

모의실험 분석중 구간평균기법의 개선을 위한 연구*

A Study on the Improvement of the Batch-Means Method in Simulation Analysis

천 영 수**
Young Soo Chun

Abstract

The purpose of this study is to make an improvement to the batch-means method, which is a procedure to construct a confidence interval (c.i.) for the steady-state process mean of a stationary simulation output process. In the batch-means method, the data in the output process are grouped into batches. The sequence of means of the data included in individual batches is called a batch-means process and can be treated as an independently and identically distributed set of variables if each batch includes sufficiently large number of observations. The traditional batch-means method, therefore, uses a batch size as large as possible in order to destroy the autocovariance remaining in the batch-means process. The c.i. procedure developed and empirically tested in this study uses a small batch size which can be well fitted by a simple ARMA model, and then utilizes the dependence structure in the fitted model to correct for bias in the variance estimator of the sample mean.

1. 서 론

시물레이션의 출력프로세스를 $\{X_i, i \geq 1\}$ 라고 하면 그의 정상상태 평균(steady-state process mean)은 다음과 같이 정의된다.

$$\mu = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i}{n}$$

이 정상상태 평균에 대해 추정하기 위하여 관찰된 데이터 사용할 때, 초기상태로 인한 영향을 제거하기 위하여 상당히 많은 수의 초기 관측치를 제외한다. 초기 관측

* 본 논문은 1993년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음

This paper was supported (in part) by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1993

** 제주대학교 경상대학 경영학과

치를 충분히 제거하고 얻은 데이터를 $\{X_1, X_2, \dots, X_n\}$ 이라 하자. 그러면 이 데이터의 표본평균 $\bar{X} = \sum X_i / n$ 은 정상상태 평균 μ 에 대한 불편 추정량(unbiased estimator)으로 간주될 수 있다.

μ 에 대한 신뢰구간을 구하려면 표본평균의 분산 $V(\bar{X})$ 에 대한 추정량인 분산추정량(variance estimator)을 사용한다. 문제의 데이터 $\{X_1, X_2, \dots, X_n\}$ 가 서로 독립적이고 동일한 확률분포를 갖는다면 다음의 식으로 정의된 분산추정량은 불편 추정량이 된다.

$$\hat{V}(\bar{X}) = \frac{1}{n(n-1)} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (1)$$

따라서 μ 에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음의 식으로 계산될 수 있다.

$$\bar{X} \pm t_{\alpha/2, f} \hat{S}(\bar{X}) \quad (2)$$

단, 여기에서 $\hat{S}(\bar{X}) = [\hat{V}(\bar{X})]^{1/2}$ 이고, $t_{\alpha/2, f}$ 는 자유도가 $f = n-1$ 인 t 분포에서 상위 $\alpha/2$ 임계값을 의미한다.

그러나, 일반적으로 시물레이션 출력프로세스에는 자동상관(Autocovariance)이 내재하기 때문에 식(1)로 정의된 분산추정량에는 편차가 존재한다. 이 편기 때문에 프로세스 평균에 대한 신뢰구간을 구하는 것은 어려운 과제가 되고 있다. 이 자동상관문제를 해결하기 위하여 여러가지 신뢰구간기법이 개발되었는데 그 중에서 가장 널리 쓰이고 있는 기법은 구간평균기법이라고 할 수 있다.

구간평균기법에서는 시물레이션 출력 데이터를 크기가 일정한 구간으로 나누어 구간평균을 구하고 구간평균의 연속(sequence)을 독립적인 확률변수의 집합으로 간주하여 이 구간평균 데이터로부터 신뢰구간을 구한다. 구간규모가 커짐에 따라 구간평균간의 상관관계가 점차 약해지기 때문에 충분히 많은 갯수의 관측치를 확보할 수 있다면 이 방법은 타당하다고 할 수 있다[2], [11]. 그리고 구간평균기법은 지금까지 개발된 모든 신뢰구간 기법중에서 가장 이해하기 쉽고 계산량이 적은 장점이 있다. 성과면에서도 다른 기법에 비하여 손색이 없음이 알려졌다[14]. 이

러한 장점들 때문에 구간평균기법은 시물레이션 결과분석에서 가장 보편적으로 사용되어 왔고, 새로운 기법이 개발될 때에는 그 기법을 평가하는 비교대상으로 채택되곤 하였다.

그러나 구간평균기법에도 결함이 있다. 표본의 크기를 사전에 임의로 정하였을 경우에는 신뢰구간이 실제로 프로세스평균을 포함하는 "포함확률(coverage)"이 의도된 수준(신뢰계수)보다 심각하게 낮을 수가 있다[14]. 반면에 신뢰구간의 실제 포함확률이 원하는 수준에 도달할 수 있도록 표본의 크기를 결정하는 경우에는 엄청나게 큰 표본이 필요하게 될 수 있다[12]. 구간평균을 구하는 과정 즉 데이터의 집계 과정은 정보의 손실을 초래한다. 구간규모가 커짐에 따라 구간평균의 프로세스가 i.i.d. 프로세스에 수렴하는 속도가 점차 느려지기 때문에 이 정보손실은 심각할 수 있다.

본 연구에서는 분산추정량에 잔류하는 편차를 보정하는 새로운 방법에 기초한 신뢰구간기법을 개발하고 그 성과를 실증적으로 분석하였다. 새 기법의 요점은 구간평균기법에서와 같이 구간평균 프로세스가 i.i.d. 프로세스로 표현될 수 있도록 구간규모를 가급적 크게 하는 대신, 구간평균 프로세스가 간단한 자동상관이동평균(ARMA)모델로 잘 대표될 수 있을 때 까지만 구간규모를 확대하는 것이다. 그리고는 추정된 ARMA 모수를 사용하여 표본평균의 분산에 대하여 추정한다. 구간평균 프로세스는 i.i.d. 프로세스에 충분히 접근하기 이전에 ARMA 프로세스로 표현될 수 있기 때문에, 새로운 기법은 구간평균기법보다 적은 구간규모를 사용하게 되고 따라서 데이터의 집계로 인한 정보의 손실¹⁾을 줄일 수 있다.

2. 기존 신뢰구간기법의 검토

시물레이션 출력 프로세스의 정상상태 평균에 대한 신뢰구간을 계산하는 기존의 신뢰구간기법들은 독립반복기법(Independent replications)[10], 구간평균기법(Batch means)[10], Regenerative method[3], [5], [6], 시계열모형(Time-series modeling)[14], [7, pp 247-262], [8], [15], 스펙트럼분

1) 구간규모가 커질 경우 표본평균의 분산을 추정하기 위하여 사용되는 구간평균치의 개수가 줄어들어 자유도가 감소하기 때문에 공식 (2)에 의하여 프로세스 평균에 대한 신뢰구간의 폭이 증가한다. 즉 모평균에 대한 추정이 덜 정밀하게 되는데, 이를 정보의 손실이라 할 수 있다.

석(Spectral analysis) [7, pp 262-267], Standard time series [16] 등과 같이 6가지 범주로 분류될 수 있다. L. Schruben [16]에 의하여 가장 늦게(1882년) 발표된 기법인 Standard time series 이외의 기법은 A. M. Law 와 D. Kelton [14]에 의하여 잘 요약되어 있다. 앞의 3가지 기법은 시뮬레이션 출력 프로세스에 존재하는 자동상관 문제를 우회하려고 하는 반면 뒤의 3가지 기법은 자동상관을 허용하고 그 대신 자동상관 구조를 파악하여 이로부터 표본분산을 추정하는 기법이다. 본 연구에서 개발하고 실증분석한 기법은 구간평균기법과 시계열 모형 특히 ARMA 표현법(ARMA Representation)의 결합에 바탕을 두고 있으므로 본 장에서는 구간평균기법과 T. J Schreiber 등[15]이 제안한 ARMA 표현법의 개요와 장단점을 검토한다.

구간평균기법에서는 표본 크기가 n 인 원래의 데이터를 크기가 m 인 k 개의 구간으로 나누는데 (단 $k = \lfloor n/m \rfloor$), 첫째 구간은 관측치 X_1, \dots, X_m 을 포함하고, 둘째 구간은 관측치 X_{m+1}, \dots, X_{2m} 을 포함하는 식으로 각 구간은 인접한 m 개의 관측치들을 포함하도록 한다. 이 때 m 을 구간규모라 하고, k 번째 구간에 포함된 관측치들의 평균 $\bar{X}_k = (1/m) \sum_{j=1}^m X_{(j-1)m+i}$ 를 구간평균이라고 한다. 그리고 구간평균의 연속(sequence)을 구간평균 프로세스라고 한다. 구간평균기법에서는 이 구간평균 프로세스를 마치 “독립적이고 동일 확률분포를 갖는(i.i.d.)” 확률변수의 집합으로 간주하여 공식(2)에 $\hat{S}(\bar{X}) = [1/k(k-1) \sum (\bar{X}_j - \bar{X})^2]^{1/2}$ 과 $f=k-1$ 을 대입하여 신뢰구간을 구한다.[13, p296]

ARMA 표현법의 분석과정은 6개 단계로 나누어지는데, 1단계에서는 시뮬레이션 출력 자료로부터 시차(time-lag) 13까지 자동상관계수를 추정한다 ($\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_{13}$ 계산). 2단계에서는 H. L. Gray 등이 제안한 D 통계량[8]을 사용하여 출력 프로세스에 적용시킬 ARMA모델의 AR 차수 p 와 MA 차수 q 를 결정한다(단 $1 \leq p \leq 3, 0 \leq q < 3$ 으로 제한). 3단계에는 앞 단계에서 차수가 결정된 ARMA모델의 AR 및 MA 계수를 maximum likelihood method에 의하여 추정한다. 4단계에는 Lyung and Box의 Q 통계량이 자유도가 $10-p-q$ 인 χ^2 분포를 갖는다는 특성을 활용하여 추정된 모델이 적합한지 검정한다.

단, $Q = n(n+2) \sum_{k=1}^{10} \frac{\hat{\rho}_k^2}{(n-k)}$ 이다.

5단계에서는 적합성 검정에서 합격한 ARMA 모델의 계

수와 분산 추정량 사이에 존재하는 다음의 극한 특성을 이용하여 표본평균의 분산을 추정한다.

$$\hat{V}(\bar{X}) = \hat{V}_e * c_n / n$$

단 \hat{V}_e 는 White noise의 분산에 대한 추정량이고 $c_n = (1 - \sum_{i=1}^q \hat{\theta}_i)^2 / (1 - \sum_{i=1}^p \hat{\phi}_i)^2$ 이다. 6단계에서는 분산추정량에 상응하는 자유도 $f = n/c_n - p - q - 1$ 를 계산하고, 공식 (2)에 의하여 신뢰구간을 구한다.

구간평균기법의 장점과 약점은 전 장에서 설명한 바 있다. ARMA 표현법은 구간평균기법과는 달리 출력 데이터의 수 n 이 적을 때에도 포함빈도가 크게 약화되지 않는 것으로 나타났다. 그러나 7개의 테스트 모델 중에서 6개가 ARMA 모델이었기 때문에 ARMA 표현법이 좋은 성과를 내었다는 의문을 피하기 어렵다. 또한 임의로 AR 차수와 MA 차수를 3 이하로 제한한 점도 적은 문제점으로 지적된다. 그리고 계산에 의해 추정된 ARMA모델이 적합성 검정에서 실패하였을 때의 대안이 제시되지 않았다.

3. 새로운 기법의 개발

모든 시계열 자료는 AR 및 MA차수만 적절하게 선정하면 ARMA프로세스로 표현될 수 있다. 그런데 ARMA프로세스에서 유도된 구간평균 프로세스 역시 ARMA프로세스이고[1], 구간규모가 커짐에 따라 AR 및 MA모수의 값은 0에 수렴하여 궁극적으로는 독립적인 프로세스가 된다. 본 연구에서는 저차 모수보다는 고차 모수가 0에 수렴하는 속도가 더 빠르기 때문에 구간규모가 커짐에 따라 저차 모수의 절대값이 상대적으로 커짐을 밝히고 이러한 현상을 “저차모수의 지배”라고 명칭하였다. 저차모수의 지배 때문에 적절한 구간규모에 상응하는 구간평균 프로세스는 차수가 낮은 ARMA프로세스로 잘 표현할 수 있고 이 추정된 프로세스의 모수로 분산추정량을 계산할 수 있으므로 구간규모를 선정하는 규칙만 설정할 수 있다면 신뢰구간을 계산하는 새로운 방법이 구성될 수 있다.

3.1 구간평균 프로세스의 AR 및 MA 모수

AR모수의 값이 $\phi_1, \phi_2, \dots, \phi_p$ 이고 MA모수의 값이 $\theta_1, \theta_2, \dots, \theta_q$ 인 ARMA(p,q) 프로세스에서 구간규모가

m 인 구간평균 프로세스를 유도하면, 유도된 프로세스는 ARMA(p, q) 프로세스가 된다. 단, q' 는 다음과 같이 결정된다 [1], [9].

$$q' = p - \lfloor (p - q) / m \rfloor \quad (3)$$

ARMA(p, q) 프로세스의 특성방정식

$$\alpha^p + \phi_1 \alpha^{p-1} + \phi_2 \alpha^{p-2} + \dots + \phi_{p-1} \alpha + \phi_p = 0 \quad (4)$$

의 근을 $\{a_1, a_2, \dots, a_p\}$ 라 하고 유도된 구간평균 프로세스의 AR모수들을 $\phi_1(m), \dots, \phi_p(m)$ 이라 하고 하자. 그러면 이 AR모수들의 값은 식 $\prod_{j=1}^p (1 - a_j^m x)$ 에서 x^1, \dots, x^p 의 계수와 같다. [1].

즉,

$$\phi_h(m) = (-1)^{h-1} \sum \delta(i_1, i_2, \dots, i_h) a_{i_1} a_{i_2} \dots a_{i_h}^m, \quad h=1, 2, \dots, p \quad (5)$$

이다. 단, $i_j \in \{1, 2, \dots, p\}$, $j=1, \dots, h$ 이고

$$\delta(i_1, i_2, \dots, i_h) = \begin{cases} 1 & i_1, i_2, \dots, i_h \text{ 가 모두 다를 경우} \\ 0 & \text{그 외의 경우} \end{cases}$$

이다.

그런데, ARMA(p, q) 프로세스의 안정성 조건으로 인하여 $|a_j| \leq 1$ ($1 \leq j \leq p$)이다. 그러므로 m 이 증가함에 따라 $|a_{i_1} a_{i_2} \dots a_{i_h}^m|$ 은 0에 수렴하고 따라서 $|\phi_h(m)|$ 의 값도 0에 수렴하게 된다.

3.2 1차 모수의 지배

또한 m 값이 커짐에 따라 구간평균 프로세스의 1차 AR모수인 $\phi_1(m)$ 의 절대값에 비하여 고차의 AR모수들인 $\phi_2(m), \dots, \phi_p(m)$ 의 절대값이 작아지는 경향이 있다(이를 1차 모수의 지배라고 하자). 왜냐하면 만일 a_j 들이 모두 1에 가까운 값을 갖지 않는다면 고차의 AR모수들은 저차의 AR모수보다 더 빨리 0에 수렴하기 때문이다.

1차모수의 지배를 증명하기 위하여, 잠정적으로 a_j ($1 \leq j \leq p$)들이 모두 실수이고 m 은 짝수라고 가정하자. 그리고, $\{a_1, a_2, \dots, a_p\}$ 를 절대값의 크기 순으로 정렬한 것을 $\{a_{(1)}, a_{(2)}, \dots, a_{(p)}\}$ 라고 하자 즉 $|a_{(1)}| \geq |a_{(2)}| \geq \dots \geq |a_{(p)}| \geq 0$. 여기에서 $|a_{(1)}| > 0$ 일 경우만 고려하면 된다. (만일

$|a_{(1)}| = 0$ 이라면, $a_1 = a_2 = \dots = a_p$ 이 되고 따라서 $\phi_1 = \phi_2 = \dots = \phi_p = 0$ 이 되기 때문이다. 이 경우에는 프로세스가 자동상관의 특성을 갖지 않기 때문에 자동상관에 대하여 추정할 필요가 없게 된다.) 그러면 공식 (5)에 의해 다음의 부등식이 성립한다.

$$\phi_1(m) = a_{(1)}^m + \dots + a_p^m \geq a_{(1)}^m > 0,$$

$$0 \leq |\phi_h(m)| \leq \left(\frac{p}{h} \right) (a_{(1)} a_{(2)} \dots a_{(h)})^m.$$

따라서 다음의 수렴이 성립한다.

$$\begin{aligned} 0 \leq |\phi_h(m) / \phi_1(m)| &\leq \left(\frac{p}{h} \right) (a_{(1)} a_{(2)} \dots a_{(p)})^m / a_{(1)}^m \\ &= \left(\frac{p}{h} \right) (a_{(2)} a_{(3)} \dots a_{(p)})^m \\ &\rightarrow 0 \text{ as } m \rightarrow \infty, \quad 2 \leq h \leq p \end{aligned}$$

왜냐하면 $\left(\frac{p}{h} \right)$ 는 상수이고 $0 \leq |a_{(2)} \dots a_{(p)}| < 1$ 이기 때문이다. 즉, $\phi_h(m)$ 은 m 이 증가함에 따라 고차의 AR모수들을 지배하게 되는 특성이 있다. 여기에서 수렴의 속도는 $|a_{(2)} \dots a_{(p)}|$ 에 의해서 결정되는데 $a_{(2)}, \dots, a_{(p)}$ 들이 모두 1에 가까운 값을 갖지 않는다면 수렴속도는 매우 빠를 것이다.

구간평균 프로세스의 AR모수 사이에 존재하는 지배관계는 MA모수간에도 존재할 것으로 추론된다. m 이 충분히 크다면, 공식(3)에 의해, $p \geq q$ 일 때에는 $q' = p$ 이 되고 $p < q$ 일 때에는 $q' = p+1$ 이 된다. 그러므로 만 $\phi_h(m)$ 의 지배가 이미 실현되었다면 이 m 값에 대하여 구간평균 프로세스는 ARMA(1, p) 혹은 ARMA(1, $p+1$) 모델로 표현될 수 있다. 이제, 보다 큰 구간규모 ms (m 의 s 배수, 단 s 는 충분히 큰 수)로 얻은 구간평균 프로세스의 MA차수를 q'' 이라 하고 q'' 에 대하여 고찰하자. 구간규모가 ms 인 구간평균 프로세스는 구간규모가 m 인 구간평균 프로세스 ARMA(1, p) 혹은 ARMA(1, $p+1$)에서 다시 구간규모가 s 인 구간평균을 취하여 얻은 프로세스로 간주할 수 있다. 그러므로 q'' 는, 식(3)에 의하여, $q'' = 1 - \lfloor (1-p) / s \rfloor$ 혹은 $q'' = 1 - \lfloor -p / s \rfloor$ 로 정해진다. 따라서 큰 s 값에 대응하는 q'' 값은 2를 넘지 못하게 된다.

〈표 1〉 구간평균 프로세스에서 1차모수 지배의 상대 빈도

(단위 : %)

m	1	2	4	8	16	32	64
$ \phi_r(m)/\phi_r(m) > 10$	5.0	12.5	22.5	37.5	64.0	83.5	90.5
$ \phi_r(m)/\phi_r(m) > 10$	6.5	33.0	63.0	91.0	98.0	100.0	100.0
$ \theta_r(m)/\theta_r(m) > 10$	-	20.5	33.0	46.5	72.0	86.5	94.5
$ \theta_r(m)/\theta_r(m) > 10$	-	-	80.5	93.0	99.5	100.0	100.0
$ \phi_r(m)/\phi_r(m) > 20$	2.5	5.5	12.5	26.5	46.5	72.5	87.5
$ \phi_r(m)/\phi_r(m) > 20$	3.0	21.0	51.0	83.0	92.5	99.0	100.0
$ \theta_r(m)/\theta_r(m) > 20$	-	7.5	18.0	30.5	55.5	76.5	88.0
$ \theta_r(m)/\theta_r(m) > 20$	-	-	64.5	91.0	97.5	99.5	99.5

주: 원 프로세스: 모수 값이 무작위로 결정된 안정적 AR(3) 프로세스
실험횟수: 200회

상대빈도: 실험된 200개의 프로세스 중에서 $|\phi_r(m)/\phi_r(m)|$ 가 10 혹은 20보다 큰 프로세스의 비율

그러나, 식(4)의 근인 a_j 가 모두 실수가 아니고 일부가 복소수값을 가질 수 있다. 그러한 경우에는 m 값이 증가함에 따라 $\phi_r(m)$ 의 값이 진동하므로 특정 m 값에서 0에 가까운 값을 갖게 되고 따라서 이 m 값에서 $\phi_r(m)/\phi_r(m)$ 비율이 확대될 수 있다. 즉, 식(4)의 근중 일부가 허수 부분을 가질 경우 $\phi_r(m)/\phi_r(m)$ 비율은 0에 수렴하지 않을 수 있다.

이러한 수렴의 실패가 얼마나 빈번히 발생하는지 확인하기 위하여 안정성 조건 $|\phi_h| < \begin{bmatrix} p \\ h \end{bmatrix}$ 을 만족시키는 AR(3) 모델을 무작위적으로 200개 선정하여 실험을 하였다. 선정된 모델 AR(3)에 대하여 구간규모가 m 인 구간평균 프로세스의 AR 및 MA 모수 $\phi_r(m), \phi_r(m), \phi_r(m), \theta_r(m), \theta_r(m)$.

$\theta_r(m)$ 값을 계산하였다[3],[16]. 실험의 결과는 〈표 1〉에 제시된 바와 같은데, 이 표는 몇 개의 m 값에 대하여 1차의 AR 및 MA모수의 값을 고차의 모수의 값으로 나눈 비율의 절대값이 일정한 크기(10혹은 20)를 초과하는 상대 빈도를 보여주고 있다. 이 표에서 볼 수 있는 바와 같이 구간규모 m 이 증가함에 따라 저차의 AR 및 MA모수가 고차의 모수를 지배하는 경향이 있음을 알 수 있다.

종합하면, 구간규모가 충분히 클 때 구간평균 프로세스는 단순한 ARMA모델(AR 및 MA 차수가 낮은 모델)로 잘

대표될 수 있다고 할 수 있다.

3.3 구간규모선정 규칙

구간평균 프로세스가 특정 ARMA 모델로 대표될 수 있도록 하는 구간규모를 선정하는 것은 어려운 일이 아니다. 구간규모가 m 인 구간평균 프로세스에 대하여 ARMA 모델을 적용하여 계수를 추정한 후 추정된 모델이 적합성 검정에서 합격하면 m 을 적정 구간규모로 간주할 수 있다. 구간규모 m 의 값으로는 $n/2$ 이하의 모든 자연수를 고려할 수 있다. 그러나, 계수의 추정과 검정에 소요되는 계산의 부담을 고려하여 다음의 집합에 포함된 r 개의 구간규모만을 검토한다고 하자.

$$S_r = \{m_1, m_2, \dots, m_r\}$$

이 집합 S_r 를 검토집합이라 하고, $1 \leq m_1 < m_2 < \dots < m_r \leq n/2$ 이라고 하자. (만일 출력자료가 많지 않다면 $m_1=1$ 로 정할 수 있으나, 출력자료가 많으면 충분히 큰 m 값에서부터 시작함으로써 구간의 갯수 $k = \lfloor n/m_i \rfloor$ 를 줄여 시계열분석을 위한 계산의 부담을 줄일 수 있다.) 우선 구간규모가 m_1 인 구간평균 프로세스에 ARMA모델을 적용하여 계수의 값을 추정한다. 그리고 추정된 ARMA 모델이 적합성 검정에 합격하면 m_1 을 적정 구간규모로 선정하고, 합격하지 못하면 보다 큰 구간규모인 m_2 에 대한 구간평균 프로세스를 만들어 모델추정과 적합성검정의 과정을 반복한다²⁾ 이 반복과정의 결과로 한 구간규모가 선정되거나 아니면 검토집합내의 모든 값이 검토된 후에도 적정 구간규모가 선정되지 않을 수 있다. 이 적합성검정에서 사용된 유의수준이 $\alpha = 0.10$ 일 경우에 이 구간규모선정규칙(이하 'BSS규칙'이라 함)을 '1st. $\alpha 10$ '규칙이라 하자. 이 규칙의 명칭이 의미하는 바는 유의수준이 $\alpha = 0.10$ 인 적합성검정에서 첫째로(1st) 합격한 (즉, 가장 작은) 구간규모를 선정한다는 것이다³⁾.

유사한 방식으로 1st. $\alpha 30$, 1st. $\alpha 05$, 혹은 1st. $\alpha 01$ 규칙

2) 본 연구에서는 $m_{k+1} = 2 * m_k$, ($k=1, 2, \dots, r-1$) 로 설정하여 점차로 큰 구간규모의 구간평균 프로세스를 도출하는 작업을 단순화하였다.
3) 이와 같이 적정규모를 찾아내는 데 적은 구간규모부터 시작하는 이유는 가급적 적은 구간규모를 선정함으로써 과도한 데이터 집계(Data aggregation)를 피하기 위함이다.

들을 정의할 수 있다. 즉 추정된 ARMA 모델이 유의수준을 각각 $\alpha = 0.30, 0.05, 0.01$ 로 설정한 적합성검정에서 처음으로 합격한 구간규모를 선정할 수 있다. 같은 맥락에서 1st. $\alpha 00$ 규칙은 테스트 집합에서 제일 작은 값인 m_1 을 구간규모로 선정하는 규칙을 의미한다. (어떠한 모델이라도 유의수준이 0인 적합성검정에 합격할 것임) 이 1st. $\alpha 00$ 규칙의 성과와 다른 구간규모선정규칙의 성과를 비교함으로써, 적합성검정을 통하여 구간규모를 선정하는 방법이 과연 유의한지 평가할 수 있을 것이다.

검토집합내에 있는 값 m_1, m_2, \dots, m_i 중에서 $m_{e_1}, m_{e_2}, \dots, m_{e_i}$ 에 해당하는 구간평균 프로세스에 적용하여 추정된 모델들이 적합성검정에 합격하였다고 하자. 이 때 다음의 집합을 후보집합이라고 정의하자.

$$S(\alpha) = \{m_{e_1}, m_{e_2}, \dots, m_{e_i}\}$$

여기에서 유의할 점은 검토집합과 추정되는 모델이 주어지면 후보집합은 데이터와 유의수준에 따라 결정된다는 점이다. 여기에서 우리는 예컨대 $\alpha = 0.10$ 에 상응하는 후보집합에 포함된 구간규모들 중에서 적합성검정시 사용되는 통계량인 Q 통계량[8]의 값이 가장 적은 구간규모를 선정하는 규칙을 생각할 수 있다. 이 규칙을 min. $\alpha 10$ 규칙이라 정하자. 또한 유의수준 $\alpha = 0.30, 0.05, 0.01$ 에 상응하는 후보집합에서 검정통계량의 값이 가장 적은 구간규모를 선정하는 규칙이라는 의미로 min. $\alpha 30, \text{min. } \alpha 05, \text{ 및 min. } \alpha 01$ 규칙을 정할 수 있다.

3.4 분산추정량

일단 특정 구간평균 프로세스에 대하여 추정된 ARMA 모델이 적합성검정에서 받아들여지면, 추정된 모수로부터 분산추정량이 계산될 수 있다. ARMA모수와 $V(\bar{X})$ 사이에 성립하는 극한특성을 이용하여, 표본분산은 다음과 같은 식으로 추정될 수 있다. [15]

$$\hat{V}_w = \frac{\hat{V}_e(m^*) (1 - \sum \hat{\theta}_i(m^*))^2}{k^* (1 - \sum \hat{\phi}_i(m^*))^2} \quad (6)$$

단 여기에서 m^* 는 선정된 구간규모를 의미한다. $\hat{V}_e(m^*)$ 는 White noise의 분산에 대한 추정량이고, $\hat{\theta}_i(m^*)$ 와 $\hat{\phi}_i(m^*)$ 는 MA모수와 AR모수에 대한 추정량을 의미한다.

그리고 $k^* = \lfloor n / m^* \rfloor$ 는 구간규모가 m^* 일 때 구간평균의 개수를 나타낸다.

3.5 구간규모선정기법의 구성

지금까지 제시한 이론들을 종합하면 하나의 신뢰구간기법을 완성할 수 있다. 그 기법을 구간규모선정기법(Batch-size Selection Method; BSS기법)이라 하고 내용을 정리하자. 위에서 언급된 바와 같이 안정적 확률 프로세스로부터 도출된 구간평균 프로세스는 구간규모가 충분히 커지면 간단한 ARMA 모델로 잘 대표될 수 있다. 간단한 ARMA 모델로는 ARMA(1,1), AR(1), AR(2) 모델등을 고려할 수 있다. 특정 ARMA 모델로 잘 대표될 수 있는 구간규모 m^* 가 일단 선정되면 식(6)에 의하여 분산의 추정량 $\hat{V}_w(m^*)$ 이 구해진다. 이 때, 유효 자유도 f^* 는 다음의 식으로 정할 수 있다 [15].

$$f^* = \max\{1, \min\{k/b(m^*) - p - q - 1, k - p - q - 1\}\}$$

단, 여기에서

$$b(m^*) = \hat{V}_w(m^*) / \hat{V}_B(m^*), \\ \hat{V}_B(m^*) = [1 / k^*(k^* - 1)] \sum_i (X_i(m^*) - \bar{X})^2.$$

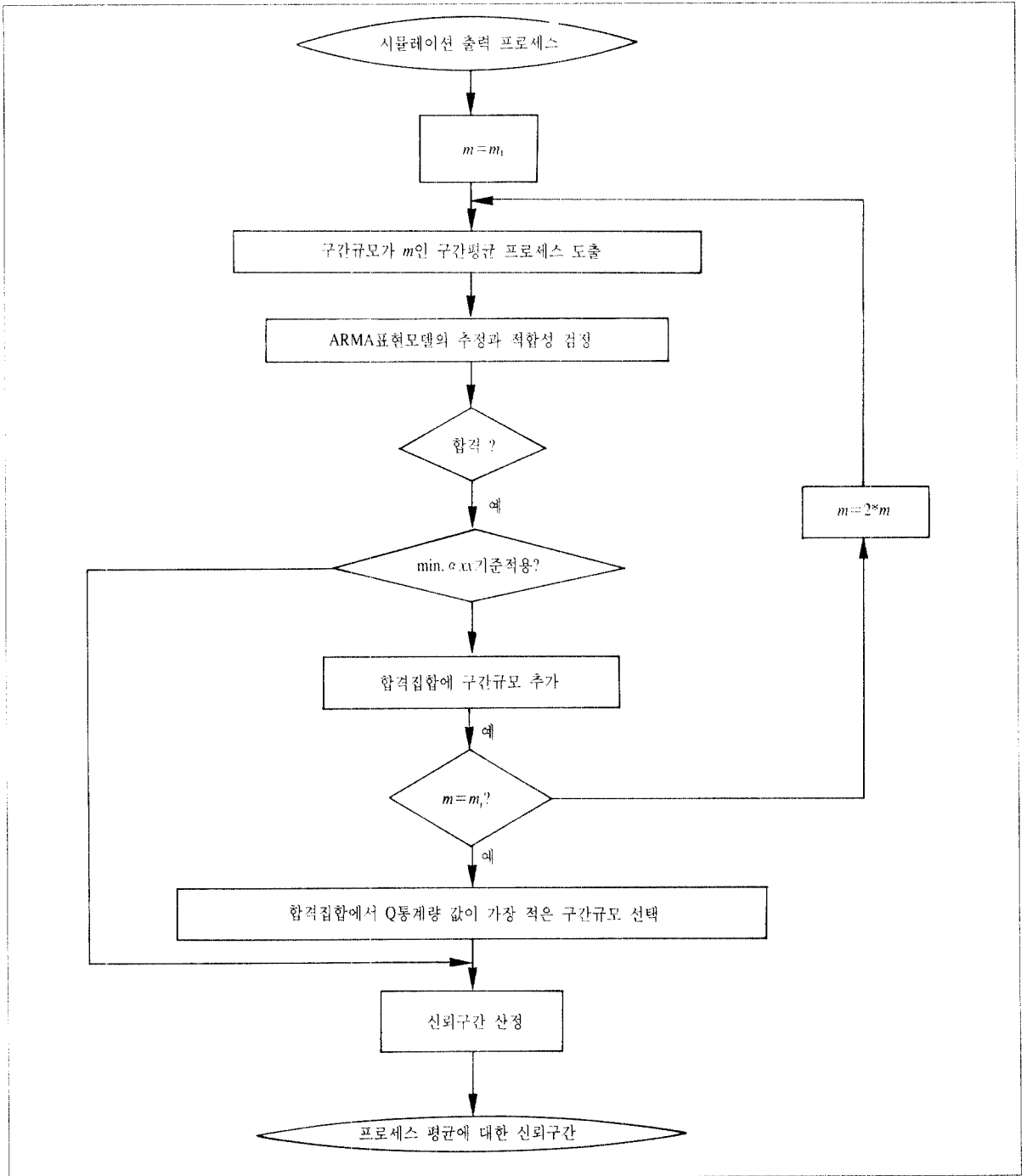
이제 식(2)에 $\hat{S}(\bar{X}) = \sqrt{\hat{V}_w(m^*)}$ 과 $f = f^*$ 을 대입함으로써 신뢰구간을 구하는 공식이 완성된다. 이상에서 설명한 구간규모선정기법을 흐름도로 표현하면 <그림 1>에 보인 바와 같다.

4. 실증분석

위에서 개발된 BSS기법이 실제 응용 상황에서 기대된 성과를 실현하는지 검토하기 위하여 다음과 같은 틀에서 실증분석을 하였다.

4.1 실험목적

민감도 분석; 우선 BSS기법의 성과가 구간평균 프로세스를 대표하는 ARMA 모델의 차수나 BSS규칙에 따라 민감하게 변하는지 검토하는 것을 목적으로 하였다. 만일



〈그림 1〉 구간규모선택기법의 흐름도

BSS기법이 이러한 모수들에 대하여 민감할 경우에는 최적의 모수값을 선정하기로 하였다.

BSS기법의 성과평가: 위의 검토과정에서 선정된 모수를 사용하여 신뢰구간을 구할 경우에 BSS기법이 좋은 성과를 내는지 평가하였다. 이 평가를 위하여 BSS기법이 기존의 BM기법에 비하여 향상된 성과를 내는지 검토하기로 하였다.

평가기준: 신뢰구간기법의 성과를 평가하는 기준으로는 “관측된 포함빈도”, “신뢰구간폭의 평균”, “신뢰구간폭의 분산”을 선정하였다.

4.2 실험설계

민감도 분석을 위해서 24종의 처리가 40종의 표본상황에서 산출하는 성과를 측정하였다. 24종의 처리는 3종의 자료표현모델, 8종의 BSS규칙을 결합한 것이고, 40종의 표본상황은 4수준의 표본규모와 10종의 테스트모델을 결합하여 얻었다. BSS기법의 성과평가를 위해서는, BSS기법과 BM기법이 4종류의 표본규모, 15종의 테스트모델로 구성된 60종의 표본상황에서 보여 주는 성과를 비교하였다. 각 표본상황에서의 성과를 관측하기 위하여 실험을 200회 반복하였다. 시뮬레이션을 위한 균등분포 난수발생기로는 $Z_i = 7^i Z_{i-1} \pmod{(2^{31} - 1)}$ 공식에 초기숫자 $Z_1 = 1,234,567$ 을 사용하였다. 실험 설계요인별 수준을 정리하면 다음과 같다.

1) 자료표현 ARMA모델: AR(1), AR(2), ARMA(1,1)

2) BSS규칙:

1st. α 30, 1st. α 05, 1st. α 01, 1st. α 00,
min. α 30, min. α 05, min. α 01, min. α 00

3) 표본규모: $n=512, 1,024, 2,048, 8,196$

4) 테스트모델:

프로세스 1. AR(1), $\phi = 0.9$

프로세스 2. AR(2), $\phi_1 = 1.3435, \phi_2 = -0.9225$

프로세스 3. M/M/1 FIFO, $\tau = 0.8$

프로세스 4. M/M/1 LIFO, $\tau = 0.8$

프로세스 5. M/M/1 SIRO(Serve-In-Random-Order Policy),
 $\tau = 0.8$

프로세스 6. 시분할 컴퓨터 시스템 ([11]에 소개된 모델)

프로세스 7. ARMA(2,1), $\phi_1 = 0.9, \phi_2 = -0.18, \theta = -0.9$

프로세스 8. M/M/1/M/1/M/1 FIFO, $\tau = 0.9, 0.7, 0.9$

프로세스 9. E4/M/1 FIFO, $\tau = 0.8$

프로세스 10. M/H2/1 FIFO, $\tau = 0.8$ (H2: Hyperexponential distribution)

프로세스 11. AR(1), $\phi = 0.5$

프로세스 12. AR(1), $\phi = 0.8$

프로세스 13. AR(2), $\phi_1 = 0.5, \phi_2 = -0.7$

프로세스 14. ARMA(1,1), $\phi = 0.7, \theta = -0.4$

프로세스 15. ARMA(2,1), $\phi_1 = 1.32, \phi_2 = -0.687, \theta = -0.4$

이 프로세스들 중에서 예컨대 프로세스 1은 자동상관계수가 큰 모델로서 선정되고 프로세스 2는 자동상관계수가 시차(time-lag)에 따라 양의 값과 음의 값 사이에서 크게 진동하는 사례로서 선정되었다. 본 연구는 이와 같이 다양한 ARMA 모델과 Queueing 모델 그리고 시분할 컴퓨터 시스템을 포함함으로써 기존의 연구사례에서보다 많은 프로세스를 테스트 모델로 사용하였다.⁴⁾

4.3 민감도분석 결과

1) 포함빈도의 분산분석

포함빈도의 분산분석에서는 3종의 자료표현 모델과 7종의 BSS규칙을 조합한 24종의 처리에 대하여 실험하였으나 BSS규칙중에서 1st. α 00 규칙의 성과가 다른 7종의 규칙에 비하여 현저하게 열등하다는 점이 분석단계에서 결론지어졌기 때문에 이 규칙을 제외한 분산분석을 다시 하였다. 즉 3종의 자료표현 모델과 7종의 BSS규칙으로 구성된 21종의 처리가 40종의 표본상황하에서 산출한 결과를 분석하였다. <표 2>는 제시된 요인들이 포함빈도의 분산에 미치는 주효과(main effect)와 상호작용효과(interaction effect)를 제시하고 있다.

이 표에서 볼 수 있는 바와 같이 표본 규모와 테스트 모델에 따라 포함빈도의 평균이 다르게 나타났는데 이 점은 기존의 모든 신뢰구간기법에서 목격된 현상과 동일하

4) 기존 연구사례에서 사용한 테스트 모델의 수: [2] 9개, [10] 1개, [11] 14개, [12] 10개, [14] 2개, [15] 7개, [16] 7개.

〈표 2〉 포함빈도의 분산분석표

분산의 원천	자유도	SS	MS	F통계값
N (표본규모)	3	1.70541	0.56847	293.0 †
P (테스트모델)	9	3.59069	0.44884	231.4 †
F (자료표현모델)	2	0.00601	0.00301	1.55
R (BSS규칙)	6	0.01032	0.00147	0.76
NF	6	0.01775	0.00296	1.53
FR	12	0.00225	0.00019	0.10
(Regression)	(38)	(5.33246)	(0.14033)	(72.3)
오차	801	1.55215	0.00194	
총계	839	6.88460		

주) SS: Sum of squares, MS: Mean of Squares, F = (각 항목의 MS) / (오차항의 MS)

관측횟수: 840회(40종의 표본상황 X 21종 처리)

† 유의수준 1% 이하

다. 그러나, 1st. $\alpha 00$ 규칙을 제외했을 경우에 BSS규칙은 포함빈도의 분산을 설명하는 요인이 아니다. 즉, 포함빈도의 평균이 BSS규칙에 따라 다르다는 주장은 통계적으로 유의하지 않다. 자료표현모델의 주효과, 자료표현모델과 BSS규칙간의 상호작용, 그리고 자료표현모델과 표본규모간의 상호작용도 통계적으로 유의하지 않은 것으로 나타났다. 이 점은 구간규모 선정기법의 강점이라고 할 수 있다. 왜냐하면 우리는 BSS규칙과 자료표현모델을 적절히 선정해야 하는 부담으로 부터 자유롭기 때문이다.

2) 신뢰구간폭의 분산분석

표본규모가 커질수록 신뢰구간의 폭이 축소된다. 그리고 테스트모델에 따라 정상상태 확률분포가 다르기 때문에 평균에 대한 신뢰구간의 폭이 테스트모델에 따라 달라지는 것은 당연한 일이다. 그러므로 신뢰구간폭의 분산을 분석할 때, 이 두 요인이 분산에 미치는 영향을 제거하는 것이 바람직하다. 이 목적을 위하여, 주어진 표본규모와 테스트모델의 조합내에서 21종의 처리(3종의 ARMA모델 X 7종의 BSS규칙)가 산출한 신뢰구간폭을 그들의 평균으로 나눈 비율(신뢰구간폭 비율)에 대한 분산을 분석하였다. 〈표 3〉은 이 비율값이 갖는 분산의 원인을 요인별로 보여주고 있다.

〈표 3〉에서 볼 수 있는 바와 같이 자료표현모델의 주효과(F)와 자료표현모델과 표본규모와의 상호작용효과(NF)는 중요한 것으로 판명되었지만, 부분 F통계량의 값 0.98

과 1.36로 보아 BSS규칙의 주효과(R)와 BSS규칙과 자료표현모델간의 상호작용효과(FR)는 통계적으로 의미가 없는 것으로 나타났다.

요약하자면, BSS규칙은 포함빈도나 구간폭에 중요한 영향을 미치지 않는다고 결론지을 수 있다. 또한 BSS규칙과 자료표현모델간의 상호작용효과가 있는 것 같지 않다.

〈표 3〉 신뢰구간폭 비율의 분산분석표

분산의 원천	자유도	SS	MS	F통계값
F (자료표현모델)	2	0.80424	0.40212	11.8 †
R (BSS규칙)	7	0.23473	0.03353	0.98
NF	6	0.60411	0.10069	2.95*
FR	14	0.65256	0.04661	1.36
(Regression)	(29)	(2.29561)	(0.07916)	(2.32)
오차	930	31.78858	0.03418	
총계	959	34.08419		

† 유의수준 1% 이하, * 유의수준 5% 이하

4.4 BSS기법의 평가

BSS규칙과 자료표현모델에 관하여 분석하는 과정에서 우리는 많은 표본상황하에서 BSS기법이 만족할 만한 포함빈도를 보여주는 것을 볼 수 있었다. 그러나 그렇지 못한 경우도 없지 않았다. BSS기법은 기존의 신뢰구간기법들에게 공통적인 문제 - 표본규모가 사전에 주어졌을 경우 신뢰구간의 실제 포함확률이 의도된 수준보다 현저하게 낮을 수도 있는 문제-의 완전한 해결책이 아님은 분명하다. 그러므로 BSS기법이 포함확률과 구간폭의 면에서 어느 정도의 개선을 이루었는지 점검하여 보았다.

개선여부를 판단하기 위하여 BM기법을 비교의 대상으로 선택하였다. BM기법의 성과는 구간규모에 매우 민감한 데, 종합적인 고려 끝에, $k=4$ 일 때의 성과로 대표하기로 하였다. 그리고 BSS기법의 성과는 BSS규칙으로 1st. $\alpha 01$ 규칙을 사용하고 자료표현모델로는 ARMA(1,1) 모델을 사용하였을 때의 성과로 대표하였다. 왜냐하면, 표본이 클 때에는 AR(1) 모델이 보다 우수한 성과를 보이고 표본이 적을 때에는 ARMA(p,q) 모델이 보다 우수한 경우도 있었지만 전반적으로 ARMA(1,1) 모델이 우수한 성과를 보였기 때문이다. 이러한 점에서 ARMA(1,1) 모델이 모든 표본 상황에서 BSS기법의 최선의 성과를 나타내는 것이 아님에 유의할 필요가 있다. 다음의 〈표 4, 5, 6〉은 조사

된 모든 표본상화에서 BSS기법과 BM기법의 성과를 대조해 보이고 있다.

1) 관측된 포함빈도의 비교: <표 4>에 의하면 포함빈도의 관점에서 BSS기법은 BM기법보다 우수하다고 할 수 없다. 오히려 몇 가지 표본상황에서 BSS기법은 BM기법보다 떨어지는 포함빈도를 보여주었다. 그러나 BSS기법은 표본이 커짐에 따라 포함빈도 성과가 상대적으로 빠르게 개선된다는 점에서 장점을 가지고 있다. 표본의 크기가 8,192일 때, BSS기법의 포함빈도는 의도된 수준에 비하여 프로세스 4에서 6.6%, 프로세스 10에서 3.5% 차이를 보이고 있다. 그러나 BSS기법의 구간폭과 관측된 포함빈도를 고려하면 이 두 사례는 BSS기법의 취약점은 나타내는 것이라고 할 수 없다. 예컨대, 프로세스4의 경우 BSS기법은 85.4%의 포함빈도를 실현하였다. 그러나 BSS기법으로 구한 신뢰구간의 구간폭의 표본평균과 표준편차는 BM기법으로 얻은 신뢰구간에 비하여 각각 72%와 60%에 머물렀다. 이 경우 포함빈도에서의 낮은 성과는 구간폭 면에서의 좋은 성과로 보상되었다고 할 수 있다. 프로세스10의 경우도 유사하게 설명될 수 있다.

<표 4> BSS기법과 BM기법의 포함빈도 비교

(단위: %)

프로세스	512		1,024		2,048		8,092	
	BSS	BM	BSS	BM	BSS	BM	BSS	BM
1	88.5	87.0	91.5	88.0	91.0	90.0	90.5	88.0
2	89.2	91.5	89.0	91.0	94.5	90.0	93.0	94.5
3	75.5	78.5	81.0	84.0	81.0	87.5	83.5	86.0
4	64.0	77.0	72.9	81.0	81.8	83.0	85.4	92.0
5	76.1	82.0	68.4	78.0	81.3	84.5	81.9	81.0
6	61.7	90.5	65.7	86.5	80.2	87.0	91.5	98.5
7	88.1	90.0	88.4	89.5	94.0	88.0	94.0	93.5
8	67.0	64.5	75.5	75.0	81.0	82.5	86.0	86.5
9	84.0	85.0	78.0	81.0	79.5	85.5	85.5	87.5
10	71.1	72.0	78.8	80.5	77.3	80.5	81.0	84.5
11	89.5	87.0	91.5	88.0	91.5	91.5	90.5	88.0
12	89.5	87.0	91.5	88.0	91.0	90.5	90.0	88.5
13	89.3	88.5	92.0	89.5	93.5	89.0	94.0	95.0
14	89.0	89.5	88.5	89.5	94.0	87.5	93.5	93.5
15	84.3	88.5	86.2	88.5	90.5	88.0	93.5	94.5

2) 구간폭의 평균 비교: <표 5>는 구간폭의 평균 면에서 BSS기법이 BM기법보다 우수함을 보여 주고 있다. BSS기법으로 얻은 신뢰구간의 구간폭 평균이 BM기법으로 구

한 신뢰구간에 비해서 적은 경우가 그 반대인 경우보다 훨씬 더 많다. BSS기법의 포함빈도 성과가 BM기법보다 다소 저조했던 프로세스 4, 5, 6에서 표본규모에 상관없이 BSS기법이 평균적으로 폭이 좁은 신뢰구간을 산출한 점은 당연하다. 그러나 두 기법의 포함빈도 성과가 거의 대등했던 프로세스 1, 2, 7에서도 BSS기법의 구간폭 평균이 적은 점은 확실히 BSS기법의 장점을 나타내고 있다. 프로세스 8과 10의 경우에는 표본규모에 따라 BSS기법의 구간폭 성과가 떨어지고 있어서 BSS기법의 한계를 보여주는 것처럼 보인다. 이 두 경우는 표본이 적을 경우에는 BSS기법이 불리할 수도 있다는 증거를 보여주는 사례라고 보인다.

<표 5> BSS기법과 BM기법의 비교(신뢰구간폭의 평균)

프로세스	512		1,024		2,048		8,092	
	BSS	BM	BSS	BM	BSS	BM	BSS	BM
1	0.72	0.88	0.51	0.64	0.36	0.47	0.18	0.23
2	0.13	0.17	0.09	0.12	0.06	0.09	0.05	0.04
3	2.46	2.22	1.43	1.72	1.03	1.36	0.56	0.79
4	1.35	2.31	1.24	1.91	1.02	1.42	0.60	0.83
5	1.78	2.11	1.47	1.75	1.18	1.40	0.58	0.76
6	2.26	4.47	1.98	3.38	1.91	2.36	1.05	1.19
7	0.50	0.62	1.35	0.44	0.25	0.32	0.12	0.16
8	23.57	8.56	18.28	8.66	12.39	8.3	4.61	5.71
9	1.53	1.34	0.88	0.98	0.56	0.77	0.29	0.39
10	14.55	6.79	10.60	6.59	6.17	5.26	2.42	3.27
11	0.14	0.18	0.10	0.13	0.07	0.09	0.04	0.05
12	0.14	0.18	0.26	0.33	0.18	0.23	0.09	0.11
13	0.10	0.12	0.07	0.09	0.05	0.06	0.02	0.03
14	0.34	0.42	0.24	0.31	0.17	0.22	0.08	0.11
15	0.36	0.46	0.26	0.33	0.17	0.24	0.09	0.12

종합하면, 표본의 크기가 8,192일 때에는, BSS기법이 프로세스2를 제외하고 모든 테스트모델에서 평균적으로 좁은 신뢰구간을 제공하였다. 이 점은 표본이 클 때 BSS기법이 구간폭 면에서 보다 우수함을 보여주는 것이다.

3) 구간폭의 표준편차 비교: 표본이 적을 때에는 BSS기법으로 얻은 신뢰구간의 폭의 표준편차가 BM기법의 경우에 비해서 큰 경우도 있으나 표본이 8,192일 경우에는 모든 프로세스에서 적었음을 <표 6>에서 알 수 있다. 표본이 클 경우에는 BSS기법이 BM기법에 비하여 신뢰구간의 폭이 보다 안정적이라는 점에서 더 신뢰할 만 하다.

<표 6> BSS기법과 BM기법의 비교 (구간폭의 분산)

프로세스	512		1,024		2,048		8,092	
	BSS	BM	BSS	BM	BSS	BM	BSS	BM
1	0.14	0.36	0.07	0.26	0.04	0.19	0.02	0.11
2	0.01	0.08	0.01	0.05	0.00	0.04	0.14	0.02
3	4.14	1.67	1.26	1.11	0.63	0.84	0.20	0.46
4	1.01	2.14	0.90	1.45	0.61	0.89	0.28	1.46
5	3.05	1.62	1.43	1.27	0.75	0.85	0.21	0.49
6	2.26	1.78	1.94	1.43	0.85	0.96	0.19	0.52
7	0.05	0.26	0.03	0.20	0.01	0.14	0.01	0.07
8	24.33	4.91	23.37	5.12	20.72	5.40	3.28	3.83
9	3.77	1.07	1.54	0.74	0.36	0.51	0.09	0.20
10	24.14	5.51	20.72	5.14	12.77	4.49	1.37	2.16
11	0.01	0.08	0.01	0.05	0.00	0.04	0.00	0.02
12	0.01	0.08	0.02	0.13	0.01	0.10	0.00	0.05
13	0.04	0.05	0.01	0.04	0.02	0.03	0.00	0.01
14	0.04	0.18	0.02	0.14	0.01	0.10	0.00	0.05
15	0.16	0.19	0.19	0.15	0.03	0.11	0.01	0.05

4.5 실증분석 결과의 해석

실증분석 결과, 표본의 크기가 증가함에 따라 BSS기법의 성과가 BM기법에 비하여 더 빠르게 개선되는 것을 발견할 수 있다. 이 결과는 다음과 같은 해석을 가능하게 한다.

BSS기법에서는 BM프로세스가 간단한 ARMA모델로 잘 표현될 수 있을 때까지 데이터를 집계한다. 이때 우리는 데이터집계의 수준과 자료표현모델의 단순성 사이에서 절충을 고려할 수 있다. 실증분석의 결과는, 표본이 클 경우(본 사례에서는 $n=8,092$ 일 경우), 매우 간단한 ARMA모델을 자료표현모델로 사용하는 것이 유리하다는 점을 보여주고 있다. ARMA(1,1)모델을 사용한 경우와 비교하면, AR(1)모델을 사용한 경우는 표본이 클 경우 실제 포함확률이 다소 우수한 것으로 나타났다. 그리고 AR(1)모델을 사용하는 것이 모수의 추정을 위한 계산이 훨씬 간단하다는 점에서 유리하다. 이론적으로 어느 표현모델이 보다 적절하냐는 점은 중요한 것 같지 않다. 만일 데이터집계가 어느 정도 진전되면 데이터에 남아있는 자동상관이 표본평균의 분산에 미치는 영향을 측정하기 위하여 자료표현모델은 AR모수 하나만 가져도 충분하다.

AR(1)모델을 사용한 BSS기법은 BM기법과 Schreiber-Andrew의 ARMA모델선정기법 사이의 절충이라고 할 수

있다. BM기법은 데이터집계의 과정을 통하여 원래의 자료에 있는 자동상관을 완전히 제거한 후 구간평균 프로세스에 ARMA(0,0)모델을 적용하려고 하는 것이다. 반면에, ARMA모델선정기법은 원래의 데이터에 있는 상호의존구조를 그대로 두고 ARMA 표현을 통하여 이 구조에 대한 정보를 도출하려 하는 것이다. AR(1)모델을 사용한 BSS기법은 적절한 정도의 데이터집계를 통하여 상호의존구조를 단순화하려고 시도하는 것이다. 이 절충방안의 잇점은 이제 실증적으로 뒷받침되었다. 왜냐하면 표본이 커짐에 따라 BSS기법의 성과가 BM기법의 성과에 비하여 더 빠르게 개선됨이 보여졌기 때문이다.

데이터집계에 의하여 구간평균 프로세스에 잔류하는 상호의존성이 약해짐에 따라 남아있는 상호의존성을 줄이는 것이 점차로 어려워진다. 그러므로 표본이 커짐에 따라 BM기법의 성과가 개선되는 속도는 점차로 떨어진다. 그러나, 실험결과에 의하면, 구간규모가 커짐에 따라 1차 AR모수가 고차의 모수를 지배하게 되어 AR(1)모델은 매우 빠른 속도로 BM프로세스의 표현에 적합하게 된다. 결론적으로, 우리는 BM기법과 ARMA모델선정기법사이에서 절충을 취함으로써 이익을 얻을 수 있다. 우리는 BSS기법을 사용함으로써 과도한 데이터집계를 피하여 보다 정밀한 신뢰구간을 구할 수 있다.

5. 결론

실증분석 결과 우리는 많은 표본상황에서 BM기법에 비하여 BSS기법의 성과가 우수하였고 특히 표본의 크기가 증가함에 따라 BSS기법의 성과가 더 빠르게 개선되는 장점을 가지고 있다고 결론지을 수 있다.

그러나 BSS기법은, 구간평균 프로세스를 대표하는 ARMA모델을 추정하고 검증하기 위하여, BM기법에 비하여 많은 계산을 필요로 한다. 그러므로 BSS기법은 출력 프로세스에서 하나의 관측치를 얻는데 계산 부담이 큰 시뮬레이션의 경우에 상대적으로 유리한 분석기법이라고 할 수 있다. 또한, 빈도는 낮으나, 검토 집합내에 있는 구간규모를 모두 검토한 후에도 적정 구간규모를 찾지 못하여 신뢰구간을 구할 수 없는 경우가 있다는 제약점을 가지고 있다.

마지막으로, BSS기법의 개선을 위하여 향후에 수행할 연구과제를 다음과 같이 제안한다.

검토집합에 대한 연구; 보다 큰 검토집합을 사용함으로써 BSS기법의 성과가 향상될 수 있다. 만일 검토집합에 보다 많은 구간규모가 포함되면 덜 적절한 모델을 선정해야 하는 기회가 줄어들 수 있다. 그러나 검토집합이 커짐에 따라 분석을 위한 계산의 부담이 증가하므로 검토할 구간규모를 추가함으로써 얻는 비용과 이익을 비교하여야 할 것이다. 이러한 관점에서 검토집합의 크기를 결정하는 문제에 관련된 실증연구는 유익할 것으로 보인다.

탐색방향에 대한 연구; 본 연구에서는 검토집합내에서 적정 구간규모를 탐색함에 있어서 작은 구간규모로부터 시작하였다. 그러나 검토집합내의 가장 큰 구간규모로부터 탐색을 시작하는 식으로 탐색방향을 바꾸면 보다 빨리 적정규모를 결정할 수 있게 되고 따라서 계산의 부담이 줄어들 것은 분명하다. 왜냐하면 구간규모가 클수록 간단한 ARMA모델로 잘 표현될 가능성이 커지기 때문이다. 탐색방향을 바꾸었을 때 BSS기법의 성과가 바람직한 방향으로 그리고 의미있는 정도로 변하는지 실증분석해 볼 필요가 있다.

순차적 기법의 개발; 표본의 크기 n 을 사전에 결정할 경우 신뢰구간의 포함확률이 심각하게 낮을 수도 있음은 이미 지적한 바와 같다. 그러므로 Law와 Carson이 동기부여한 [11] 순차적(sequential)기법을 개발하는 것이 바람직하다고 생각된다. 순차적인 BSS기법의 예로서, 구간의 수 k 를 고정시키고 구간평균간의 자동상관계수들이 특정조건을 만족시킬 때까지 구간규모 m 을 (따라서 표본의 크기 $n = km$)을 증가시키는 방법을 고려할 수 있다. 이러한 방식을 구체화하고 그 성과를 측정해 보는 실증연구는 가치있을 것으로 기대된다.

참고문헌

- [1] Amemiya T. and R.Y. Wu, "The effect of Aggregation on Prediction in the Autoregressive Model", Journal of the American Statistical Association, 67(1972), 628-632.
- [2] Brillinger, D. R., "Estimation of the Mean of a Stationary Time Series by Sampling", Journal of Applied Statistics, 10(1973), 419-431.
- [3] Crane, M. A. and D. L. Iglehart, "Simulating Stable Stochastic Systems III: Regenerative Processes and Discrete-Event Simulations", Operations Research, 23(1975), 33-45.
- [4] Fishman, G. S. "Estimating Sample Size in Computer Simulation Experiments", Management Science, 18(1971), 21 - 38.
- [5] Fishman, G. S. "Statistical Analysis for Queuing Simulations", Management Science, 20(1973), 363-369
- [6] Fishman, G. S. "Estimation in Multiserver Queuing Simulations", Operations Research, 22(1974), 72-78.
- [7] Fishman, G. S., Principles of Discrete Event Simulation, Wiley, New York, 1978.
- [8] Gray, H. L., G. D. Kelley, and D. D. McIntire, "A New Approach to ARMA Modeling", Communication in Statistics-Simulation and Computation, B7(1)(1978), 1-77.
- [9] Kang K. and B. Schmeiser, "Properties of Batch Means from Stationary ARMA Time Series", Operations Research Letters, 1987. 19-24
- [10] Law, A. M. "Confidence Intervals in Discrete event Simulation: A Comparison of Replication and Batch Means", Naval Research Logistics Quarterly, 24(1977), 667-678.
- [11] Law, A. M. and J. S. Carson, "A Sequential Procedure for Determining the Length of a Steady-State Simulation", Operations Research, 27(1979), 1011-1025.
- [12] Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-State Simulations; II. A Survey of Sequential Procedures", Management Science, 28(1982), 550-562.
- [13] Law, A. M. and W. D. Kelton. Simulation Modeling and Analysis. McGraw-Hill, New York, 1982.
- [14] Law, A. M. and W. D. Kelton, "Confidence Intervals for Steady-State Simulations; I. A Survey of Fixed Sample Size Procedures", Operations Research. 32(1984), 1221-1239.
- [15] Shreiber T. J. and R. W. Andrews, "ARMAS-based Confidence Intervals for Simulation Output Analysis", American Journal of Mathematical and Management Sciences. 4(1984), 345-373.
- [16] Schruben, L. "Confidence Interval Estimation Using Standardized Time Series", Operations Research. 30(1983), 569-590.

● 저자소개 ●

**천영수**

현재 제주대학교 경상대학 경영학과에 재직 중이다. 서울대학교 물리학과를 졸업하고(1972년) 학사편입하여 경제학과를 졸업하였다(1975년). 그 후 한국과학기술연구소에서 연구원으로, 두산산업에서 과장으로 근무하였고, 한국전자통신연구소에서 선임 연구원으로 재직하던 중에 미국의 Minnesota 대학에서 경영학 박사학위를 취득하였다(1989년). 연구관심분야는 경영과학과 경영정보론이다.