

자기상관 공정에 대한 누적합관리도에서 설계모수 값의 결정

이재준 · 이종선[†]

인하대학교 자연과학대학 통계학과

A note on CUSUM design for autocorrelated processes

Jae June Lee, Jong Seon Lee[†]

Department of Statistics, Inha University

Key Words : CUSUM, Autocorrelation, Reference value, ARMA(1,1)

Abstract

It is common to use CUSUM charts for detecting small level shifts in processes control, in which reference value(k) and decision interval(h) are the design parameters to be determined. To control process with autocorrelation, CUSUM charts could be applied to residuals obtained from fitting ARIMA models. However, constant level shifts in processes lead to varying mean shifts in residual processes and thus standard CUSUM charts may need to be modified. In this paper, we study the performance of CUSUM charts with various design parameters applied to autocorrelated processes, especially focussing on ARMA(1,1) models, and propose how they can be determined to get better performance in terms of the average run length.

1. 서 론

일반적으로 자료가 서로 독립인 공정과 달리 자기상관(autocorrelation)이 존재하는 공정의 경우 크게 두 가지 방법을 통해 공정관리를 하게 된다. 첫째는 공정 자료를 적절한 모형에 적합한 후에, 자기상관이 제거된 잔차(residual)를 관리도에 적용하는 방법이다(Alwan과 Robert, 1988). Montgomery와 Mastrangelo (1990)는 자기상관을 갖는 공정 자료를 EWMA로 적합하여 구한 잔차에 대하여 공정을 관리하는 방법을 제안하였다. 두 번째 방법으로, Runger와 Willemain(1996)는 배치(batch)의 크기를 배치평균 사이의 자기상관이 소멸되도록 결정하고 배치평균을 \bar{X} 관리도와 같은 슈하르트 관리도에 적용하는 방법을 제안하였다.

모형적합으로 구한 잔차를 슈와르트 관리도에 적용하는 방법은 일반적으로 공정의 작은 평균 변화에 둔감

한 단점이 있다. 이에 대한 대안으로 Runger et al. (1995)는 누적합관리도(CUSUM)를 잔차에 적용하여 작은 공정의 변화에 대해서도 빨리 공정 변화를 검출할 수 있는 방법을 제안하였다. 하지만, 일반적으로 자기상관이 존재하는 공정에 지속적인 평균변화(persistent Mean Shift; MS)가 발생하는 경우, 모형을 적합하여 구한 잔차의 평균 수준은 시점에 따라 달라지는데, 모형과 모수에 따라 일정한 형태의 패턴을 나타내게 된다(Hu와 Roan, 1996). 예를 들어, ARMA(1,1) 과정을 따르는 공정에서는 잔차의 평균이 점진적으로 증가하거나 감소하는 형태, 또는 진동(oscillation)의 패턴을 나타낸다.

Runger et al.(1995)는 AR(1) 과정에 대하여 누적합관리도의 설계모수 중 하나인 참고값(k)을 결정 방법을 제안했는데, 이 방법이 일반적인 ARMA 모형에 대해서도 타당한 지에 대한 검토가 필요하다. 본 연구에서는 ARMA(1,1) 과정을 중심으로, 자기상관 공정에서 MS 의 발생을 탐지할 수 있는 기존의 누적합관리도 설계방

[†] 교신저자 jslee@stat.inha.ac.kr

* 이 논문은 인하대학교의 지원에 의해 연구되었음.

법과 그 효용성을 분석하고, 몇 가지 추가적인 방법을 제시하고자 한다.

2. 자기상관 자료의 CUSUM관리도 적용

2.1 ARMA(1,1) 모형과 잔차

ARMA(1,1) 과정을 따르는 Z_t 는 다음과 같은 모형으로 표현된다.

$$(1 - \phi B)Z_t = (1 - \theta B)a_t \quad (1)$$

여기서 a_t 는 서로 독립인 정규분포 $N(0, \sigma^2)$ 를 따르고 $\phi \neq \theta$ 를 가정한다. 이러한 ARMA(1,1) 과정을 따르는 자료에 대해 시점 t 에서 구한 1-시차 후 평균제곱 오차예측값(Minimum Mean Square Error Forecast; MMSEF) $\hat{Z}_t(1)$ 과 예측오차 $e_t(1)$ 는 다음과 같이 표현된다.

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_t \quad (2)$$

여기서, $\hat{Z}_t(1) = \phi Z_t + \theta e_{t-1}$ 이다. 즉, 모형의 모수 (ϕ, θ) 를 알 때 구해지는 잔차는 $e_t(1)$ 로서, 식(2)에 의해 서로 독립인 과정이 된다.

2.2 잔차 평균의 변화 패턴

자기상관이 존재하는 공정 자료 Z_t 에서 이 Z_t 에 일정한 크기의 지속되는 평균변화(MS)가 발생하면, 잔차 과정(e_t)에는 다양한 형태의 평균변화가 나타나게 된다. Hu와 Roan(1996)은 자기상관이 존재하는 ARMA 모형에서 모수가 달라짐에 따른 다양한 잔차 평균의 변화 패턴을 제시하였다. <그림 1>은 식(1)의 ARMA(1,1) 모형에서, 다양한 ϕ 와 θ 에 대하여 잔차 평균이 어떻게 변화되는지를 나타낸 그림이다. 잔차 평균의 변화 패턴은 크게 세 가지 모습으로 나타나게 된다. 구체적으로, 점진적으로 감소하는 형태, 점진적으로 증가하는 형태, 또는 진폭이 점점 작아지는 진동(oscillation)의 형태를 나타내는데 궁극적으로 어떤 수준으로 수렴하게 된다. 이런 잔차 평균의 변화 패턴은 다음과 같이 구체적으로 표현될 수 있다.

만약, 시점 T_0 에서 Z_t 에 크기가 $\omega (= c \cdot \sigma_z)$ 인 평

균변화가 발생한 경우에, 시점 T_0 에서의 잔차, e_t 에는 ω 만큼의 평균 변화가 나타나지만 시점 T_0 이후에서는 시점별 평균변화의 크기가 식 (4)와 같게 된다 「Lee와 Cho, 2004」.

$$\omega \cdot \left[1 - \frac{(\phi - \theta)(1 - \theta^j)}{1 - \theta} \right] \quad (4)$$

여기서, $j (> T_0 + k, k = 1, 2, \dots)$ 는 공정평균의 발생 후 경과된 시차이다. 따라서 시간이 충분히 경과한 후($j \rightarrow \infty$) 잔차에 나타나는 평균 변화의 크기는 식(5)의 값으로 수렴하게 되고, 이를 $MS_\infty(e)$ 라 표기한다.

$$MS_\infty(e) = \omega \cdot \left[1 - \frac{(\phi - \theta)}{1 - \theta} \right] \quad (5)$$

2.3 누적합관리도의 적용

ARMA(1,1) 모형을 따르는 과정 Z_t 에서 평균변화의 발생을 탐지하기 위해 Z_t 자료를 모형에 적합하여 구한 잔차에 누적합관리도를 적용하기 위해서는 참고값(k)의 설정이 필요하다. 표준편차가 σ 인 iid(independent and identically distributed) 공정에서 참고값은 관심 있는 평균변화의 크기, $\delta (= |\mu_1 - \mu_0| / \sigma)$ 를 고려하여 식 (6)과 같이 설정되는데 흔히 $\delta = 1\sigma$ 를 적용한다.

$$k = \frac{\delta}{2\sigma} \quad (6)$$

자기상관이 존재하는 공정의 관리에서, 공정 자료를 모형에 적합하여 구한 잔차를 CUSUM 관리도를 적용하는 경우에 공정에 지속되는 평균 변화가 발생하더라도 잔차의 평균은 식(5)와 같이 시간이 흐름에 따라 달라진다. 따라서 자기상관 공정관리에서는 식(6)과는 다른 타당한 참고값의 설정이 필요하다. 이런 배경에서 Runger *et al.*(1995)는 AR(1) 과정의 잔차 평균의 변화 특성을 고려하여 CUSUM 관리도의 참고값으로 식 (7)을 제안하였다.

$$k_\infty = \frac{E[e_\infty]}{2\sigma} \quad (7)$$

여기서, $E[e_\infty]$ 는 잔차 평균의 궁극적 수렴값으로,

ARMA(1,1) 과정의 경우 식(5)와 같이 구체적으로 표현된다. 여기에서 σ 는 누적합관리도에 적용되는 잔차의 표준편차로서 식(1)에 정의되어 있다.

AR(1) 과정의 경우 잔차 평균은 T_0 시점에서 ω 가 되고 그 이후 시점부터 같은 크기의 평균이 지속되는 <그림 1>의 ($\phi = 0.6, \theta = -0.3$)와 유사한 패턴을 보인다. 반면에 ARMA(1,1) 모형의 경우 AR(1) 과정과 달리 모수에 따라 잔차 평균이 시점에 따라 크기와 그 형태가 다양하게 나타나므로 식(7)의 참고값이 모든 모형에서 최적인지를 검토할 필요가 있다. 따라서 본 논문에서는 ARMA(1,1) 과정을 따르는 공정에 대해 식(7)의 참고값 보다 민감도가 높은 누적합관리도를 설계할 수 있는 참고값이 존재하는지를 살펴보기로 한다. 구체적으로 잔차 평균의 수렴값으로 설정한 식(7)의 참고값과 T_0 시점의 잔차 평균으로 설정한 식(8)의 참고값을 각각 CUSUM 관리도에 적용하여 두 관리도의 효용성(성능)에 어떠한 차이가 나타나는지를 실증적으로 분석한다.

$$k_Z = \sigma_Z / 2\sigma \tag{8}$$

여기서 σ_Z 는 ARMA(1,1) 과정을 따르는 공정 자료

의 표준편차로서, 식(9)로부터 구해진다.

$$\sigma_Z^2 = \frac{1 + \theta^2 - 2\phi\theta}{1 - \phi^2} \sigma_e^2 \tag{9}$$

따라서 잔차에 대한 누적합관리도는 식(7) 또는 (8)을 참고값으로 식(10)과 같이 작성된다.

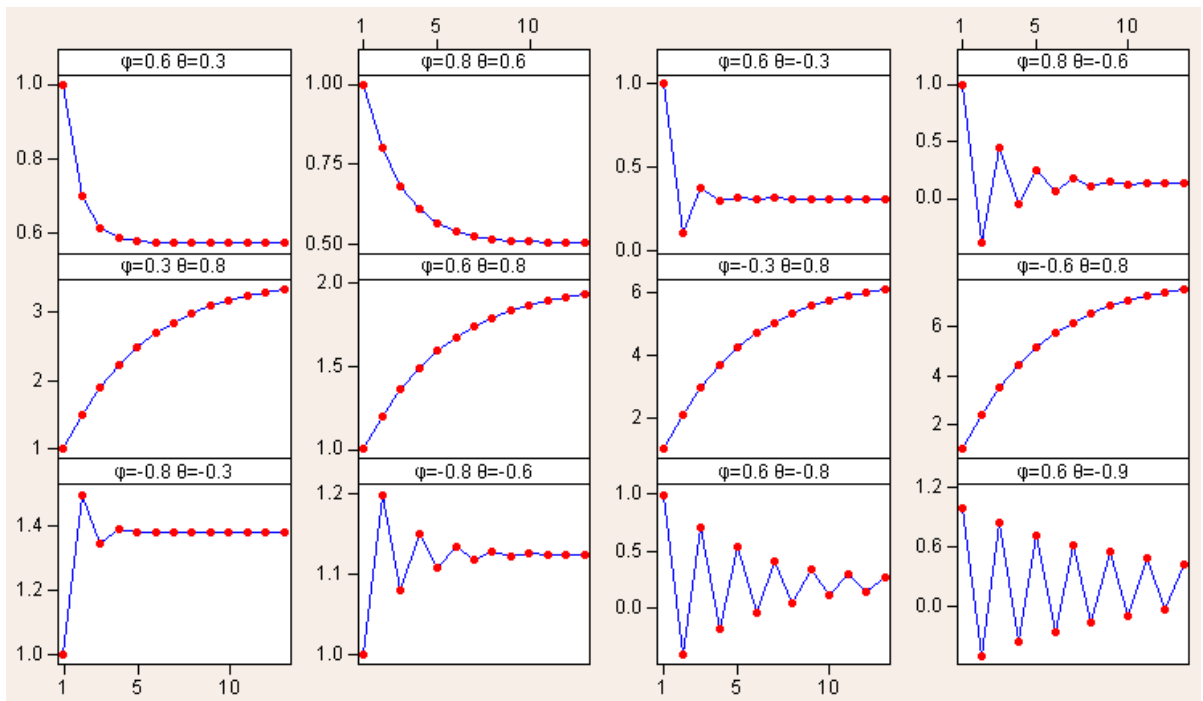
$$\begin{aligned} S_n^+ &= \max(S_{n-1}^+ + e_n - k, 0) \\ S_n^- &= \max(S_{n-1}^- - e_n - k, 0) \end{aligned} \tag{10}$$

여기서, e_n 은 시점 n 에서의 잔차이고 $S_0^+ = 0, S_0^- = 0$ 으로 설정한다. 누적합관리도의 관리한계 h 는 식(7)과 식(8)의 참고값에 대하여 관리상태의 평균 런 길이가(ARL_0)가 각각 370이 되도록 Hawkins와 Owell (1997)의 방법으로 구하였다.

3. 모의실험

Runger *et al.*(1995)이 제안한 누적합관리도의 참고값(k_∞)과 식(8)에서 제시한 참고값(k_Z)을 대상으로, k_∞ 을 사용하는 것이 모든 형태의 모형에서 최적인지 알아

<그림 1> ARMA(1,1) 과정을 따르는 공정 평균 변화에 따른 잔차 평균의 변화 패턴



보기 위해 다음과 같이 모의실험을 수행하였다. 이 과정에서 대표적인 두 참고값 (k_{∞} 와 k_Z)과 함께 그 값들에 각각 ± 0.05 단위만큼 가감한 참고값들을 추가로 고려하였고, 각 대상 참고값에 대해 Hawkins와 Owell (1997)의 방법으로 관리한계를 구한 결과는 <표 1>과 <표 2>의 첫 3 열에 요약하였다. 두 표에서 k_{∞} 와 k_Z 의 경우는 굵은 글씨체로 표시하였다.

단계1: 먼저 부록의 <표 1>과 <표 2>에 제시한 다양한 ARMA(1,1) 모형에 대하여 그 모형을 따르는 Z_t 를 생성하고 Z_t 에 크기가 $\omega = c\sigma_Z$ ($c = 0.5, 0.75, 1.0, 1.5, 2.0, 3.0$)인 MS를 각각 부과하여 모의자료를 생성하였다. 여기에서 ARMA(1,1)의 오차 a_t 는 $N(0,1)$ 을 따르도록 생성하고 σ_Z 는 식(9)로 계산하였다.

단계2: 이렇게 생성한 MS 자료를 모형에 적합하여 잔차를 구하고, 그 잔차를 <표 1>과 <표 2>에서 제시된 설계모수로 설정한 누적합관리도에 적용하여 관리이탈이 탐지되는 시점까지의 런 길이를 측정한다.

단계3: 각 모형과 각 누적합관리도의 설계방법(모수) 대하여, 단계1과 단계2의 과정을 5000회의 모의실험을 실시하여 평균 런 길이(ARL_1)를 구한다.

대표적인 패턴을 나타내는 ARMA(1,1) 모형에 대하여 모의실험을 수행하여 구한 ARL_1 값들을 정리한 결과는 <표 1>과 <표 2>과 같다. 이들로부터, 모의실험에서 고려된 설계방법 중에서 Runger *et al.*(1995)의 참고값(k_{∞})는 항상 최적인지, 아니면 어느 모형에 어떤 설계모수가 가장 좋은 지에 대해 다음과 같은 사실을 확인할 수 있다.

(i) 잔차의 평균이 증가하는 패턴의 모형

<표 1>과 <표 2>의 $\phi = 0.6, \theta = 0.8$ 인 모형의 경우, 최초 공정평균의 크기가 관리도 설계과정에서 고려한 $1.0\sigma_Z$ 보다 작은 $0.5\sigma_Z$ 의 평균변화가 발생했을 때 k_Z 를 적용하여 구한 $ARL_1 = 12.5$ 는 k_{∞} 를 적용하여 구한 $ARL_1 = 16.5$ 보다 더 좋은 결과를 보인다. 이와 같은 모형에서는 k_{∞} 보다는 k_Z 를 이용하는 것이 $1.0\sigma_Z$ 미만의 작은 평균변화에서 ARL_1 이 작게 나타났다. 또한 잔차

평균이 증가하는 패턴의 모형에서는 다른 패턴의 모형보다도 ARL_1 이 전반적으로 작게 나타나는 것을 알 수 있다. 이와 같은 현상은 참고값 k_{∞} 가 참고값 k_Z 보다 크게 설정됨에 따라 참고값 k_{∞} 의 경우 상대적으로 부적합이 작게 되어 관리이탈이 늦어지게 되기 때문이다.

반면에 최초 공정평균의 크기가 관리도 설계과정에서 고려한 $1.0\sigma_Z$ 보다 같거나 큰 모든 경우에서는 참고값을 k_{∞} 로 설정한 설계의 관리한계 h 가 참고값을 k_Z 로 설정한 설계보다 더 작은 값으로 결정되어 상대적으로 더 작은 ARL_1 을 보인다.

(ii) 잔차의 평균이 감소하는 패턴의 모형

잔차의 평균이 감소하는 $\phi = 0.6, \theta = 0.3$ 인 모형은 Runger *et al.*(1995)이 고려한 AR(1) 모형과 유사한 패턴의 경우로서, 일반적으로 k_{∞} 가 더 작은 ARL_1 을 보인다. 그러나 공정평균의 변화 크기가 설계과정에서 고려한 $1.0\sigma_Z$ 보다 훨씬 큰 경우에 k_Z 를 적용하는 것이 더 좋은 결과를 나타냄을 확인할 수 있다. 예를 들어 공정에서 $\omega = 3.0\sigma_Z$ 의 평균 변화가 발생한 경우에 k_{∞} 는 $ARL_1 = 4.0$ 이지만 k_Z 는 $ARL_1 = 2.8$ 로서, k_{∞} 보다 k_Z 가 더 좋은 결과를 보인다.

(iii) 잔차의 평균이 진동 패턴의 모형

<그림 1>에서 보이듯이, $\phi = 0.6, \theta = -0.3$ 의 모형과 $\phi = 0.6, \theta = -0.8$ 그리고 $\phi = 0.6, \theta = -0.9$ 인 모형은 초기에 큰 폭의 잔차 평균의 변화가 발생한 후에 수렴값을 중심으로 진폭이 작아지며 진동이 발생하는 패턴을 보인다. 이런 모형에서는 전반적으로 k_{∞} 로 설정한 관리도가 상대적으로 작은 ARL_1 을 보이지만, $2.0\sigma_Z$ 이상의 큰 평균변화가 발생한 경우에 k_Z 를 적용하는 것이 k_{∞} 보다 ARL_1 이 더 작게 나타난다. 예를 들어, $\phi = 0.6, \theta = -0.9$ 의 모형에서 $2.0\sigma_Z$ 의 평균변화가 발생한 경우, k_Z 를 이용한 ARL_1 은 2.0인데 비해서 k_{∞} 를 이용한 경우는 ARL_1 이 10.8로 k_Z 가 훨씬 더 좋은 결과를 나타냈다.

모의실험 결과를 요약하면, 참고값 k_{∞} 는 k_Z 보다 일반적으로 더 좋은 결과를 보이지만, ARMA(1,1) 모형에 따라 설계과정에서 고려한 공정 평균의 변화 크기인 $1.0\sigma_Z$ 보다 더 작거나 큰 평균의 변화가 발생한 경우에 k_Z 가 더 좋은 결과를 보이는 경우가 혼함을 확인하였

다. 또한 < 표 1>과 < 표 2>에서 확인할 수 있듯이, k_{∞} 와 k_z 에 각각 ± 0.05 단위만큼 가감한 참고값들 중에서 더 작은 ARL_1 이 얻어지는 경우가 발생할 수 있음을 알 수 있다.

4. 결 론

자기상관이 존재하는 공정자료에 대하여 잔차를 관리도에 적용하여 공정을 관리하는 방법이 많이 사용되

고 있다. 특히, 누적합관리도는 공정평균의 작은 변화를 탐지하기 위해 제안된 방법으로, 최적의 참고값(k)을 결정하는 것이 중요한 과제가 된다. 자기상관 자료로부터 구한 잔차를 누적합관리도에 적용하는 공정관리 방법에서, 기존의 Runger *et al.*(1995)이 제시한 참고값(k_{∞})이 주로 적용되고 있는데, 본 논문의 모의실험을 통해 그 참고값의 효용성을 검토하였다. 다양한 모수값의 ARMA(1,1) 모형에 초점을 두고 살펴볼 때, 참고값 k_{∞} 가 일반적으로 좋은 결과를 보이지만, 잔차 평균의

<표 2> $k_z = \sigma_z/2\sigma$ 를 이용한 ARL_1

	k_z	h_z	$0.5\sigma_z$	$0.75\sigma_z$	$1.0\sigma_z$	$1.5\sigma_z$	$2.0\sigma_z$	$3.0\sigma_z$
$\phi = 0.6$ $\theta = 0.8$	0.4654	5.062	12.3	7.8	5.9	4.1	3.2	2.3
	0.5154	4.654	12.4	7.7	5.8	3.9	3.1	2.2
	0.5654	4.303	12.5	7.7	5.6	3.8	3.0	2.1
$\phi = 0.6$ $\theta = 0.3$	0.4840	4.903	85.5	40.1	22.4	10.2	6.0	3.0
	0.5340	4.518	92.4	43.0	23.2	10.1	5.8	2.8
	0.5840	4.185	99.3	47.2	24.7	10.3	5.6	2.7
$\phi = 0.6$ $\theta = -0.3$	0.7026	3.548	166.2	95.3	52.7	19.5	8.9	2.3
	0.7526	3.328	181.6	99.9	55.5	20.9	8.7	2.1
	0.8026	3.131	183.4	102.7	59.5	22.2	9.2	2.0
$\phi = 0.6$ $\theta = -0.8$	0.9578	2.629	213.4	123.2	73.4	20.4	4.6	1.0
	1.0078	2.496	217.4	131.5	74.3	20.6	4.2	1.0
	1.0578	2.374	219.8	217.1	137.0	21.7	4.1	1.0
$\phi = 0.6$ $\theta = -0.9$	1.0125	2.484	211.0	120.5	67.3	13.4	2.3	1.0
	1.0625	2.364	213.0	128.5	69.3	13.6	2.0	1.0
	1.1125	2.252	218.4	130.4	71.8	13.2	2.0	1.0

<표 1> $k_{\infty} = E[e_{\infty}]/2\sigma$ 를 이용한 ARL_1

	k_{∞}	h_{∞}	$0.5\sigma_z$	$0.75\sigma_z$	$1.0\sigma_z$	$1.5\sigma_z$	$2.0\sigma_z$	$3.0\sigma_z$
$\phi = 0.6$ $\theta = 0.8$	0.9818	2.565	16.0	8.2	5.5	3.5	2.5	1.7
	1.0318	2.439	16.5	8.3	5.7	3.5	2.5	1.7
	1.0818	2.320	16.7	8.4	5.7	3.5	2.5	1.7
$\phi = 0.6$ $\theta = 0.3$	0.2551	7.904	62.1	31.7	20.7	11.4	7.7	4.5
	0.3051	6.981	66.6	32.4	20.4	10.8	7.2	4.0
	0.3551	6.249	71.8	34.4	20.3	10.5	6.7	3.6
$\phi = 0.6$ $\theta = -0.3$	0.1816	9.811	83.4	47.3	31.6	17.7	12.0	6.9
	0.2316	8.426	91.0	49.3	30.8	16.6	10.9	6.0
	0.2816	7.386	95.7	50.6	31.7	16.0	10.0	5.3
$\phi = 0.6$ $\theta = -0.8$	0.1740	10.062	89.7	49.7	32.6	18.6	12.5	6.9
	0.2240	8.608	94.4	51.1	32.0	17.1	11.1	5.8
	0.2740	7.526	100.9	52.5	32.2	16.5	10.2	4.9
$\phi = 0.6$ $\theta = -0.9$	0.1737	10.078	88.9	49.6	32.3	18.4	12.2	6.5
	0.2237	8.617	95.1	50.5	31.9	17.0	10.8	5.1
	0.2737	7.535	101.6	53.4	32.3	16.1	9.5	4.3

변화 패턴과 공정평균의 변화 크기, ω 에 따라 k_{∞} 보다 k_Z 를 사용한 참고값이 더 좋은 결과를 보일 수 있음을 알 수 있다. 따라서 자기상관 자료의 누적합관리도 적용 시에는 k_{∞} 와 k_Z 구간을 세분화하여 참고값 대상의 폭을 확대하고, 모의실험을 통해 최적의 설계모수를 결정하는 것이 필요하다고 판단한다.

참고문헌

- [1] Alwan, L. C. and Roberts, H. V.(1988), "Time-Series Modeling for Statistical Process Control", *Journal of Business & Economic Statistics*, Vol. 6, pp. 87-95.
- [2] Douglas C. Montgomery and Christina M. Mastrangelo (1991), "Some Statistical Process Control Methods for Autocorrelated Data", *Journal of Quality Technology*, Vol. 23, No. 3, 179-197
- [3] Douglas M. Hawkins and David H. Owell (1997), "*Cumulative Sum Charts and Changing for Quality Improvement*", Springer, New York
- [4] Jae Jun Lee, Sinsup Cho (2004), "Problems of Special Causes in Feedback Adjustment", *Journal of the Korean Society for Quality management*, Vol. 32, No. 2, pp. 201-211
- [5] George C Runger, Thomas R. Willemain and Sharad Prabhu (1995), "Average Run Length for CUSUM Control Charts Applied to Residuals", *Communications Statistics-Theory and Method*, Vol.24,No.1, pp.273-282
- [6] Runger G. C. and T. R. Willemain (1996), "Batch Mean Control Charts for Autocorrelated Data", *IIE Transaction*, Vol. 28(6), pp. 483-487
- [7] S. Jack Hu and Chinmo Roan (1996), "Change Patterns of Time Series Based Control Charts", *Journal of Quality Technology*, Vol. 28, No. 3, pp. 302-312