000) has introduced diagonal systematic sampling (DSS) as an alternative to LSS and proved that DSS performs better than LSS and SRSWOR for estimating the population mean in the presence of linear trend. Later Subramani (2009), Subramani (2010) has extended the DSS scheme and introduced generalized diagonal systematic sampling (GDSS) for estimating the finite population mean with a linear trend among the population values. For further discussions on estimating the finite population mean in the presence of linear trend the readers are referred to Bellhouse and Rao (1975), Bellhouse (1984), Bellhouse (1988), Chang and Huang (2000), Cochran (2007), Fountain and Pathak (1989), Gupta and Kabe (2011), Khan et al. (2013), Khan et al. (2014), Khan et al. (2015), Madow (1953), Mukerjee and Sengupta (1990), Murthy et al. (1967), Murthy and Rao (1988),Singh et al. (1968), Subramani (2012). Subramani (2013) Subramani (2009), Subramani (2014) Subramani et al. (2014), Subramani and Gupta (2014),Subramani et al. (2014), Subramani and Singh (2014), Subramani and Tracy (1999) , Sukhatme et al. (1970) and the references cited there in. For the sake of ready reference the procedure of selecting random samples from the LSS and DSS schemes are explained with the help of numerical examples.

1.1 Linear Systematic Sampling (LSS) Scheme

If N=nk, the first unit is selected at random and the remaining units get selected automatically according to some pre assigned patterns is known as linear systematic sampling. The steps involved in LSS for selecting a sample of size n with sampling interval k are given below: Arrange the N population units $U = (U_1, U_2, .., U_N)$ in a linear array Select a random number r such that $1 \leq r \leq k$ For selecting a linear systematic sample of size n select every $k^t h$ elements from the random start r in the linear array until n elements are accumulated The selected units $U = (U_r, U_{r+k}, , , , U_{r+(n-1)k})$ be the linear systematic sample of size n for the random start r. The variance of the linear systematic sample mean is obtained as given below: $V(\bar{y}_{lss}) = \frac{1}{k} \sum_{i=1}^{k} (\bar{Y}_i^2 - \bar{Y}^2)$

For the labelled population and N=kn one may get the explicit expression for the variances of SRSWOR and LSS sample means as given below:

Equation (1.1)

$$V(\bar{y}_r) = \frac{(k-1)(N+1)b^2}{12}$$

Equation (1.2)

$$V(\bar{y}_{lss}) = \frac{(k^2-1)b^2}{12}$$

Example 1.1: The procedure of obtaining the linear systematic samples is explained for the fixed value of sample size n = 3 and the population size N = 18. If $N = 18$ and $n = 3$ then $k = 6$. That is $k = [N/n] = [18/3] = 6$ The selected LSS samples, their means, expected value and the variance are given for the sampling interval $k = 6$ in the following table:

Table 1: LSS Samples and their Means for the Sampling Interval $k = 6$

Random Start	**Sample Values**	**LSS Mean**
1	1 7 13	7
2	2 8 14	8
3	3 9 15	9
4	4 10 16	10
5	5 11 17	11
6	6 12 18	12
Total		57

For the case of $N = 18$, $n = 3$ and $k = 6$, it is obtained that $E(\bar{y}_{lss}) = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_i = \frac{57}{6} = 9.5 = \bar{Y}$

$V(\bar{y}_{lss}) = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_i^2 - \bar{Y}^2 = \frac{559}{6} - 9.5^2 = 93.16667 - 90.25 = 2.916667$

The value $V(\bar{y}_{lss}) = 2.916667$ is coincided with the value obtained in (1.2)

2 Diagonal Systematic Sampling Scheme(DSS)

The diagonal systematic sampling scheme consists of drawing n units from the matrix M of order nxk, systematically such that the selected n units are the diagonal elements or broken diagonal elements of matrix M. The steps involved in DSS for selecting a random sample of size n with sampling interval k are given below.

Arrange the N population units $U = U_1, U_2, .., U_N$ in an nxk matrix Select a random number r such that $1 \leq r \leq k$
The selected sampling units are
$S_r = U_r, U_{r+k}, ..., U_{r+(n-1)k} if r \leq k - n + 1$
$S - r = U_r, U_{r+k}, ..., U_t(k+1) + r = (t+1)k, U_{(t+1)k+1}, .., U_{(n-1)k+(n-t-1)}$
if $r + n - 1 > k$ then $r + n - 1$ is reduced to $\mod k$.

The variance of the diagonal systematic sample mean is obtained as given below:

Equation (1.3)

$$V(\bar{y}_{dss}) = \frac{1}{k} \sum_{i=1}^{k} (\bar{y}_i - \bar{Y})^2$$

For the labelled population $Y_i = a + ib, i = 1, 2, ..., N$ and $N = kn$ one may get the explicit expression for the variance given in (1.3) as:

Equation (1.4)

$$V(\bar{y}_{dss}) = \frac{(k-n)[n(k-n)+2]b^2}{12n}$$

When $N = 2k$ and $n = 2$, the variance expression (1.4) is reduced to

Equation (1.5)

$$V(\bar{y}_{dss}) = \frac{(k-1)(k-2)}{12n} b^2$$

Example 1.2: The procedure of obtaining the diagonal systematic samples is explained for the fixed value of sample size $n = 3$ and the population size $N = 18$. If $N = 18$ and $n = 3$ then $k = 6$. That is $k = [N/n] = [18/3] = 6$
The selected DSS samples, their means, expected value and the variance are given for the sampling interval $k = 6$ in the following table:

For the case of $N = 18$, $n = 3$ and $k = 6$, it is obtained that $E(\bar{y}_{dss}) = \frac{1}{k} \sum_{i=1}^{k} \bar{y}_i = 57/6 = 9.5 = \bar{Y}$
$V(\bar{y}_{dss}) = \frac{1}{k} \sum_{i=1}^{k} (\bar{y}_i^2 - \bar{Y})^2 = 547/6 - 9.5^2 = 91.16667 - 90.25 = 0.916667$

The value $V(\bar{y}_{dss}) = 0.916667$ is coincided with the value obtained in (1.4).

If the population size N is a multiple of sample size $n(N = kn)$ then LSS and DSS are providing better estimators than the SRSWOR for the population mean in the presence of linear trend. It is to be noted that the sampling interval is fixed in both the LSS and DSS schemes. However the sampling interval is fixed as k in LSS whereas the actual sampling interval is fixed as $k + 1$ in DSS but it varies sometimes. Now, the problems in the systematic sampling are the following: The choice for the sampling interval k.

Table 2: LSS Samples and their Means for the Sampling Interval k=6

Random Start	**Sampled Units**	**Sample Values**	DSS Mean
1	U_1 U_8 U_{15}	1 8 15	8
2	U_2 U_9 U_{16}	2 9 16	9
3	U_3 U_{10} U_{17}	3 10 17	10
4	U_4 U_{11} U_{18}	4 11 18	11
5	U_5 U_{12} U_{19}	5 12 19	10
6	U_6 U_{13} U_{20}	6 13 20	9
Total			57

Is it necessary to choose the equal sampling interval between the selected units? Is it possible to choose unequal sampling intervals between the selected units? If yes, what will be the choices for the unequal sampling intervals? If unequal sampling intervals are possible, will they ensure the simplicity of the systematic sampling; distinct units in the samples and maintain the minimum variance. Is it possible to derive explicit expression for the variance of the proposed sampling which is useful to assess the efficiency with other sampling schemes like LSS, DSS and SRSWOR algebraically? The points noted above are motivating the present study, which deals with the following:

1. To propose a linear systematic sampling with unequal sampling intervals, which addresses all the problems raised above;
2. To derive the explicit expressions for the proposed linear systematic sample means and its variance for the population with a perfect linear trend among the population values;
3. To derive the explicit expressions for the proposed linear systematic sample means and its variance for the population with a perfect linear trend among the population values;
4. To assess the relative performance of proposed systematic sampling with that of simple random sampling without replacement, linear systematic sampling and the diagonal systematic sampling algebraically and also for certain natural populations.

3 Proposed Linear Systematic Sampling with Unequal Sampling Intervals (LSSU)

As stated earlier, the LSS is used when the population size N is a multiple of sample size $n(N = kn)$ for selecting a sample of fixed size n with a fixed sampling interval k.

That is, the distance between any two successive (selected) units in the LSS sample is the same. In the proposed LSS with unequal sampling intervals, the distance between any two selected units is not the same. The steps involved in LSSU for selecting a random sample of size n with unequal sampling interval are given below: Arrange the N population units $U = U_1, U_2, .., U_N$ in a linear array Select a random number r such that $1 \leq$ r $\leq$ k. For selecting a linear systematic sample with unequal sampling intervals of size n select every elements from the random start r in the linear array until n elements are accumulated, where $i = 1, 2, 3, ..., (n-1)$
The labels for selected sampling units are given below:

$$i_1 = r, i_2 = i_1 + k + 1, i_3 = i_2 + k + 2, ..., i_n = i_n - 1 + k + (n-1)$$

If $i_{l+1} > (l+1)k$ then select the item corresponding to $i_{i+1} - K$. That is, the value of the label is reduced by k. Then

$$U_{i_1}, U_{i_2}, U_{i_l}, U_{i_l+1}, ..., U_{i_n}$$

is the LSSU sample of size n for the random start r. The variance of the LSSU sample mean is obtained as given below:

Equation 2.1

$$V(\bar{y}_{lssu}) = \frac{1}{k}\sum_{i=1}^{k}(\bar{y}_i - \bar{Y})^2 = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_i^2 - \bar{Y}^2$$

The diagrammatic representations of the selection of LSSU samples for $N = 18, k = 6, n = 3$ with different random starts are given below:

	1	2	3	4	5	6
1	X					
2		X				
3				X		

Figure 1: When the random start is r = 1

	1	2	3	4	5		6
1			X				
2				X			
3							X

Figure 2: When the random start is r = 3

	1	2	3	4	5		6
1				X			
2					X		
3	X						

Figure 3: When the random start is r = 4

	1	2	3	4	5	6
1						X
2	X					
3			X			

Figure 4: When the random start is r = 6

Remark 2.1: The method discussed in this paper maintains distinct sampling intervals between the selected units in the sample and also the simplicity of the linear systematic sampling. The other choices may be the sampling intervals are in arithmetic progression, geometric progression or Fibonacci series. However the problems are the

proposed methods have to maintain the simplicity of the LSS as well as the derivation of explicit expression for the sample mean ant its variance so as to compare the efficiency with other existing competitive estimators.
For the sake of the convenience of the readers the above said procedure for drawing LSSU samples is explained with the help of a numerical example as given below.

Example 2.1: The procedure of obtaining the LSSU samples is explained for the fixed values of sample size n and the population size N and N=kn. If $N = 18$ and $n = 3$ then $k = 6$. That is $k = [N/n] = [18/3] = 6$ The selected LSSU samples, their means, expected value and the variance are given in the following table:

Table 3: LSS Samples and their Means

Random Start	**Sampled Units**	**Sample Values**	DSS Mean
1	$U_1U_8U_{16}$	1 8 16	8.3333
2	$U_2U_9U_{17}$	2 9 17	9.3333
3	$U_3U_{10}U_{18}$	3 10 18	10.3333
4	$U_4U_{11}U_{13}$	4 11 13	9.3333
5	$U_5U_{12}U_{14}$	5 12 14	10.3333
6	$U_6U_{13}U_{15}$	6 7 15	9.3333
Total			57.0000

For the case of $N = 18$, $n = 3$ and $k = 6$, it is obtained that:

$$E(\bar{y}_{lssu}) = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_i = 57/6 = 9.5 = \bar{Y}$$

$$V(\bar{y}_{lssu}) = \frac{1}{k}\sum_{i=1}^{k}(\bar{y}_i^2 - \bar{Y})^2 = 544.33333/6 - 9.5^2 = 90.72222 - 90.25 = 0.472222$$

4 Computation of LSSU Sample Means

Consider the labelled population with the N population values $Y_j = a+jb, j = 1, 2, ..., N$.

Equation(2.2)
The population mean is $\bar{Y} = a + [((N + 1))/2]b$. For the labelled population defined in (2.2), the LSSU sample means are obtained as:

$$\bar{y}_{i.} = a + [i + \frac{(n-1)(3k+n+1)}{6}]b, i = 1, 2, 3...k\frac{n(n-1)}{2} = L_O(say)$$

$$a+[i+\frac{(n-1)(3k+n+1)}{6}-\frac{k}{n}]b, i=L_0+1, L_0+2, ..., L_0+(n-1)=L_1$$

$$a+[i+\frac{(n-1)(3k+n+1)}{6}-\frac{2k}{n}]b, i=L_0+1, L_0+2, ..., L_0+(n-2)=L_2$$

$$a+[i+\frac{(n-1)(3k+n+1)}{6}-\frac{3k}{n}]b, i=L_0+1, L_0+2, ..., L_0+(n-3)=L_3$$

$$a+[i+\frac{(n-1)(3k+n+1)}{6}-\frac{(n-1)k}{n}]b, i=1,2,3...k\frac{n(n-1)}{2}=L_{n-2}+1=L_{(n-1)}$$

Where:

Equation(2.3)

$L_0=(k-\frac{n(n-1)}{2})$ and $L_i-L_(i-1)=(n-i), i=1,2,3,...,(n-1)$

Remark 2.2: Since $L_0=(k-\frac{n(n-1)}{2})$ and $L_i-L(i-1)=(n-i), i=1,2,....,(n-1)$, by successive substitution one may get $L_i=L_o+ni-i(i+1)/2$ and $L(n-1)=k$

From the above expressions the average of the LSSU sample means is obtained as

$$\frac{1}{k}\sum_{i=1}^{k}\bar{y_i}=\frac{1}{k}\left\{\sum_{i=1}^{k}\left[a+\left[i+\frac{(n-1)(3k+n+1)}{6}\right]b\right]-\frac{kb}{n}\sum_{n-1}^{j=1}\sum_{n-j}^{i=1}j\right\}$$

$$a+\left(\frac{(k+1)}{2}+\frac{(n-1)(3k+n+1)}{6}\right)b-\left(\frac{n(n-1)}{2}+\frac{(n-1)(2n-1)}{6}\right)b$$

$$=a+\left(\frac{3(k+1)}{6}+\frac{(n-1)(3k+n+1)}{6}-\frac{3n(n-1)}{6}+\frac{(n-1)(2n-1)}{6}\right)b$$

$$=a+\left(\frac{3(k+1)}{6}+\frac{(n-1)(3k+n+1-3n+2n-1)}{6}\right)b$$

$$=a+\left(\frac{3(k+1)}{6}+\frac{(3k(n-1)}{6}\right)b$$

$$=a+\left(\frac{3(kn+1)}{6}\right)b$$

That is,

$$\frac{1}{k}\sum_{i=1}^{k}\bar{y_i}=a+\left(\frac{N+1}{2}\right)b=\bar{Y}$$

That is, the LSSU sample mean is an unbiased estimator for its population mean.

5 Computation of Variance of LSSU Sample Mean

For the labelled population and the corresponding LSSU sample means defined in Section 2.1, the derivation of the variance of LSSU sample mean is given below. By substituting the LSSU sample means and the population mean in (2.1), the variance of LSSU sample mean for the labelled population is obtained as:

Equation(2.5)

$$V(\bar{y}_{lssu}) = \frac{b^2}{k}\left\{\sum_{i=1}^{N}\left(i + \frac{(n-1)(3k+n+1)}{6}\right)^2 + \frac{k^2}{n^2}\right\}$$

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-j} j^2 - \frac{2k}{n}\sum_{j=1}^{n-1}\sum_{i=L_{(j-1)}+1}^{L_j} j\left(i + \frac{(n-1(3k++1))}{6} - \frac{k(N=1)^2}{4}\right)$$

After a little algebra, the variance of LSSU sample mean is obtained as:

$$V(\bar{y}_{lssu}) = \left\{\frac{(k+1)(2k+1)}{6} + \frac{(n^2-3k-4)(n^2-3k+2)}{36} - \frac{(n^2-1)(15k+n^2-4)}{180}\right\} b^2$$

By simplifying the above expression one may get the **Equation(2.6)**:

$V(\bar{y}_{lssu}) = \left\{\frac{(k-1)(k+1)}{12} + \frac{(n^2-1)(15k-4n^2+1)}{180}\right\} b^2$

The variance expression given in (2.6) can be rewritten as:

Equation(2.7)

$V(\bar{y}_{lssu}) = V(\bar{y}_{lss})\left\{\frac{(n^2-1)(15k-4n^2+1)}{180}\right\} b^2$

Remark 2.2: By setting $n = 2$ in (2.6), the variance expression is reduced to:

$$V(\bar{y}_{lssu}) = \left\{\frac{(n^2-1)(15k-4n^2+1)}{180}\right\} b^2$$

The variance expression (2.8) is exactly the same as the variance of DSS sample mean given in (1.5). That is, when n = 2 the efficiency of DSS and LSSU is the same.

6 Comparison of the Efficiency of LSSU with LSS and SRSWOR Sample Means

By comparing the variance expressions for SRSWOR sample mean (1.1) and a LSSU sample mean (2.6) one can easily show that:

$$V(\bar{y}_{lss}) - V(\bar{y}_{lssu}) = \frac{(k-1)(k+1)}{12} - \left\{ \frac{(k-1)(k+1)b^2}{12} - \frac{(n^2-1)(15k-4n^2+1)}{180} \right\} b^2$$

$$\left\{ \frac{(N-k)(N+1)b^2}{12} + \frac{(n^2-1)(15k-4n^2+1)}{180} \right\} b^2$$

By comparing the variance expressions for LSS sample mean (1.2) and a LSSU sample mean (2.7) one can easily show that:

Equation(3.2)

$$V(\bar{y}_{lss}) - V(\bar{y}_{lssu}) = \left\{ \frac{(n^2-1)(15k-4n^2+1)}{180} \right\} b^2 \geq 0$$

By comparing the variance expressions for DSS sample mean (1.4) and a LSSU sample mean (2.6) one can easily show that:

Equation(3.3)

$$V(\bar{y}_{dss}) - V(\bar{y}_{lssu}) = \left\{ \frac{(n^2-1)(n-2)(15k-4n^2+8n)}{180n} \right\} b^2 \geq 0$$

Subramani (2000) has already shown that for the labelled populations $Y_i = a + ib, i = 1, 2, ..., N$ and $N = kn$ the following inequalities given in (3.4) are always hold good.

$$V(\bar{y}_{dss}) \leq V(\bar{y}_{lss}) \leq V(\bar{y}_r)$$

From (3.3) and (3.4) it is obtained that the following inequalities given in (3.5) are always true for the labelled population defined in Section 1.1.

$$V(\bar{y}_{lssu}) \leq V(\bar{y}_{dss}) \leq V(\bar{y}_{lss}) \leq V(\bar{y}_r)$$

7 Yates Type End Corrections on Circular Systematic Sample Means

It has been shown in Section 3 that the LSSU performs better than the SRSWOR, LSS and DSS. However it is not a trend free sampling (Mukerjee and Sengupta, 1990) which can be achieved by introducing Yates type end corrections (Yates, 1948) as given below:

The modification involves the usual LSSU but the modified sample mean is defined as:

$$V(\bar{y}^*_{lssu}) = \bar{y}_{lssu} + \alpha(y_1 - y_n)$$

That is, the first and the last units in the selected samples are given the weights $n^{-1} + \alpha$ and $n^{-1} - \alpha$ respectively whereas the remaining units get the same weight n^{-1}.

By equating $\bar{y}^*_{lssu} = \bar{Y}$ for the population with a perfect linear trend, one may get the values for α from (4.1). Here one may have the following two situations:

(i). The random start i is less than or equal to $k - \frac{n(n-1)}{2} = L_0$. (ii). The random start i is greater than $k - \frac{n(n-1)}{2} = L_0$

Case (i): When the random start i is less than or equal to $k - \frac{n(n-1)}{2} = L_0$.
By setting (4.1) is equal to $\bar{Y}$ one may get:

$$\left[i + \frac{(n-1)(3k+n+1)}{6}\right] + \alpha(y_1 - y_n) = \frac{(N+1)}{2}, i = 1, 2, ..., k - \frac{n(n-1)}{2}$$

By putting:

$$(y_1 - y_n) = i - (i + (n-1)k + \frac{n(n-1)}{2}) = -(k(n-1) + \frac{n(n-1)}{2})$$

Equation(4.2)

$$\alpha = \frac{6i + (n^2-1) - 3(k+1)}{3(n-1)(2k+n)}, i = 1, 2, ..., k - \frac{n(n-1)}{2} = L_0$$

Case (ii): When the random start i is greater than:

$$k - \frac{n(n-1)}{2} = L_0$$

Let the random start i lies between L_{m-1} and L_m. By setting (4.1) is equal to $\bar{Y}$ we get:

$$\left[i + \frac{(n-1)(3k+n+1)}{6} - \frac{mk}{n}\right] + \alpha(y_1 - y_n) = \frac{(N+1)}{2},$$

i = $L_{m-1} + 1$ to L_m. By putting:

$$(y_1 - y_n) = i - (i + (n-1)k + \frac{n(n-1)}{2} - mk) = \frac{2mk - (n-1)(2k+n}{2})$$

$$\alpha = \frac{n[6i + (n^2-1) - 3(k+1)] - 6mk}{3n[(n-1)(2k+n) - 2mk]}$$

i = L_{m-1} + 1 to L_m.

Remark 4.1: In the presence of a perfect linear trend the modified LSSU sample mean $\bar{y}^*_{lssu}$ becomes the population mean $\bar{Y}$ and hence the $V(\bar{y}^*_{lssu}) = 0$. In this case, the LSSU becomes a completely trend free sampling, see (Mukerjee and Sengupta, 1990).

8 Numerical Comparison of LSSU for certain Natural Population

It has been shown in Section 3 that the proposed sampling scheme LSSU performs well, compared to simple random sampling, linear systematic sampling and diagonal systematic sampling schemes whenever there exists a perfect linear trend among the population values. In fact this is an unrealistic assumption in real life situations. Consequently an attempt has been made to study the efficiency of LSSU sampling for certain natural populations with a linear trend among the population values but not a perfect linear trend. In this connection the populations given in Subramani (2014) are considered. The first 5 populations are pertaining to the road accidents occurred in Tamilnadu, India during the years from 1990 to 2004 and its impacts. The data are classified as 5 populations of size 15 each, which measured the number of accidents; number of persons killed; number of registered motor vehicles and so on. The population 6 is the data, collected for assessing the process capability of the turning operation performed on the component Torsion bar in Frontier CNC Lathe Machine, one of the key components in integrated power steering system, from an auto ancillary manufacturing unit located in Tamilnadu. The 50 measurements based on the order of the production are collected and are arranged in an ascending order to create a linear trend and considered the first 48 measurements for assessing the relative performances of various sampling schemes. When the following 3 combinations (24, 2), (16, 3), (12, 4) for are considered. The simple random sampling, linear systematic sampling, diagonal systematic sampling and linear systematic sampling with unequal sampling intervals are used for computing the variances and also the percentage relative efficiencies of the proposed LSSU estimators and the results are presented in Table 5.1. The percentage relative efficiency (PRE) of the proposed estimator (p) with respect to an existing estimator (e) is computed as $PRE(p) = \frac{V(e)}{V(p)} \times 100$.

Table 4: Comparison of simple random, linear systematic and diagonal systematic and LSSU sample means for the populations given in Subramani and Singh [29]

popln	**k**	**n**	V($\bar{y_r}$)	V($\bar{y_{lss}}$)	$V(\bar{y_{dss}})$	$V(\bar{y_{lssu}})$	$\bar{y_r}$	$\bar{y_{lss}}$	$\bar{y_{dss}}$
1	5	3	728170278.23	161361946.86	14531417.48	31335121.53	2323.82	514.96	46.37
1	5	3	37525740.30	16299828.15	4859489.57	1600207.88	2345.05	1018.61	303.68
3	5	3	76934478.99	30829696.77	9953615.71	4390654.82	1752.23	702.17	226.70
4	5	3	195.92	112.60	47.09	26.11	750.26	431.22	180.33
5	5	3	6.89	2.66	0.73	0.51	1342.67	518.21	142.90
6	24	2	41.05	19.62	16.79	16.79	244.50	116.88	100.00
6	16	3	31.33	12.64	9.47	7.64	410.22	165.47	124.01
6	12	4	24.55	5.14	2.48	1.02	2409.44	504.77	243.10

It is seen from the table values that the proposed LSSU is the most efficient. By comparing with SRSWOR, the PREs are ranging from 244.50 to 2409.44 whereas the PREs are ranging from 116.88 to 1018.61 and from 46.37 to 303.68 respectively for LSS and DSS sampling schemes. In general, it is observed that for the populations with a linear trend among the population values, the following inequalities are always true $V(\bar{y}_{lssu}) \leq V(\bar{y}_{dss}) \leq (\bar{y}_{lss}) \leq (\bar{y}_r)$.

Acknowledgements

The author wishes to record his gratitude to the reviewer and the Editors, particularly Prof. M.R. Srinivasan for their constructive comments, which has improved the presentation of the paper.

References

Bellhouse, D. (1984). On the choice of the sampling interval in circular systematic sampling. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 247–248.

Bellhouse, D. (1988). Systematic sampling, handbook of statistics, vol. pr krisnaiah, cr rao.

Bellhouse, D. and Rao, J. (1975). Systematic sampling in the presence of a trend. *Biometrika*, pages 694–697.

Chang, H.-J. and Huang, K.-C. (2000). Remainder linear systematic sampling. *Sankhyā: The Indian Journal of Statistics, Series B*, pages 249–256.

Cochran, W. G. (2007). *Sampling techniques.* John Wiley & Sons.

Fountain, R. L. and Pathak, P. K. (1989). Systematic and nonrandom sampling in the presence of linear trends. *Communications in statistics-Theory and Methods*, 18(7):2511–2526.

Gupta, A. K. and Kabe, D. G. (2011). *Theory of sample surveys.* World Scientific.

Khan, Z., Gupta, S., and Shabbir, J. (2014). A note on diagonal circular systematic sampling. *Journal of Statistical Theory and Practice*, 8(3):439–443.

Khan, Z., Shabbir, J., and Gupta, S. (2013). A new sampling design for systematic sampling. *Communications in Statistics-Theory and Methods*, 42(18):3359–3370.

Khan, Z., Shabbir, J., and Gupta, S. (2015). Generalized systematic sampling. *Communications in Statistics-Simulation and Computation*, 44(9):2240–2250.

Madow, W. G. (1953). On the theory of systematic sampling, iii. comparison of centered and random start systematic sampling. *The Annals of Mathematical Statistics*, pages 101–106.

Mukerjee, R. and Sengupta, S. (1990). Optimal estimation of a finite population mean in the presence of linear trend. *Biometrika*, pages 625–630.

Murthy, M. and Rao, T. (1988). Systematic sampling with illustrative examples. *Handbook of statistics*, 6:147–185.

Murthy, M. N. et al. (1967). Sampling theory and methods. *Sampling theory and methods.*

Singh, D., Jindal, K., and Garg, J. (1968). On modified systematic sampling. *Biometrika*, 55(3):541–546.

Subramani, J. (2000). Diagonal systematic sampling scheme for finite populations. *Journal of the Indian Society of Agricultural Statistics*, 53(2):187–195.

Subramani, J. (2009). Further results on diagonal systematic sampling scheme for finite populations. *Journal of the Indian Society of Agricultural Statistics*, 63(3):277–282.

Subramani, J. (2010). Generalization of diagonal systematic sampling scheme for finite populations. *Model assisted statistics and applications*, 5(2):117–128.

Subramani, J. (2012). A modification on linear systematic sampling for odd sample size. *Bonfring International Journal of Data Mining*, 2(2):32–36.

Subramani, J. (2013). A modification on linear systematic sampling. *Model Assisted Statistics and Applications*, 8(3):215–227.

Subramani, J. (2014). Star-type systematic sampling schemes for finite populations. *Communications in Statistics-Theory and Methods*, 43(1):175–190.

Subramani, J. and Gupta, S. N. (2014). Generalized modified linear systematic sampling scheme for finite populations. *Hacettepe Journal of Mathematics and Statistics*, 43(3):529–542.

Subramani, J., Gupta, S. N., and Prabavathy, G. (2014). Circular systematic sampling in the presence of linear trend. *American Journal of Mathematical and Management Sciences*, 33(1):1–19.

Subramani, J. and Singh, S. (2014). Estimation of population mean in the presence of linear trend. *Communications in Statistics-Theory and Methods*, 43(15):3095–3116.

Subramani, J. and Tracy, D. (1999). Determinant sampling scheme for finite populations. *International Journal of Mathematical and Statistical Sciences*, 8:27–42.

Sukhatme, P., Sukhatme, B., Sukhatme, S., and Asok, C. (1970). Sampling theory of surveys with applications. ames.

Yates, F. (1948). Systematic sampling. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 241(834):345–377.