Journal of the Korean
Data & Information Science Society
2000, Vol. 11, No.1, pp. 91 ∼ 101

A New Estimator of Population Mean Based on Centered Balanced Systematic Sampling[1]

Hyuk Joo Kim[2]

Abstract

We propose a new method for estimating the mean of a population which has a linear trend. The suggested estimator is based on the centered balanced systematic sampling method and the concept of interpolation and extrapolation. The efficiency of the proposed method is compared with that of existing methods.

Key Words and Phrases : Population with a linear trend, Centered balanced systematic sampling, Interpolation, Extrapolation

1. Introduction

In performing sample surveys, we occasionally encounter a population with a linear trend. For example, suppose we wish to estimate the average sales of the supermarkets in a certain city. If we arrange the supermarkets in that city in increasing or decreasing order of the number of employees, there is expected to be a linear trend in this population.

In estimating the mean of a population which has a linear trend, ordinary systematic sampling (OSS) is known to be much better than simple random sampling (SRS). Several versions of systematic sampling have been suggested so far. Among them, end corrections (EC) method proposed by Yates (1948), centered systematic sampling (CSS) proposed by Madow (1953), balanced systematic sampling (BSS) proposed by Sethi (1965) and Murthy (1967), and modified systematic sampling (MSS) proposed by Singh et al. (1968) are well-known methods. Combining the concepts of CSS and BSS, Kim (1985) proposed centered balanced systematic sampling (CBSS). Kim (1998, 1999a, 1999b) also developed some estimation methods by applying the concepts of interpolation and extrapolation to BSS and MSS.

---

[1]This paper was supported by Wonkwang University in 1999.

[2]Associate Professor, Division of Mathematical Science, Wonkwang University, Iksan, Chonbuk, 570-749, Korea

In this paper, a new method is developed for estimating the mean of a population which has a linear trend. The method, based on CBSS, interpolation and extrapolation, will be developed for use in the case when the sample size n (≥ 3) is an odd number and k (the reciprocal of the sampling fraction) is an even number, and will be compared with existing methods under the expected mean square error criterion based on the infinite superpopulation model introduced by Cochran (1946).

2. Development of the method

Suppose we have a population of size $N = kn$, the units of which are denoted by $U_1, U_2, \cdots, U_N$. We wish to select a sample of size n from this population.

Let us define the clusters C_i' $(i = 1, 2, \cdots, k)$ by

$$C_i' = \{U_{i+2(j-1)k} : j = 1, 2, \cdots, n/2\} \cup \{U_{2jk+1-i} : j = 1, 2, \cdots, n/2\}$$

for n even, and

$$C_i' = \{U_{i+2(j-1)k} : j = 1, 2, \cdots, (n+1)/2\} \cup \{U_{2jk+1-i} : j = 1, 2, \cdots, (n-1)/2\}$$

for n odd. For example, if $N = 20$, $n = 5$ and $k = 4$, then the four clusters are $C_1' = \{U_1, U_8, U_9, U_{16}, U_{17}\}, C_2' = \{U_2, U_7, U_{10}, U_{15}, U_{18}\}, C_3' = \{U_3, U_6, U_{11}, U_{14}, U_{19}\}$, and $C_4' = \{U_4, U_5, U_{12}, U_{13}, U_{20}\}$.

We briefly review the CBSS method proposed by Kim (1985). If k is odd, $C'_{(k+1)/2}$ is selected. So CBSS is the same as CSS in this case. If k is even (let us consider only this case from now on), either $C'_{k/2}$ or $C'_{k/2+1}$ is selected with respective probability 1/2.

The sample mean $\overline{y}_{CBSS}$ obtained by CBSS has mean square error

$$MSE(\overline{y}_{CBSS}) = \frac{1}{2}\{(\overline{y}'_{k/2} - \overline{Y})^2 + (\overline{y}'_{k/2+1} - \overline{Y})^2\},$$

where $\overline{y}_i'$ is the mean value for the units in C_i' $(i = 1, 2, \cdots, k)$.

Throughout this paper the following notation will be used :

y_i : value for the ith unit in the population $(i = 1, 2, \cdots, N)$,

$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$: population mean to be estimated,

y_{ij}' : value for the jth unit in C_i' $(i = 1, 2, \cdots, k;\ j = 1, 2, \cdots, n)$, that is,

$$\begin{aligned} y_{ij}' &= y_{i+(j-1)k} \ (j = 1, 3, 5, \cdots, n-1) \\ y_{ij}' &= y_{1-i+jk} \ (j = 2, 4, 6, \cdots, n) \end{aligned}$$

for n even, and

$$y'_{ij} = y_{i+(j-1)k} \; (j = 1, 3, 5, \cdots, n)$$
$$y'_{ij} = y_{1-i+jk} \; (j = 2, 4, 6, \cdots, n-1)$$

for n odd,

$\overline{y}'_i = \frac{1}{n}\sum_{j=1}^{n} y'_{ij}$: mean for the units in C'_i $(i = 1, 2, \cdots, k)$.

Now we introduce a new method for estimating the population mean $\overline{Y}$. This method involves the same sampling method as CBSS, but it estimates $\overline{Y}$ by an adjusted estimator, not by the sample mean itself. We consider only the case when n is an odd number ($n \geq 3$) and k is an even number, because the method is defined and has a practical meaning in this case.

Consider again the case of $N = 20$, $n = 5$ and $k = 4$. Either C'_2 or C'_3 is selected with respective probability 1/2. We notice that the sums of the numbers assigned to the units in C'_2 and C'_3 are 52 and 53, respectively. When a linear trend exists in the population, it would be desirable to remove such a difference. Our idea is to replace y_{18} or y_{19} by "$y_{18.5}$" according as C'_2 or C'_3 is selected. Here, of course, $y_{18.5}$ is an imaginary value which does not actually exist.

If C'_2 is selected, then we can "estimate" $y_{18.5}$ by use of y_{15} and y_{18}. By the extrapolation method, $y_{18.5}$ is estimated by $(1/6)(7y_{18} - y_{15})$. Therefore, by using this value in place of y_{18}, we can estimate $\overline{Y}$ by

$$\begin{aligned}\overline{y}'^{*}_2 &= \frac{1}{5}\{y_2 + y_7 + y_{10} + y_{15} + \frac{1}{6}(7y_{18} - y_{15})\} \\ &= \overline{y}'_2 + \frac{1}{30}(y_{18} - y_{15}).\end{aligned}$$

Note that $\overline{y}'^{*}_2$ can also be expressed as

$$\overline{y}'^{*}_2 = \overline{y}'_2 + \frac{1}{30}(y'_{25} - y'_{24}),$$

where 24 and 25 subscript to y' are two-dimensional.

Suppose now that the selected cluster is C'_3. Then we can estimate $y_{18.5}$ by using y_{14} and y_{19}. We now need to use the method of interpolation because 18.5 is between 14 and 19. Using the resultant value in place of y_{19}, we can estimate $\overline{Y}$ by

$$\begin{aligned}\overline{y}'^{*}_3 &= \frac{1}{5}\{y_3 + y_6 + y_{11} + y_{14} + \frac{1}{10}(y_{14} + 9y_{19})\} \\ &= \overline{y}'_3 - \frac{1}{50}(y_{19} - y_{14}) \\ &= \overline{y}'_3 - \frac{1}{50}(y'_{35} - y'_{34}).\end{aligned}$$

The above method can be generalized as follows. Either one of the two clusters $C'_{k/2}$ and $C'_{k/2+1}$ is selected with respective probability 1/2. The population mean $\overline{Y}$ is estimated by $\overline{y}'^*_{k/2}$ or $\overline{y}'^*_{k/2+1}$ according as the selected cluster is $C'_{k/2}$ or $C'_{k/2+1}$, where

$$\overline{y}'^*_{k/2} = \overline{y}'_{k/2} + \frac{1}{2n(k-1)}(y'_{k/2,n} - y'_{k/2,n-1}) \tag{1}$$

and

$$\overline{y}'^*_{k/2+1} = \overline{y}'_{k/2+1} - \frac{1}{2n(k+1)}(y'_{k/2+1,n} - y'_{k/2+1,n-1}). \tag{2}$$

Let us denote this method and the resultant estimator of $\overline{Y}$ by CBIE and $\overline{y}_{CBIE}$, respectively. Here "CBIE" represents centered balanced systematic sampling, interpolation and extrapolation. Then $\overline{y}_{CBIE}$ is biased for $\overline{Y}$ and it is clear that $\overline{y}_{CBIE}$ has bias

$$B(\overline{y}_{CBIE}) = \frac{1}{2}(\overline{y}'^*_{k/2} + \overline{y}'^*_{k/2+1}) - \overline{Y}$$

and mean square error

$$MSE(\overline{y}_{CBIE}) = \frac{1}{2}\{(\overline{y}'^*_{k/2} - \overline{Y})^2 + (\overline{y}'^*_{k/2+1} - \overline{Y})^2\}.$$

3. Expected mean square error of $\overline{y}_{CBIE}$

In this section, the expected mean square error of $\overline{y}_{CBIE}$ is obtained by using Cochran's (1946) infinite superpopulation model.

The finite population is regarded as a sample from an infinite superpopulation. First, as a general case, we set up the model as

$$y_i = \mu_i + e_i \ (i = 1, 2, \cdots, N), \tag{3}$$

where μ_i is a function of i and the random error e has properties $E(e_i) = 0$, $E(e_i^2) = \sigma^2$, $E(e_ie_j) = 0 \ (i \neq j)$. The operator E denotes the expectation over the infinite superpopulation.

From now on, with regard to μ and e also we will use the same style of notation as adopted for y. That is,

$$\overline{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mu_i,$$

$$\begin{aligned}\mu'_{ij} &= \mu_{i+(j-1)k}\ (j = 1, 3, 5, \cdots, n)\ (n\ :\ \text{odd}),\\ \overline{\mu}'_i &= \frac{1}{n}\sum_{j=1}^{n}\mu'_{ij},\\ \overline{\mu}'^{*}_{k/2} &= \overline{\mu}'_{k/2} + \frac{1}{2n(k-1)}(\mu'_{k/2,n} - \mu'_{k/2,n-1}),\end{aligned}$$

and so on.

The following theorem is very important in evaluating the efficiency of $\overline{y}_{CBIE}$. The proof of this theorem is omitted here, because the method of the proof is similar to that used in the proof of Theorem 1 in Kim (1999a).

Theorem 1. Assuming the model expressed as (3), the expected mean square error of $\overline{y}_{CBIE}$ for k even and n odd $(n \geq 3)$ is

$$E\{MSE(\overline{y}_{CBIE})\} = \frac{1}{2}\sum_{i=k/2}^{k/2+1}(\overline{\mu}'^{*}_i - \overline{\mu})^2 + \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2}{4n^2}\left\{\frac{1}{(k-1)^2} + \frac{1}{(k+1)^2}\right\}. \quad (4)$$

Now, let us consider the case when the population has a linear trend, namely, $\mu_i = a + bi$, where a and b are constants with $b \neq 0$. In other words, the assumed model is

$$y_i = a + bi + e_i\ (i = 1, 2, \cdots, N). \quad (5)$$

In this case, as a preparatory stage for obtaining $E\{MSE(\overline{y}_{CBIE})\}$ we get the following formulas :

$$\overline{\mu} = a + \left(\frac{b}{2}\right)(N+1) \quad (6)$$

$$\overline{\mu}'_i = a + \left(\frac{b}{2}\right)(N+1) + \left(\frac{b}{n}\right)\left(i - \frac{k+1}{2}\right)\ (i = 1, 2, \cdots, k) \quad (7)$$

$$\mu'_{in} = \mu_{i+(n-1)k} = a + b\{i + (n-1)k\}\ (i = 1, 2, \cdots, k) \quad (8)$$

$$\mu'_{i,n-1} = \mu_{1-i+(n-1)k} = a + b\{1 - i + (n-1)k\}\ (i = 1, 2, \cdots, k) \quad (9)$$

Thus we have

$$\begin{aligned}\overline{\mu}'^{*}_{k/2} &= \overline{\mu}'_{k/2} + \frac{1}{2n(k-1)}(\mu'_{k/2,n} - \mu'_{k/2,n-1})\\ &= a + \left(\frac{b}{2}\right)(N+1) \quad (10)\end{aligned}$$

and

$$\begin{aligned}\overline{\mu}'^{*}_{k/2+1} &= \overline{\mu}'_{k/2+1} - \frac{1}{2n(k+1)}(\mu'_{k/2+1,n} - \mu'_{k/2+1,n-1}) \\ &= a + \left(\frac{b}{2}\right)(N+1). \qquad (11)\end{aligned}$$

Now using (6), (10), (11) and the result of Theorem 1, we obtain the following theorem:

Theorem 2. For a population characterized by (5), the expected mean square error of $\overline{y}_{CBIE}$ is

$$E\{MSE(\overline{y}_{CBIE})\} = \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2}{4n^2}\left\{\frac{1}{(k-1)^2} + \frac{1}{(k+1)^2}\right\} (k: \text{even}, n: \text{odd}, n \geq 3). (12)$$

4. Comparison of efficiency with existing methods

In this section, the efficiency of $\overline{y}_{CBIE}$ is compared with that of estimators resulting from existing methods. First, let us consider SRS, OSS, MSS, BSS, CSS and CBSS. Bellhouse and Rao (1975) also discussed on comparisons of the performances of OSS, MSS, BSS and CSS.

For a population having a linear trend represented by (5), the following were obtained in Kim (1985) :

$$E\{MSE(\overline{y}_{SRS})\} = \left(\frac{b^2}{12}\right)(N+1)(k-1) + \frac{\sigma^2}{n}\frac{N-n}{N} \qquad (13)$$

$$E\{MSE(\overline{y}_{OSS})\} = \left(\frac{b^2}{12}\right)(k+1)(k-1) + \frac{\sigma^2}{n}\frac{N-n}{N} \qquad (14)$$

$$E\{MSE(\overline{y}_{MSS})\} = E\{MSE(\overline{y}_{BSS})\} = \left(\frac{b^2}{12n^2}\right)(k+1)(k-1) + \frac{\sigma^2}{n}\frac{N-n}{N} \quad (n: \text{odd}) \qquad (15)$$

$$E\{MSE(\overline{y}_{CSS})\} = \frac{b^2}{4} + \frac{\sigma^2}{n}\frac{N-n}{N} \; (k: \text{even}) \qquad (16)$$

$$E\{MSE(\overline{y}_{CBSS})\} = \frac{b^2}{4n^2} + \frac{\sigma^2}{n}\frac{N-n}{N} \quad (k : \text{even},\ n : \text{odd}) \tag{17}$$

Here $\overline{y}_{SRS}$, $\overline{y}_{OSS}$, $\overline{y}_{MSS}$, $\overline{y}_{BSS}$, $\overline{y}_{CSS}$ and $\overline{y}_{CBSS}$ denote the sample mean, which is used as the estimator of $\overline{Y}$, obtained from SRS, OSS, MSS, BSS, CSS and CBSS, respectively.

On the basis of formulas (12) through (17), we can arrange the methods under consideration according to the magnitude of the expected mean square error as the following theorem. For simplicity's sake, $E\{MSE(\overline{y}_{CBIE})\}$ is abbreviated as "$CBIE$", $E\{MSE(\overline{y}_{OSS})\}$ as "OSS", and so on. Thus, for example, "$CBIE < OSS$" means that CBIE is more efficient than OSS.

Theorem 3. Let A_k denote $(k-1)^{-2} + (k+1)^{-2}$. For a population having a linear trend represented by (5), the following hold :

(1) The case of $k = 2$ and $n = 3, 5, 7, \cdots$

(i) If $\sigma^2 < 9b^2/10$, then $CBIE < CBSS = MSS = BSS < CSS = OSS < SRS$.

(ii) If $9b^2/10 \le \sigma^2 < 9b^2n^2/10$, then $CBSS = MSS = BSS \le CBIE < CSS = OSS < SRS$.

(iii) If $9b^2n^2/10 \le \sigma^2 < 3b^2n^2(N+1)/10$, then $CBSS = MSS = BSS < CSS = OSS \le CBIE < SRS$.

(iv) If $3b^2n^2(N+1)/10 \le \sigma^2$, then $CBSS = MSS = BSS < CSS = OSS < SRS \le CBIE$.

(2) The case of $k = 4, 6, 8, \cdots$, $n = 3, 5, 7, \cdots$ and $n < \sqrt{(k^2-1)/3}$

(i) If $\sigma^2 < b^2/A_k$, then $CBIE < CBSS < CSS < MSS = BSS < OSS < SRS$.

(ii) If $b^2/A_k \le \sigma^2 < b^2n^2/A_k$, then $CBSS \le CBIE < CSS < MSS = BSS < OSS < SRS$.

(iii) If $b^2n^2/A_k \le \sigma^2 < b^2(k^2-1)/3A_k$, then $CBSS < CSS \le CBIE < MSS = BSS < OSS < SRS$.

(iv) If $b^2(k^2-1)/3A_k \le \sigma^2 < b^2n^2(k^2-1)/3A_k$, then $CBSS < CSS < MSS = BSS \le CBIE < OSS < SRS$.

(v) If $b^2n^2(k^2-1)/3A_k \le \sigma^2 < b^2n^2(N+1)(k-1)/3A_k$, then $CBSS < CSS < MSS = BSS < OSS \le CBIE < SRS$.

(vi) If $b^2n^2(N+1)(k-1)/3A_k \le \sigma^2$, then $CBSS < CSS < MSS = BSS < OSS < SRS \le CBIE$.

(3) The case of $k = 4, 6, 8, \cdots$, $n = 3, 5, 7, \cdots$ and $n = \sqrt{(k^2-1)/3}$ (for example, $k = 26$ and $n = 15$)

(i) If $\sigma^2 < b^2/A_k$, then $CBIE < CBSS < CSS = MSS = BSS < OSS < SRS$.

(ii) If $b^2/A_k \le \sigma^2 < b^2n^2/A_k$, then $CBSS \le CBIE < CSS = MSS = BSS < OSS < SRS$.

(iii) If $b^2n^2/A_k \le \sigma^2 < b^2n^2(k^2-1)/3A_k$, then $CBSS < CSS = MSS = BSS \le CBIE < OSS < SRS$.

(iv) If $b^2n^2(k^2-1)/3A_k \leq \sigma^2 < b^2n^2(N+1)(k-1)/3A_k$, then $CBSS < CSS = MSS = BSS < OSS \leq CBIE < SRS$.

(v) If $b^2n^2(N+1)(k-1)/3A_k \leq \sigma^2$, then $CBSS < CSS = MSS = BSS < OSS < SRS \leq CBIE$.

(4) The case of $k = 4, 6, 8, \cdots, n = 3, 5, 7, \cdots$ and $n > \sqrt{(k^2-1)/3}$

(i) If $\sigma^2 < b^2/A_k$, then $CBIE < CBSS < MSS = BSS < CSS < OSS < SRS$.

(ii) If $b^2/A_k \leq \sigma^2 < b^2(k^2-1)/3A_k$, then $CBSS \leq CBIE < MSS = BSS < CSS < OSS < SRS$.

(iii) If $b^2(k^2-1)/3A_k \leq \sigma^2 < b^2n^2/A_k$, then $CBSS < MSS = BSS \leq CBIE < CSS < OSS < SRS$.

(iv) If $b^2n^2/A_k \leq \sigma^2 < b^2n^2(k^2-1)/3A_k$, then $CBSS < MSS = BSS < CSS \leq CBIE < OSS < SRS$.

(v) If $b^2n^2(k^2-1)/3A_k \leq \sigma^2 < b^2n^2(N+1)(k-1)/3A_k$, then $CBSS < MSS = BSS < CSS < OSS \leq CBIE < SRS$.

(vi) If $b^2n^2(N+1)(k-1)/3A_k \leq \sigma^2$, then $CBSS < MSS = BSS < CSS < OSS < SRS \leq CBIE$.

Example. Suppose that we wish to draw a sample of size $n = 15$ from a population consisting of $N = 300$ units. We have $k = 300/15 = 20$. Assume that the slope of the linear trend is $b = 0.8$.

Then, by (4) of Theorem 3, the efficiency of the estimation methods can be compared as follows :

(i) If $\sigma^2 < 127.04$, then $CBIE < CBSS < MSS = BSS < CSS < OSS < SRS$.

(ii) If $127.04 \leq \sigma^2 < 16896.74$, then $CBSS \leq CBIE < MSS = BSS < CSS < OSS < SRS$.

(iii) If $16896.74 \leq \sigma^2 < 28584.72$, then $CBSS < MSS = BSS \leq CBIE < CSS < OSS < SRS$.

(iv) If $28584.72 \leq \sigma^2 < 3801767.37$, then $CBSS < MSS = BSS < CSS \leq CBIE < OSS < SRS$.

(v) If $3801767.37 \leq \sigma^2 < 54491998.96$, then $CBSS < MSS = BSS < CSS < OSS \leq CBIE < SRS$.

(vi) If $54491998.96 \leq \sigma^2$, then $CBSS < MSS = BSS < CSS < OSS < SRS \leq CBIE$.

We can see from this example that CBIE is relatively efficient as compared with other methods unless σ^2 is preposterously large.

Now let us compare CBIE with methods which estimate $\overline{Y}$ by a weighted mean, not by the simple mean, of the sample values. The methods and the expected mean square errors of the resultant estimators are as follows :

(1) End corrections (EC) (See Yates (1948).)

$$E\{MSE(\overline{y}_{EC})\} = \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2(k^2-1)}{6k^2(n-1)^2}.$$

(2) Modified systematic sampling with interpolation (MI) (See Kim (1998).)

$$E\{MSE(\overline{y}_{MI})\} = \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2}{12n^2}(4 - 12B_k + 6kC_k - \frac{1}{k^2}) \quad (k : \text{even}, n : \text{odd}, n \geq 3),$$

where

$$\begin{aligned}
B_k &= \sum_{i=1}^{k/2} \frac{1}{2k+1-2i} = \frac{1}{2}\left\{\psi\left(k+\frac{1}{2}\right) - \psi\left(\frac{k+1}{2}\right)\right\} \\
C_k &= \sum_{i=1}^{k/2} \frac{1}{(2k+1-2i)^2} = -\frac{1}{4}\left\{\psi^{(1)}\left(k+\frac{1}{2}\right) - \psi^{(1)}\left(\frac{k+1}{2}\right)\right\} \\
\psi(x) &= \frac{d}{dx}\ln\Gamma(x)\ (x>0) : \text{ polygamma function} \\
\Gamma(x) &= \int_0^\infty t^{x-1}e^{-t}dt\ (x>0) : \text{ gamma function} \\
\psi^{(1)}(x) &= \frac{d}{dx}\psi(x).
\end{aligned}$$

(3) Balanced systematic sampling with interpolation and extrapolation (BIE) (See Kim (1999a).)

$$E\{MSE(\overline{y}_{BIE})\} = \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2}{2n^2}(1 - \gamma - 2ln2 + D_k) \quad (k : \text{even},\ n : \text{odd},\ n \geq 3),$$

where $\gamma = 0.577215\cdots$ is the Euler constant, and $D_k = \frac{k}{8}\{\pi^2 - 2\psi^{(1)}(k+\frac{1}{2})\} - \psi(k+\frac{1}{2})$.

(4) Balanced systematic sampling with interpolation (BI) (See Kim (1999b).)

$$E\{MSE(\overline{y}_{BI})\} = \frac{\sigma^2}{n}\frac{N-n}{N} + \frac{\sigma^2}{2n^2}(1 - 4B_k + 2kC_k) \quad (k : \text{even},\ n : \text{odd},\ n \geq 5).$$

For various k, the values of the second terms of $E\{MSE(\cdot)\}$'s for EC, MI, BIE, BI and CBIE are given in Table 1. Note that the first terms are all the same for the five methods. Table 1 clearly shows that CBIE is the most efficient of the five methods, regardless of the value of σ^2.

5. Concluding remarks

In this paper, a new method was developed for estimating the mean of a population which has a linear trend, for the case of k even and n odd ($n \geq 3$). The proposed method, CBIE, consists of selecting a sample of size n by CBSS, and then estimating the population mean by using the concept of interpolation and extrapolation.

CBIE turned out to be relatively efficient as compared with existing methods if σ^2, the variance of the random error term in the infinite superpopulation model, is not preposterously large. Moreover, it was found to be more efficient than EC, MI, BIE and BI.

Table 1. The values of the second terms of $E\{MSE(\cdot)\}$'s for EC, MI, BIE, BI and CBIE

k	EC	MI	BIE	BI	CBIE
4	$0.1563\sigma^2/(n-1)^2$	$0.1061\sigma^2/n^2$	$1.1668\sigma^2/n^2$	$0.0559\sigma^2/n^2$	$0.0378\sigma^2/n^2$
8	$0.1641\sigma^2/(n-1)^2$	$0.1103\sigma^2/n^2$	$3.2882\sigma^2/n^2$	$0.0566\sigma^2/n^2$	$0.0082\sigma^2/n^2$
12	$0.1655\sigma^2/(n-1)^2$	$0.1111\sigma^2/n^2$	$5.5529\sigma^2/n^2$	$0.0567\sigma^2/n^2$	$0.0035\sigma^2/n^2$
16	$0.1660\sigma^2/(n-1)^2$	$0.1114\sigma^2/n^2$	$7.8765\sigma^2/n^2$	$0.0568\sigma^2/n^2$	$0.0020\sigma^2/n^2$
20	$0.1663\sigma^2/(n-1)^2$	$0.1115\sigma^2/n^2$	$10.2324\sigma^2/n^2$	$0.0568\sigma^2/n^2$	$0.0013\sigma^2/n^2$
∞	$0.1667\sigma^2/(n-1)^2$	$0.1117\sigma^2/n^2$	∞	$0.0569\sigma^2/n^2$	0

References

1. Bellhouse, D. R. and Rao, J. N. K. (1975). Systematic sampling in the presence of a trend, *Biometrika*, 62, 694-697.

2. Cochran, W. G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations, *Annals of Mathematical Statistics*, 17, 164-177.

3. Kim, H. J. (1985). New systematic sampling methods for populations with linear or parabolic trends, *Unpublished Master Thesis*, Department of Computer Science and Statistics, Seoul National University.

4. Kim, H. J. (1998). Estimation of population mean using interpolation in modified systematic sampling, *Korean Annals of Mathematics*, 15, 217-231.

5. Kim, H. J. (1999a). A study on estimating population mean by use of interpolation and extrapolation with balanced systematic sampling, *Journal of the Korean Data and Information Science Society*, 10, 91-102.

6. Kim, H. J. (1999b). Improved estimation of population mean using balanced systematic sampling and interpolation, *Proceedings of the Autumn Conference of the Korean Statistical Society,* 138-143.

7. Madow, W. G. (1953). On the theory of systematic sampling, III. Comparison of centered and random start systematic sampling, *Annals of Mathematical Statistics*, 24, 101-106.

8. Murthy, M. N. (1967). *Sampling Theory and Methods*, Statistical Publishing Society, Calcutta, India.

9. Sethi, V. K. (1965). On optimum pairing of units, *Sankhya*, B27, 315-320.

10. Singh, D. Jindal, K. K. and Garg, J. N. (1968). On modified systematic sampling, *Biometrika*, 55, 541-546.

11. Yates, F. (1948). Systematic sampling, *Philosophical Transactions of the Royal Society of London*, A241, 345-377.