Estimation of Population Mean Using Centered Modified Systematic Sampling and Interpolation

Hyuk Joo Kim[1]) and Byoung Chul Choi[2])

Abstract

A method is proposed for efficiently estimating the mean of a population which has a linear trend. The proposed estimator is based on the centered modified systematic sampling method and the concept of interpolation. Using the expected mean square error criterion, it is shown that the proposed method is more efficient than conventional methods in most real cases.

Keywords : Linear trend, Centered modified systematic sampling, Interpolation, Infinite superpopulation model.

1. 서 론

유한모집단의 평균을 추정하는 문제를 생각해 보자. 모집단의 단위들이 특성값에 관해 랜덤한 순서로 배열되어 있지 않고 어떤 추세를 가지고 있는 경우가 있다. 이 추세는 선형 또는 곡선형일 수도 있고 그밖의 형태일 수도 있다.

본 논문에서는 선형 추세를 갖는 모집단의 평균을 효율적으로 추정하기 위한 새로운 방법을 제시하고자 한다. 선형 추세는 3절에서 통계적 모형에 의하여 정의될 것이다. 제시되는 방법은 표본크기 n이 3 이상의 홀수이고 추출률의 역수인 k가 짝수인 경우에 사용할 수 있는 방법이다.

2. 새로운 모평균 추정 방법

크기가 $N=kn$ 인 모집단을 생각하자. 모집단을 구성하는 N 개의 단위에는 1 부터 N 까지의 번호가 붙어 있다고 하고, 이 N 개의 단위들을 $U_1, U_2, \cdots, U_N$ 으로 나타내기로 하자. 이 모집단으로부터 크기 n 인 표본을 뽑으려 한다.

---

1) Professor, Division of Mathematics and Informational Statistics, Wonkwang University, Iksan, Chonbuk, 570-749, Korea.
2) Professor, Division of Mathematics and Statistical Informatics, Chonbuk National University, Chonju, Chonbuk, 561-756, Korea.

집락 $S'_i\,(i=1,2,\cdots,k)$ 를 다음과 같이 정의하자.

n 이 짝수일 때

$S'_i = \{\, U_{i+(j-1)k} \;:\; j=1,2,\cdots,n/2 \,\} \bigcup \{\, U_{N+1-i-(j-1)k} : j=1,2,\cdots,n/2 \,\}$

n 이 홀수일 때

$S'_i = \{\, U_{i+(j-1)k} : j=1,2,\cdots,(n+1)/2 \,\} \bigcup \{\, U_{N+1-i-(j-1)k} : j=1,2,\cdots,(n-1)/2 \,\}$

예를 들어 N=20, n=5, k=4이면 $S'_1 = \{\, U_1, U_5, U_9, U_{16}, U_{20} \,\}$, $S'_2 = \{\, U_2, U_6, U_{10}, U_{15}, U_{19} \,\}$, $S'_3 = \{\, U_3, U_7, U_{11}, U_{14}, U_{18} \,\}$, $S'_4 = \{\, U_4, U_8, U_{12}, U_{13}, U_{17} \,\}$ 이다.

이제 Kim(1985)에 의하여 제시된 중심변형계통추출(centered modified systematic sampling: CMSS) 방법을 살펴보자. 이 방법은 Singh 등(1968)의 변형계통추출(modified systematic sampling: MSS)과 Madow(1953)의 중심계통추출(centered systematic sampling: CSS)을 결합한 것이다. k가 홀수이면 $S'_{(k+1)/2}$를 뽑는다. 따라서 이 경우 CMSS는 CSS와 같은 방법이다. k가 짝수(앞으로는 이 경우만 생각하기로 하자)이면 각각 1/2의 확률로 두 집락 $S'_{k/2}$와 $S'_{k/2+1}$ 중 하나를 뽑는다. 이 방법에 의한 모평균 $\overline{Y}$ 의 추정량 $\overline{y}_{cm}$ 은 다음과 같은 평균제곱오차를 갖는다.

$$MSE(\overline{y}_{cm}) = \frac{1}{2}\left\{ (\overline{y}'_{k/2} - \overline{Y})^2 + (\overline{y}'_{k/2+1} - \overline{Y})^2 \right\} \tag{2.1}$$

단 $\overline{y}'_i$ 는 S'_i 안의 단위들의 특성값의 평균이다($i=1,2,\cdots,k$).

이제 본 논문에서 사용될 기호들을 정의한다.

y_i : 모집단 안의 i 번째 단위 U_i 가 가지고 있는 특성값 ($i=1,2,\cdots,N$)

$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$: 추정하고자 하는 모집단 평균

y'_{ij} : S'_i안의 j 번째 단위의 특성값 ($i=1,2,\cdots,k$; $j=1,2,\cdots,n$)

즉, n 이 짝수일 때

$y'_{ij} = y_{i+(j-1)k} \quad (j=1,2,\cdots,n/2)$

$y'_{ij} = y_{N+1-i-(n-j)k} = y_{1-i+jk} \quad (j=n/2+1\ ,\ n/2+2,\ \cdots,n)$

n 이 홀수일 때

$y'_{ij} = y_{i+(j-1)k} \quad (\,j=1,2,\cdots,(n-1)/2,\,(n+1)/2\,)$

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

$\overline{y}'_i = \frac{1}{n}\sum_{j=1}^{n} y'_{ij}$: S'_i안의 단위들의 특성값의 평균 ($i=1,2,\cdots,k$)

이제 k가 짝수이고 n이 홀수($n \geq 3$)인 경우 모평균 $\overline{Y}$에 대한 새로운 추정 방법을 제시하고자 한다. 이 방법은 CMSS를 사용하여 표본을 추출한 뒤 보간법(i n t e r p o l a t i o n)을 이용하여 표본평균 $\overline{y}_{cm}$보다 수정된 추정량을 써서 $\overline{Y}$를 추정하는 것이다.

k가 짝수이고 n이 3 이상의 홀수인 경우 CMSS에 의하여 하나의 집락을 뽑는다. 즉, 각각 1/2의 확률로 두 집락 $S^{'}_{k/2}$와 $S^{'}_{k/2+1}$ 중 하나를 뽑는다. $\overline{Y}$ 의 추정값은 $S^{'}_{k/2}$가 뽑힌 경우

$$\overline{y}^{'*}_{k/2} = \overline{y}^{'}_{k/2} + \frac{1}{2n(k+1)}(y^{'}_{k/2,(n+3)/2} - y^{'}_{k/2,(n+1)/2}) \qquad (2.2)$$

로 하며, $S^{'}_{k/2+1}$이 뽑힌 경우에는

$$\overline{y}^{'*}_{k/2+1} = \overline{y}^{'}_{k/2+1} - \frac{1}{2nk}(y^{'}_{k/2+1,(n+1)/2} - y^{'}_{k/2+1,(n-1)/2}) \qquad (2.3)$$

로 한다.

이 방법을 CMI로 나타내고 CMI에 의한 $\overline{Y}$의 추정량을 $\overline{y}_{cmi}$ 로 표시하자. $\overline{y}_{cmi}$ 는 다음과 같은 편의와 평균제곱오차를 갖는다는 것을 쉽게 알 수 있다.

$$B(\overline{y}_{cmi}) = \frac{1}{2}(\overline{y}^{'*}_{k/2} + \overline{y}^{'*}_{k/2+1}) - \overline{Y} \qquad (2.4)$$

$$MSE(\overline{y}_{cmi}) = \frac{1}{2}\{(\overline{y}^{'*}_{k/2} - \overline{Y})^2 + (\overline{y}^{'*}_{k/2+1} - \overline{Y})^2\} \qquad (2.5)$$

3. 무한초모집단 모형 하에서의 평균제곱오차의 기대값

이 절에서는 Cochran(1946)이 소개한 무한초모집단 모형을 사용하여 $\overline{y}_{cmi}$ 의 평균제곱오차의 기대값을 구한다.

무한초모집단 모형이란, 주어진 유한모집단을 무한초모집단으로부터 뽑힌 하나의 표본으로 간주하는 것으로서 다음과 같이 세워진다.

$$y_i = \mu_i + e_i \qquad (i=1,2,\cdots,N) \qquad (3.1)$$

여기서 μ_i는 i의 함수이며, e_i는 오차항으로서 $E(e_i)=0$, $E(e_i^2)=\sigma^2$, $E(e_ie_j)=0$ ($i\neq j$일 때)이다. E는 무한초모집단에 걸친 기대값을 나타낸다.

이제부터 μ에 관해서도 y에 관해서 정의된 것과 같은 양식의 기호를 사용하기로 한다. 예컨대

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

등이다.

이러한 기호들과 위의 가정들, 그리고 식 (2.5)를 이용하면 다음과 같은 정리를 얻을 수 있다. 증명은 생략하기로 한다.

정리 1. 모형 (3.1)을 가정할 때, $\bar{y}_{cmi}$ 의 평균제곱오차의 기대값은 다음과 같다. 단, 여기서 $A=\sigma^2(k-1)/N$ 이며, k 는 짝수, n 은 3 이상의 홀수이다.

$$EMSE(\bar{y}_{cmi})=\frac{1}{2}\left\{(\bar{\mu}'^{*}_{k/2}-\bar{\mu})^2+(\bar{\mu}'^{*}_{k/2+1}-\bar{\mu})^2\right\}+A+\frac{\sigma^2}{4n^2}\left\{\frac{1}{k^2}+\frac{1}{(k+1)^2}\right\} \quad (3.2)$$

이제 $\mu_i=a+b\,i$ (a와 b는 상수, $b\neq0$)인 경우, 즉 가정된 무한초모집단 모형이

$$y_i=a+bi+e_i \quad (i=1,2,\cdots,N) \quad (3.3)$$

인 경우를 생각해 보자. 이 경우가 바로 모집단에 선형 추세가 존재하는 경우이다. 이 경우 정리1을 이용하면 다음과 같은 정리를 얻게 된다. 앞에서와 같이 $A=\sigma^2(k-1)/N$ 이다.

정리 2. 모집단이 식 (3.3)과 같은 선형 추세를 가질 때, $\bar{y}_{cmi}$ 의 기대평균제곱오차는 다음과 같다.

$$EMSE(\bar{y}_{cmi})=A+\frac{\sigma^2}{4n^2}\left\{\frac{1}{k^2}+\frac{1}{(k+1)^2}\right\}\ (k\text{: 짝수},\ n\text{: 3 이상의 홀수}) \quad (3.4)$$

4. 무한초모집단 모형 하에서의 효율성 비교

이제 본 논문에서 얻어진 추정량인 $\bar{y}_{cmi}$ 의 효율성을 다음과 같은 기존의 여러 방법들에 의한 추정량들과 비교해 보자.

(1) 보통의 계통추출(ordinary systematic sampling)

$$EMSE(\bar{y}_{sy})=\frac{b^2(k^2-1)}{12}+A \quad (4.1)$$

(2) 끝값수정법(end corrections) (Yates, 1948)

$$EMSE(\bar{y}_{ec})=A+\frac{\sigma^2(k^2-1)}{6k^2(n-1)^2} \quad (4.2)$$

(3) 변형계통추출(modified systematic sampling) (Singh et al., 1968)

$$EMSE(\bar{y}_{mod})=\begin{cases}A & (n\text{: 짝수})\\ \dfrac{b^2(k^2-1)}{12n^2}+A & (n\text{: 홀수})\end{cases} \quad (4.3)$$

(4) 균형계통추출(balanced systematic sampling) (Sethi, 1965; Murthy, 1967)

$$EMSE(\bar{y}_{bal})=\begin{cases}A & (n\text{: 짝수})\\ \dfrac{b^2(k^2-1)}{12n^2}+A & (n\text{: 홀수})\end{cases} \quad (4.4)$$

(5) 중심계통추출(centered systematic sampling) (Madow, 1953)

$$EMSE(\bar{y}_{cen}) = \begin{cases} \dfrac{b^2}{4} + A & (k\text{: 짝수}) \\ A & (k\text{: 홀수}) \end{cases} \tag{4.5}$$

(6) 중심변형계통추출(centered modified systematic sampling) (Kim, 1985)

$$EMSE(\bar{y}_{cen}) = \begin{cases} A & (k\text{: 짝수},\ n\text{: 짝수}) \\ \dfrac{b^2}{4n^2} + A & (k\text{: 짝수},\ n\text{: 홀수}) \end{cases} \tag{4.6}$$

(7) 중심균형계통추출(centered balanced systematic sampling) (Kim, 1985)

$$EMSE(\bar{y}_{cb}) = EMSE(\bar{y}_{cm}) \tag{4.7}$$

(8) 중심균형계통추출과 보간법을 이용한 방법(centered balanced systematic sampling with interpolation: CBI) (김혁주와 석은양, 2000)

$$EMSE(\bar{y}_{cbi}) = A + \frac{\sigma^2}{2n^2(k+1)^2} \quad (k\text{: 짝수},\ n\text{: 5 이상의 홀수}) \tag{4.8}$$

(9) 중심변형추출(centered modified sampling) (Fountain과 Pathak, 1989)

$$EMSE(\bar{y}_{cmFP}) = \begin{cases} A & (n\text{: 짝수}) \\ A & (k\text{: 홀수},\ n\text{: 홀수}) \\ \dfrac{b^2}{4n^2} + A & (k\text{: 짝수},\ n\text{: 홀수}) \end{cases} \tag{4.9}$$

(10) 중심균형추출(centered balanced sampling) (Fountain과 Pathak, 1989)

$$EMSE(\bar{y}_{cbFP}) = EMSE(\bar{y}_{cmFP}) \tag{4.10}$$

(11) 양끝추출(two-end sampling) (Fountain과 Pathak, 1989)

$$EMSE(\bar{y}_{tes}) = EMSE(\bar{y}_{cmFP}) = EMSE(\bar{y}_{cbFP}) \tag{4.11}$$

모집단에 선형 추세가 존재하는 경우 식(3.4)와 식(4.1)부터 식(4.11)까지를 사용하여 여러 방법들의 효율성을 비교할 수 있다. 본 논문에서 제시된 방법이 k가 짝수이고 표본크기 n이 3 이상의 홀수인 경우에 사용하기 위한 것이므로, 이 절에서 고려하는 경우는 모두 이러한 경우이다.

먼저 본 논문에서 제시된 방법인 CMI와 보통의 계통추출법을 비교해 보자. (3.4)와 (4.1)로부터, $\bar{y}_{cmi}$가 $\bar{y}_{sy}$ 보다 효율적일 조건, 즉

$$EMSE(\bar{y}_{cmi}) < EMSE(\bar{y}_{sy}) \tag{4.12}$$

일 필요충분조건은

$$\sigma^2 < \frac{b^2 n^2 (k^2-1)}{3\{1/k^2 + 1/(k+1)^2\}} \tag{4.13}$$

임을 얻게 된다. 뒤에 나올 예1에서 보여지겠지만, 이 조건 (4.13)은 오차항의 분산 σ^2(모집단의

분산이 아님)이 터무니없이 큰 값을 갖지 않는 한 성립하는 조건이다.

다음으로 CMI와 Yates(1948)의 끝값수정법을 비교해 보자. (3.4)와 (4.2)로부터

$$EMSE(\bar{y}_{ec}) - EMSE(\bar{y}_{cmi}) = \frac{\sigma^2}{12k^2(k+1)^2 n^2 (n-1)^2} \{ (2k^4+4k^3-6k^2-10k-5)n^2 + 3(2k^2+2k+1)(2n-1) \} \quad (4.14)$$

을 얻게 되는데, 이 값은 k가 짝수이고 n이 3 이상의 홀수이면 항상 0보다 큰 값이라는 것을 쉽게 알 수 있다. k에 2, 4, 6, …을 넣어서 생각해 보면 명백하다. 따라서 기대평균제곱오차의 관점에서 보았을 때 $\bar{y}_{cmi}$는 항상 $\bar{y}_{ec}$ 보다 효율적이라는 사실을 알 수 있다.

이번에는 CMI를 김혁주와 석은양(2000)의 CBI와 비교해 보자. (3.4)와 (4.8)로부터

$$EMSE(\bar{y}_{cbi}) - EMSE(\bar{y}_{cmi}) = \frac{\sigma^2}{4n^2}\left\{ \frac{1}{(k+1)^2} - \frac{1}{k^2} \right\} \quad (4.15)$$

이 된다. 이 값은 엄밀히 부호를 따지면 음수이긴 하나 대부분의 현실적인 경우에 0과 거의 같은 값이며(예를 들면 (a) $k=4$, $n=7$인 경우 $-1.15\times10^{-4}\sigma^2$, (b) $k=10$, $n=25$인 경우 $-6.94\times10^{-7}\sigma^2$, (c) $k=20$, $n=35$인 경우 $-4.74\times10^{-8}\sigma^2$), 더욱이 EMSE $(\bar{y}_{cmi})$와 EMSE $(\bar{y}_{cbi})$에 공통으로 포함되어 있는 항인 $A=\sigma^2(k-1)/N$의 값에 대한 비율을 생각해 보면 거의 완전히 무시할 수 있는 수준이다(절대값의 비율이 (a) $k=4$, $n=7$인 경우 1.07×10^{-3}, (b) $k=10$, $n=25$인 경우 1.93×10^{-5}, (c) $k=20$, $n=35$인 경우 1.75×10^{-6}). 따라서 CMI는 CBI와 거의 같은 정도의 효율성을 가지고 있다고 볼 수 있다.

이러한 방식으로 모든 경우에 대하여 CMI를 여러 방법들과 비교한 결과를 다음과 같이 정리할 수 있다. 간결한 표현을 위하여 $EMSE(\bar{y}_{sy})$, $EMSE(\bar{y}_{cmi})$ 등을 각각 sy, cmi 등으로 나타내기로 한다. 따라서, 예컨대 " $cmi < sy$ " 라는 표현은 $\bar{y}_{cmi}$가 $\bar{y}_{sy}$ 보다 효율적이라는 것을 의미한다. 그리고 k가 짝수이고 n이 홀수인 모든 경우에 $cm=cb=cmFP=cbFP=tes$이므로 간편성을 위하여 대표로 cm 한 가지만 표시하겠다. 여기서 $A_k = 1/\{1/k^2 + 1/(k+1)^2\}$ 이다.

(1) $k=2$ 이고 n 이 3 이상의 홀수인 경우

(i) $\sigma^2 < 36b^2/13$ 이면 $cmi < cm = mod = bal < cen = sy$

(ii) $36b^2/13 \le \sigma^2 < 36b^2n^2/13$ 이면 $cm = mod = bal \le cmi < cen = sy$

(iii) $36b^2n^2/13 \le \sigma^2$ 이면 $cm = mod = bal < cen = sy \le cmi$

(2) k 가 4 이상의 짝수, n 이 3 이상의 홀수이고 $n < \sqrt{(k^2-1)/3}$ 인 경우

(ⅰ) $\sigma^2 < b^2 A_k$ 이면 $cmi < cm < cen < mod = bal < sy$

(ⅱ) $b^2 A_k \le \sigma^2 < b^2 n^2 A_k$ 이면 $cm \le cmi < cen < mod = bal < sy$

(ⅲ) $b^2 n^2 A_k \le \sigma^2 < b^2 (k^2-1) A_k / 3$ 이면 $cm < cen \le cmi < mod = bal < sy$

(ⅳ) $b^2 (k^2-1) A_k / 3 \le \sigma^2 < b^2 n^2 (k^2-1) A_k / 3$ 이면 $cm < cen < mod = bal \le cmi < sy$

(ⅴ) $b^2 n^2 (k^2-1) A_k / 3 \le \sigma^2$ 이면 $cm < cen < mod = bal < sy \le cmi$

(3) k 가 4 이상의 짝수, n 이 3 이상의 홀수이고 $n = \sqrt{(k^2-1)/3}$ 인 경우
(예를 들면 $k=26$, $n=15$)

(ⅰ) $\sigma^2 < b^2 A_k$ 이면 $cmi < cm < cen = mod = bal < sy$

(ⅱ) $b^2 A_k \le \sigma^2 < b^2 n^2 A_k$ 이면 $cm \le cmi < cen = mod = bal < sy$

(ⅲ) $b^2 n^2 A_k \le \sigma^2 < b^2 n^2 (k^2-1) A_k / 3$ 이면 $cm < cen = mod = bal \le cmi < sy$

(ⅳ) $b^2 n^2 (k^2-1) A_k / 3 \le \sigma^2$ 이면 $cm < cen = mod = bal < sy \le cmi$

(4) k 가 4 이상의 짝수 , n 이 3 이상의 홀수이고, $n > \sqrt{(k^2-1)/3}$ 인 경우

(ⅰ) $\sigma^2 < b^2 A_k$ 이면 $cmi < cm < mod = bal < cen < sy$

(ⅱ) $b^2 A_k \le \sigma^2 < b^2 (k^2-1) A_k / 3$ 이면 $cm \le cmi < mod = bal < cen < sy$

(ⅲ) $b^2 (k^2-1) A_k / 3 \le \sigma^2 < b^2 n^2 A_k$ 이면 $cm < mod = bal \le cmi < cen < sy$

(ⅳ) $b^2 n^2 A_k \le \sigma^2 < b^2 n^2 (k^2-1) A_k / 3$ 이면 $cm < mod = bal < cen \le cmi < sy$

(ⅴ) $b^2 n^2 (k^2-1) A_k / 3 \le \sigma^2$ 이면 $cm < mod = bal < cen < sy \le cmi$

예. (모의실험을 이용한 설명)
모형

$$y_i = 4 + 0.6\, i + e_i \ (i = 1, 2, \cdots, 36) \tag{4.16}$$

를 설정하자(즉 $a=4$, $b=0.6$). 오차항 e_i 의 값들을 실제로 발생시킴으로써 무한초모집단으로부터 크기 $N=36$ 인 모집단을 생성한 다음, 다시 이 모집단으로부터 크기 $n=9$ 인 표본을 뽑아 모평균을 추정하는 문제를 생각해 보기로 하자. $k=4$ 이며, 오차항 e_i 는 3절에서와 같은 조건을 만족시킨다. e_i 의 분산 σ^2 의 값은 4로 하였고, e_i 의 분포의 형태는 정규분포로 정하였으며, 미니탭(MINITAB)의 RANDOM 명령문을 이용하여 e_i 의 값들을 발생시켰다. 생성된 모집단은 다음과 같다.

3.3756	5.0503	4.1700	7.6870	4.6081	5.6075	10.1837	7.3378
11.6498	7.4627	10.6502	14.0271	10.3177	13.6694	13.1529	15.5815
17.4959	12.3033	12.8413	15.8322	20.6432	15.2415	17.0501	15.7510
18.2948	16.2124	22.0999	20.2150	22.1852	22.9331	19.0335	22.3455
19.6729	22.7010	25.9393	26.8678				

이 모집단의 평균은 $\overline{Y}=14.7275$ 이며, 이 모집단은 대체적으로 증가하는 선형 추세를 가지고 있음을 볼 수 있다. 앞에 열거된 기존의 열한 가지 방법에 의한 $\overline{Y}$의 추정량들의 평균제곱오차는 아래와 같다.

$$MSE(\overline{y}_{sy})=1.0058,\quad MSE(\overline{y}_{ec})=0.3421,\quad MSE(\overline{y}_{mod})=0.3537,$$
$$MSE(\overline{y}_{bal})=0.5757,\quad MSE(\overline{y}_{cen})=0.8383,\quad MSE(\overline{y}_{cm})=0.2387,$$
$$MSE(\overline{y}_{cb})=0.5474,\quad MSE(\overline{y}_{cbi})=0.4962,\quad MSE(\overline{y}_{cmFP})=0.2852,$$
$$MSE(\overline{y}_{cbFP})=0.4263,\quad MSE(\overline{y}_{tes})=0.2522$$

한편 본 논문에서 제시된 방법인 CMI를 사용하면, 모평균 $\overline{Y}$를 다음과 같은 두 개의 값 중 하나로 추정하게 된다(확률은 각각 1/2).

$$\overline{y}'^{*}_{2}=14.2990$$
$$\overline{y}'^{*}_{3}=14.2361$$

따라서 CMI에 의한 $\overline{Y}$의 추정량 $\overline{y}_{cmi}$ 의 평균제곱오차는

$$MSE(\overline{y}_{cmi})=0.2126$$

이다. 이 값은 위의 열한 가지를 포함한 열두 가지의 방법에 의한 평균제곱오차 중 가장 작은 값이므로, CMI가 열두 가지의 방법 중 가장 효율적이라는 것을 보여 준다.

참 고 문 헌

[1] 김혁주, 석은양(2000), “선형추세를 갖는 모집단에 대한 효율적인 모평균 추정: 계통추출의 확장”, *응용통계연구*, 제13권 제2호, 457-476.

[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] 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, 2511-2526.

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

[5] 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.

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

[7] Sethi, V. K. (1965), "On optimum pairing of units," *Sankhya,* **B27,** 315-320.

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

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