JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 25 No. 12, pp. 73-82, December 2020
https://doi.org/10.9708/jksci.2020.25.12.073

A Study of The reference value of the CUSUM control chart that can detect small average changes in the process

Sang-Pyo Jun*

*Professor, College of General Education, Namseoul Unversity, Cheonan, Korea

[Abstract]

Most process date such as semiconductor and petrochemical processes, autocorrelation often exists between observed data, but when the existing SPC(Statistical process control) is applied to these processes, it is not possible to effectively detect the average change of the process.

In this paper, when the average change of a certain size occurs in the process data following a specific time series model, the average of the residuals changes according to the passage of time, and the change pattern of the average is introduced around the ARMA(1,1) process. Based on this result, the reference value required in the design process of the CUSUM (Cumulative sum) control chart is appropriately considered by considering the type of the time series model of the process data of the CUSUM control chart that can detect small mean changes in the process and the width of the process mean change of interest. It was confirmed through simulation that it should be selected and used.

▸ **Key words**: SPC, Autocorrelation, CUSUM chart, ARMA(1,1) process, Reference values

[요 약]

반도체나 석유화학 공정과 같이 프로세스 중심의 장치 산업에서는 흔히 관측된 자료들 사이에 자기상관(Autocorrelation)이 존재하는데, 이러한 공정에 기존의 SPC(Statistical process control)를 적용하는 경우 공정의 평균 변화를 효과적으로 검출하지 못하는 문제가 발생할 수 있다.

본 논문에서는 특정 시계열 모형을 따르는 공정자료에 일정한 크기의 평균 변화가 발생할 때, 잔차는 시간의 흐름에 따라 그 평균이 달라지게 되는데, ARMA(1,1) 과정을 중심으로 평균의 변화 패턴을 소개하고, 이 결과를 바탕으로 공정의 작은 평균 변화를 검출할 수 있는 CUSUM(Cumulative sum) 관리도의 공정 자료가 갖는 시계열 모형의 형태와 관심 있는 공정 평균 변화의 폭을 고려하여 CUSUM 관리도의 설계 과정에서 필요한 참고값이 적절히 선택되어 사용되어야 함을 모의실험을 통해 확인하였다.

▸ **주제어**: 통계적 공정관리, 자기상관, CUSUM 관리도, ARMA(1,1), 참고값

• First Author: Sang-Pyo Jun, Corresponding Author: Sang-Pyo Jun
*Sang-Pyo Jun (spjun7129@naver.com), College of General Education, Namseoul Unversity
• Received: 2020. 11. 19, Revised: 2020. 11. 26, Accepted: 2020. 11. 27.

I. Introduction

컴퓨터의 자료 처리 속도가 증가하여 일반 제조업체에서는 생산되는 모든 제품에 대해 자동 측정이 가능하게 되었다. 컴퓨터 기술의 발전으로 생산품 전체를 측정하면서 신속한 양품과 불량품 판정에는 실효를 거두고 있지만, 전통적인 통계적 공정관리 측면에서는 몇 가지 고려해야 하는 문제가 야기되었다. 그 중 하나가 측정되는 개별 자료들 사이에 기존의 공정관리에서 가정된 독립성이 위배되고 자기상관(autocorrelation)이 존재하는 변화를 둘 수 있다.

통계적 공정관리(statistical process control)에 대한 논의는 1920년대 초 Walter A. Shewhart에 의해 관리도가 제안되면서부터라고 할 수 있다. 이러한 통계적 공정관리의 목적은 우연원인과 특수원인에 의해 나타나는 공정 내의 산포(variation)를 반영하여 공정에 변화가 발생되는지를 모니터링 하는 것이다. 특히, 특수원인에 의한 공정 변화를 조기에 검출함으로써 원인을 파악하고, 거기에 맞는 적절한 조치(corrective action)를 취하기 위함이다. 이를 통해, 공정을 통계적으로 관리상태(In-control) 하에 놓이도록 하기 위한 것이다.

현장에서 공정관리를 실시함에 있어서 공정 변화의 검출은 기업의 손익에 막대한 영향을 끼치게 된다. 사용하고 있는 관리도가 공정의 변화를 제때 검출하지 못하게 되면 제품에 산포가 크게 발생하여 공정능력을 떨어뜨리고, 심지어 다량의 불량을 내재한 제품이 검출되지 못하고 고객에게 도달되는 문제가 발생할 수 있다. 반면에, 공정에 특이한 변화가 없음에도 불구하고 공정에 관리이탈이 발생된 것으로 잘못 판단하게 되면 불분명한 원인을 파악하기 위해 시간을 낭비하거나 잘못된 원인 검출로 공정조건을 변경하여 다량의 불량품이 발생할 가능성이 생기게 된다. 이처럼 공정관리 과정에서 발생되는 잘못된 경보(false alarm)는 시간적으로나 경제적으로 기업에 큰 손실을 줄 수 있기 때문에 공정 변화에 대한 신속하고 올바른 검출 방법은 공정관리에 있어서는 아주 중요한 요소가 된다.

이러한 공정관리 방법 중에서 가장 흔하게 사용되고 있는 것은 슈하르트가 제안한 관리도이다. 이 관리도는 흔히 "슈하르트 관리도" 혹은 "3-시그마 관리도" 라고 불리는데 공정에서 관측된 자료들이 서로 독립인 가정하에서 적용될 수 있다. 하지만 자기상관이 존재하는 공정 자료에 대해 슈하르트 관리도를 그냥 적용하게 되면 앞서 언급한 잘못된 경보가 발생할 가능성이 높다. 이러한 자기상관이 존재하는 공정 자료에 대해 슈하르트 관리도를 적용하는 방법에 대해 Alwan & Robert는 관측된 자료를 시계열 모형으로 추정하고, 추정된 모형으로부터 구한 잔차에 슈하르트 관리도를 적용하는 방법을 제안했다[1]. 또한, Runger & Willemain은 뱃치 평균을 이용해 뱃치들 사이에 자기상관을 없애고 뱃치 평균에 슈하르트 관리도를 적용하는 방법 등을 제안하였다[2].

Chen & Liu은 공정에 AO(Additive Outlier), IO(Innovational Outlier), LS(Level Shift) 등과 같은 다양한 형태의 공정 변화가 발생했을 때에 잔차에 나타나는 변화를 연구하였다[3]. Antienza et al 은 이런 변화를 고려하여 잔차에 슈하르트 관리도를 적용했을 때의 민감도(sensitivity)를 분석하였다. 이 연구를 통해 공정 자료에 대해 Box-Jenkins의 ARIMA (Autoregressive Integrated Moving Average) 모형을 추정하고, 이를 통해 구한 잔차에 슈하르트 관리도를 적용하는 경우에 높은 양의 자기상관이 존재하는 모형에서는 민감도가 떨어지는 것을 보였다. 이러한 민감도 문제 해결을 위해 Runger et al., Lu & Reynolds은 잔차에 대해 CUSUM 관리도와 EWMA 관리도를 적용하는 방안을 제안하였다[2,4,5,6]. 또 선택적인 누적합 관리도에 대한 의견도 제시되었다 [7,8].

특히, ARMA(1,1) 과정과 같이 특정한 모형을 따르는 공정에서 지속적인 평균변화가 발생하면 시계열 모형으로 구한 잔차의 평균이 다양한 형태로 변하게 된다[9]. 이런 경우에 있어서 잔차에 대해 CUSUM 관리도를 적용할 때 CUSUM 관리도의 설계 모수인 참고값을 어떻게 결정해야 하는지에 대한 문제가 발생하게 된다[10].

잔차에 대해 CUSUM 관리도를 적용하는 경우에 잔차의 평균 변화를 고려한 CUSUM 관리도의 설계 모수인 참고값에 대한 새로운 제안을 하고자 한다.

II. Process management using residuals

잔차에 대한 공정관리 방법에는 MA(1,1)과정을 따르는 공정 자료에 대한 EWMA 통계량은 최소평균제곱오차예측(Minimum Mean Square Error Forecast; MMSEF)이 되고 이로부터 도출된 잔차의 평균은 0이고 표준편차가 σ_e인 iid(independent and identically distributed) 형태의 과정이 되므로 MMSEF로부터 구한 잔차에 대해 슈하르트 관리도를 적용할 수 있게 된다. 이와 같은 잔차에 대한 공정관리 방법은 두 가지의 문제점을 내포하게 된다. 첫째로, 이러한 접근 방법은 관측된 자료로부터 모형 추정이 올바르게 이뤄진다는 것을 전제로 한다는 것이다. 하지만 현실적으로 올바른 모형 추정이 어려운 경우가 많이 발

생하게 된다. 잔차에 대해 슈하르트 관리도를 적용하기 위해서는 모형 추정이 올바르게 되고 예측오차인 잔차가 iid의 조건을 만족해야 하는데, 모형 추정이 잘못되는 경우에는 잔차에 여전히 자기상관이 존재한다는 것이다. 두 번째 문제로 설사 모형 추정이 올바르게 되었다 하더라도 잔차에 적용된 슈하르트 관리도는 원공정의 작은 평균 변화를 효과적으로 검출하는데 한계가 있을 수 있다.

1. ARMA(p,q) Process

일반적으로 자기상관이 존재하는 공정은 Box et al.가 제안한 식(1)의 ARMA(p,q) 모형을 따른다고 가정한다[9].

$$\phi(B) = Z_t = \theta(B)a_t \qquad t = 1,2,\cdots,n \qquad (1)$$

여기서, $\phi(B) = (1-\phi_1 B - \cdots - \phi_p B^p)$
$\theta(B) = (1-\phi_1 B - \cdots - \phi_q B^q)$이다.

여기서 B는 후진작용소(backshift operator)로써 $B^k X_t = X_{t-k}$이다. 본 논문에서는 공정 모형이 정상성(stationarity)을 만족한다고 가정한다. 식(1)의 ARMA(p,q) 모형은 다음과 같이 무한 AR(Autoregressive)모형으로 표현 된다.

$$\pi(B)Z_t = a_t$$

여기서,
$\pi(B) = \phi(B)/\theta(B) = (1-\pi_1 B - \pi_2 B^2 - \cdots)$로써, ARMA(1,1)의 경우 $\pi_1 = 1+\phi-\theta$이고, $k \geq 2$일 때 $\pi_k = (1-\theta)(\theta-\phi)\theta^{k-2}$가 된다. 단, $\pi_0 =- 1$로 표기한다.

또한, 식(1)을 무한 MA(Moving Average)모형으로 표현하면 다음과 같다.

$$Z_t = \psi(B)a_t$$

여기서,
$\psi(B) = \theta(B)/\phi(B) = (1+\psi_1 B + \psi_2 B^2 + \cdots)$로써, ARMA(1,1)의 경우 $\psi_1 = \theta-\phi-1$이고 $k \geq 2$일 때 $\psi_k =- (1-\theta)(\theta-\phi)\theta^{k-2}$가 된다 단, $\psi_0 = 1$로 표기한다.

자기상관이 존재하는 자료에 대한 공정관리 방법으로는 ARMA(p,q) 과정을 따르는 관측치 Z_t을 직접 이용하는 방법과 추정된 모형으로부터 구한 잔차(residual)를 관리도에 적용하는 방법으로 구분된다. 여기에서 잔차는

$$e_t = Z_t - \hat{Z}_t \qquad (2)$$

로 표현되는데, $\hat{Z}_t$는 시점 t의 관측값에 대한 예측치이고, $\hat{Z}_t$이 1시차 후 MMSEF 인 경우 $e_t = a_t$가 된다.

1.1. If the model is knowledgeable

공정으로부터 수집된 표본 자료를 식(1)과 같은 ARMA(p,q) 모형으로 추정할 수 있다면 추정된 모형을 이용하여 예측을 수행하고 식(2)에 의해 잔차를 구할 수 있다. 이와 같이 모형의 형태와 모수가 알려져 있는 경우에는 잔차 e_t는 근사적으로 평균이 0이고 일정한 상수의 분산을 갖는 정규분포의 형태를 따르게 된다. 이러한 배경에서 잔차 e_t에 대해 슈하르트 관리도를 적용할 수 있게 된다.

이러한 잔차를 슈하르트 관리도에 적용하는 경우에 두 가지의 문제점이 발생할 수 있다. 첫 번째 문제는 슈하르트 관리도를 적용하는 경우 공정의 작은 변화에 대한 검출력이 떨어진다는 것이다[11,12]. 이러한 문제점에 대한 대안으로 시계열 모형을 적합하여 구한 잔차를 CUSUM 관리도에 적용하는 방법이 고려될 수 있는데, 여기에 필요한 참고값은 관리 대상인 원 공정이 독립성을 만족하지 않기 때문에 Runger et al.는 다음과 같은 새로운 참고값을 제시하였다[13].

$$k = \frac{\sigma(1-\phi)}{2} \qquad (3)$$

여기서, σ는 잔차의 표준편차이고, ϕ는 AR(1) 모형의 모수이다. 위 식(5)의 참고값은 원 공정이 AR(1) 과정을 따르는 경우에 대하여, 원 공정에서 발생한 σ크기의 지속적인 평균 변화가 잔차에서는 시간이 지남에 따라 잔차 평균이 $\sigma(1-\phi)$로 수렴됨을 반영한 것이다.

하지만, 잔차를 CUSUM 관리도에 적용하여 공정의 평균변화를 검출하기 위해 제안된 식(3)의 참고값은 모든 모형에 대해서 타당한지에 대한 근거가 부족한 문제가 있다. 이유는 공정 자료 Z_t에 지속적인 평균 변화가 나타나게 되면, 잔차에서도 모형에 따라 잔차의 평균이 다양한 형태로 변하게 됨이 연구 결과로 보고되었는데, 또 다른 문제로, 대부분의 경우에 공정 자료로부터 참 모형(true model)을 추정한다는 것이 어려운 면이 있고, 이렇게 추정한 모형으로부터 구한 잔차가 iid의 성질을 만족하지 못하는 경우가 흔히 발생하게 된다[14].

1.2. If the model is unknowledgeable

현실적으로 정확한 시계열 모형을 추정하는 것은 어려운 일이고, 또한 수많은 품질 특성에 대해 일일이 모형을 추정하여 관리한다는 것은 어려운 일이다. 따라서 대부분의 공정 자료에 대한 모형이 정확히 추정되지 못한 채 관리될 가능성이 있다.

모형이 미지인 경우에는 모형에 무관하게 관측치 만을 이용한 공정관리 방법이 사용될 수 있다. 관측치를 이용하는 방법의 가장 대표적인 UBM을 이용하는 것이다. 다만 UBM을 사용하는 경우에는 공정의 자기상관 정도에 따라서 뱃치 평균들이 독립에 근사하도록 뱃치 크기를 다르게 결정해야 하는 문제와 뱃치 크기에 따라 공정 평균의 변화를 검출하는데 까지 소요되는 자료의 수가 비현실적으로 늘어날 수 있는 문제가 있다.

2. Process control performance comparison

자료가 서로 독립인 경우, 슈하르트 관리도를 적용할 때 발생하는 제1종오류(α) 를 고려한 관리상태의 평균-런-길이(in-control Average Run Length: ARL_0)는 다음과 같이 정의된다.

$$ARL_0 = \frac{1}{\alpha}$$

여기에서, $\alpha = \Pr[Z_t \notin (LCL, UCL) \mid \mu = \mu_0]$

만약 관리도에 적용되는 자료가 정규분포를 따른다고 가정하고, 관리한계의 폭을 $\pm 3\sigma_x$로 잡는다면 $\alpha = 0.0027$이 되므로 ARL_0는 약 370 정도가 된다. 즉, 공정에 아무런 이상 원인에 의한 영향이 없는 경우에도 평균적으로 370번 마다 한 번쯤은 관리한계를 벗어나는 관리이탈(out-of-control)현상이 발생한다.

ARL_0는 여러 가지 관리도의 성능을 비교해 각 관리도의 관리한계를 공정 평균의 변화가 없는 경우에 대하여 가정하여 ARL_0가 모두 같도록 설정하는데 적용될 수 있는데, 슈하르트 관리도의 경우를 고려하여 일반적으로 약 370 정도가 된다.

관리도의 성능은 공정에 이상 원인이 발생했을 때, 관리도 상에서 얼마나 빨리 검출할 수 있는지를 통해 비교하게 되는데, 관리이탈 시 평균-런-길이(out of control Average Run Length,)를 통해 비교하게 된다. 일반적으로 ARL_1은 다음과 같이 정의된다.

$$ARL_1 = \frac{1}{1-\beta}$$

여기서, $\beta = \Pr[LCL < Z_t < UCL \mid \mu = \mu_0 + \delta\sigma]$로 나타나는 제 2종 오류. 즉, 공정 평균이 μ_0에서 $\mu_1 (= \mu_0 + \delta\sigma, \delta \neq 0)$으로 변경되었을 때, 관리한계를 이탈하는 데 까지 소요되는 런의 길이를 의미한다. 만약, $1-\beta$가 1에 가까워 진다면 ARL_1은 1에 가까운 값이 나오게 되는데, 이것은 공정에 이상 원인이 발생하게 되면 평균적으로 그 다음 시점에서 관리도에 관리이탈로 검출된다.

III. Average change in autocorrelation process

1. Mean change in time series model

식(1)과 같은 ARMA(p,q) 과정을 따르는 공정 자료에 특수원인으로 인한 공정 변화가 발생한다면, 원래의 공정 자료 Z_t는 특수원인의 영향으로 오염된(contaminated) 형태의 공정 모형 N_t로 다음과 같은 모형으로 표현된다.

$$\begin{aligned} N_t &= f(t) + Z_t \\ &= \omega\xi(B) I_t(T) + Z_t \end{aligned} \tag{4}$$

여기서, T는 공정 변화가 발생한 시점이고, $\omega (= c \cdot \sigma_z)$는 원 공정의 표준편차의 배수로 표현한 시점 T에서의 공정 변화의 크기이며, $\xi(B) I_t(T)$는 특수원인에 의해 발생 이후 시점 등에서 공정에 미친 변화의 형태를 의미한다. 여기에서 $\xi(B) = 1 + \xi_1 B + \xi_2 B^2 + \cdots$이고 $I_t(T)$는

$$I_t(T) = \begin{cases} 1 & \mathrm{for}\, t = T \\ 0 & \mathrm{for}\, t \neq T \end{cases}$$

인 지수함수이다. 예를 들어, 평균 변화(mean shift)가 발생한 경우에는 $\xi(B) = 1/1(1-B)$로 표현할 수 있다[3].

ARMA(p,q) 과정을 따르는 공정에 시점 에서 식(4)과 같이 크기가 ω인 평균 변화가 발생한다면, 1시점 후 MMSE 예측으로 구한 잔차는 예측오차에 특수원인의 영

향으로 인해 평균 수준이 변한 형태 표현될 수 있다(Lee & Cho, [11]). 따라서, 식(5)은 $t < T$에서 $N_t = Z_t$이고, $t \geq T$에서 $N_t = Z_t + \omega$가 관측되는 모형이 된다.

$$e_t = a_t + \omega(1 - \sum_{i=1}^{t-T} \pi_i) \qquad (5)$$

식(6)은 공정의 평균변화가 발생한 시점 에서는 공정 변화의 영향이 ω만큼의 잔차의 평균 변화로 나타나게 되고, 시점 $T+k(k=1,2,\cdots)$에서는 $\omega(1-\sum_{i=1}^{k}\pi_1)$의 크기로 잔차의 평균이 변하게 된다.

1.1. Process change in AR(1) process

공정자료 Z_t가 AR(1) 과정을 따른다면 Z_t는 다음과 같이 표현될 수 있다.

$$Z_t = \phi Z_{t-1} + a_t, \qquad t = 1, 2, \cdots, \ (6)$$

여기서, a_t는 평균이 0이고 분산이 σ^2인 서로 독립인 백색잡음(white noise)과정이다.

식(6)에서 ϕ의 참 모수값을 안다면, 시점 t에서의 MMSE 예측오차는 다음과 같이 표현할 수 있다.

$$e_t = Z_t - \widehat{Z}_t \qquad (7)$$

식(7)은 여기서, MMSEF는 $\widehat{Z}_t = \phi Z_{t-1}$이고, $e_t = a_t$가 된다. 즉, 잔차는 평균이 0이고 표준 편차가 σ인 되는 서로 독립인 백색잡음과정이 된다.

만약, 시점 T에서 공정 평균이 μ_0에서 $\mu_1 = \mu_0 + \omega$로 변하게 되면 실제 공정 값은 Z_t가 아닌 N_t가 발생하고, 따라서 잔차는 다음과 같이 변하게 된다.

$$\begin{aligned} e_T &= N_T - Z_{T-1}(1) \\ &= (\omega + \phi Z_{T-1} + a_T) - \phi Z_{T-1} \\ &= \omega + a_T \end{aligned}$$

여기서, $\widehat{Z_T} = Z_{T-1}(1)$은 시점 $(T-1)$에서 구한 1 시점 후 MMFE 예측치로서 $Z_{T-1}(1) = \phi Z_{T-1}$이다. 시점 $T+1$의 경우, 예측에 적용되는 시점 T에서의 공정 값은 N_T이므로 따라서,

$$\begin{aligned} e_{T+1} &= N_{T+1} - Z_T(1) \\ &= (\omega + \phi Z_T + a_{T+1} - \phi N_T) \\ &= (\omega + \phi Z_T + a_{T+1} - \phi(\omega + Z_T) \\ &= \omega(1-\phi) + a_{T+1} \\ &\vdots \end{aligned}$$

여기에서 $Z_T(1) = \phi N_T = \phi(\omega + Z_T)$이다.

위의 결과를 요약하면 식(8)과 같다.

$$\begin{aligned} e_t &= a_t & t &< T \\ &= \omega + a_t & t &= T \\ &= \omega(1+\phi) + a_t, & t &\geq T+k \quad k = 1, 2, \cdots \end{aligned} \qquad (8)$$

식(8)로 보면, 원래 공정 Z_t에 시점 T에서 ω만큼의 평균 변화가 발생하면, 잔차에서는 시점 T에서는 ω이고 그 이후 시점들 $T+k, k=1,2,\cdots$에서는 모든 시점에서 같은 크기의 $\omega(1-\phi)$만큼의 평균이 변하게 된다.

1.2. Process changes in the ARMA(1,1) process

공정자료 Z_t가 ARMA(1,1) 과정을 따른다면 Z_t의 모형은 다음과 같이 식(9)으로 표현된다.

$$(1-\phi B)Z_t = (1-\theta B)a_t \qquad (9)$$

여기서, a_t는 평균이 0이고, 표준편차는 σ인 백색잡음 과정이다.

ARMA(1,1) 과정을 따르는 과정 Z_t에서 시점 t에서 구한 시점 $(t+1)$에서의 공정 값 Z_{t+1}에 대한 1-시점 후 예측값 $Z_t(1)$은 다음과 같이 표현된다.

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

여기서, a_t는 시점 t에서의 오차이다.

이로부터 시점 t에서의 1-시점 후 MMSE 예측오차 $e_t(1)$는 $(t+1)$시점의 잔차로 다음과 같이 구해진다.

$$e_{t+1} = e_t(1) = Z_{t+1} - Z_t(1)$$

따라서 ARMA(1,1) 과정을 따르는 공정에 시점 T에서 $\omega(=\delta\sigma)$의 평균변화가 발생하는 경우에, 잔차의 평균은 시점별로 다음과 같이 나타낼 수 있다.

$$E(e_t) = \begin{cases} 0 & \text{for}\, t < T \\ \omega & \text{for}\, t = T \end{cases}$$

그리고, $t > T+k$, $k = 1,2,\cdots$인 경우에는 $\pi_1 = 1+\phi-\theta$, $\pi_i = (1-\theta)(\theta-\phi)\theta^{i-1}$, $i \geq 2$ 이므로

$$E(e_t) = \omega \cdot \left[1 - \frac{(\phi-\theta)(1-\theta^k)}{1-\theta}\right] \quad (10)$$

가 된다. 여기서, ω는 시점 T에서 발생한 원래 공정에서의 평균 변화 크기이다.

평균이 변화한 시점 T이후에 충분한 시간이 흐르게 되면 $k \rightarrow +\infty$이 되어, e_t평균은 다음과 같은 형태로 수렴하게 된다.

$$E(e_t) = \omega \cdot \left[1 - \frac{(\phi-\theta)}{1-\theta}\right] \quad (11)$$

따라서, 식(11)에 의하면 ARMA(1,1) 과정을 따르는 공정 자료에 평균 변화가 발생하면 MMSE 예측 후의 잔차의 평균은 모수에 따라 변화 크기와 형태가 변하게 된다.

공정 평균의 변화 크기 ω를 공정 관측치 Z_t의 표준편차 σ_z의 배수로 $\omega = c \cdot \sigma_z$로 표기하였는데, ARMA(1,1) 과정을 따르는 Z_t의 σ_z는 $N(0,\sigma^2)$을 따른다는 가정하에서 식(13)와 같이 표현된다

$$\sigma_z = \sigma_a \sqrt{\frac{1-2\phi\theta+\theta^2}{1-\phi^2}} \quad (12)$$

본 논문에서는 ARMA(1,1) 모형의 모수값(ϕ,θ)과 ω의 크기에 따른 잔차의 평균 변화 형태를 검토하고, 이러한 잔차를 CUSUM 관리도에 적용할 때의 성능을 모의실험을 통해 살펴보기로 한다.

2. Variety of residual mean change

θ가 양수이고 $|\phi| > \theta$인 경우에는 잔차의 평균이 시점 T에서는 크기가 ω인 값을 갖지만 그 이후 시점에서 점진적으로 감소하여 식(12)의 값으로 수렴하는 모습을 보인다. 반면에 θ가 양수이고 $|\phi| < \theta$인 경우에는 잔차의 평균이 크기가 ω인 값으로 시작하여 그 이후 시점에서 점진적으로 증가하여 식(12)의 값으로 수렴하는 형태를 나타내게 된다. 특히, θ가 음수인 경우에는 잔차의 평균이 진동(oscillation)하는 패턴으로 나타나는 것을 알 수 있다.

이처럼 자기상관이 존재하는 공정에 대해 관리도를 적용하는 경우에 MMSE 예측에 의한 잔차가 그 평균이 시점에 따라 달라지게 되어 서로 다른 형태의 패턴을 보이는 특징이 있음을 확인할 수 있다. 반면에 iid를 만족하는 공정의 경우 공정 상에 같은 크기의 평균 변화를 검출하는 것이 관리도 적용의 주된 목적인데, 자기상관이 있는 공정의 경우에 공정에서 같은 크기의 평균변화가 발생할 때 잔차의 평균이 시점에 따라 달라진다는 면에서 관리도의 효용성은 두 경우에서 차이가 있을 수 있다. 즉, 이러한 현상이 관리도의 성능에 차이를 보일 것으로 생각된다. 이러한 배경에서 본 연구에서는 다양한 잔차 패턴의 차이에 따라 기존에 제시된 관리 방법의 성능을 비교하고 자기상관이 존재하는 공정에 더 유효한 관리도의 선택을 위한 기준을 제시하고자 한다.

IV. CUSUM chart design

1. Proposal background

잔차를 관리도에 적용하는 대표적인 관리 방법으로 SCC(special cause charts)가 사용되는데, SCC의 단점은 작은 공정 변화에 둔감하다는 것이다. 이런 단점을 해결하기 위해 CUSUM 관리도가 고려될 수 있으나 3장의 ARMA(1,1) 과정의 예에서 보여지듯이 모형 추정이 올바르게 되었다 하더라도 추정된 모형으로부터 도출된 잔차에 시점에 따라 평균이 변하는 특성이 나타나기 때문에 CUSUM 관리도의 설계 모수들, 특히 참고값의 경우에 iid공정에 타당한 참고값이, 잔차의 경우에도 타당하다는 보장은 없다. 따라서, 잔차를 CUSUM 관리도에 적용할 때 참고값을 어떻게 결정하는 것이 좋은 지에 대한 논의가 필요하다. 4장에서는 이러한 문제의 해결을 위해 CUSUM 관리도의 참고값 결정 방법을 비교, 검토하고자 한다.

2. CUSUM chart design

CUSUM 관리도는 공정의 품질 특성치 자료를 ARMA모형에 적합하여 구한 잔차에 적용할 수 있는 모니터링 방법으로, 슈하르트 관리도에 비하여 CUSUM 관리도는 공정의 작은 평균 변화를 더 효율적으로 검출할 수 있다고 알려져 있다. CUSUM 관리도는 기본적으로 iid 공정에 대해 작은 평균의 변화를 검출하기 위한 방법으로 3장에서 다룬 ARMA(1,1) 공정과 같이 자기상관이 존재하는 공정에서는 관측치에 일정 수준의 지속적인 평균 변화가 발생

한 경우에 잔차의 평균은 시점에 따라 그 평균이 달라지는 특성이 있어 이를 고려한 참고값 결정이 필요하다.

2.1. Determine reference value

CUSUM 관리도에서의 참고값(reference value)은 관리상태(in-control) 하에서의 공정 평균, μ_0로부터 관리이탈로 판단할 수 있는 공정의 평균, $\mu_1 (= \mu_0 + \delta\sigma)$의 차에 의해 결정된다. 자기상관이 존재하는 공정 자료를 시계열 모형에 적합하여 구한 잔차에 대해서 Runger et al. 가 AR(1) 공정에 대한 분석을 바탕으로 식(13)과 같은 참고값 결정 방법을 제안하였다[2].

$$k_\infty = \frac{E[e_\infty]}{2}\sigma \qquad (13)$$

여기서, σ는 잔차의 표준편차이며, $E[e_\infty]$는 식(11)에서 제시된 잔차 평균의 궁극적인 수렴 값이다. 즉, 참고값은 잔차 평균의 수렴 값을 기준으로 제시된 것이다.

AR(1) 과정으로부터 도출된 잔차의 경우는 잔차의 평균이 식(10)에 나와 있는 것처럼 공정 평균이 변경된 시점 T 다음부터 $\omega(1-\phi)$의 크기로 시점에 무관하게 일정하지만, ARMA(1,1) 과정의 경우에는 잔차의 평균이 T이후의 시점에 따라 그 크기가 달라지게 된다. ARMA(1,1)과정을 따르는 공정 자료로부터 도출된 잔차에 대해 CUSUM 관리도에 적용하는 경우에는 식(13)의 참고값 k_∞가 최적인지에 대한 검토가 필요하다.

잔차의 평균이 공정 특성치의 평균이 발생한 이후 시점부터 증가 혹은 감소하는 경우와 진동의 패턴을 발생하는 경우 등을 구분하여 참고값을 어떻게 결정하는 것이 타당한 지에 대한 검토가 필요하다. 이를 위하여 식(13)이 잔차 평균의 수렴 값에 바탕을 둔 참고값이란 점을 고려하여, 평균 변화의 발생 시점 T에서 잔차에 발생하는 평균 변화의 크기 ω에 바탕을 둔 새로운 참고값으로 식(14)을 대안으로 제시하고, 식(13)의 참고값과 효용성을 비교하기로 한다.

$$k_z = \frac{\sigma_z}{2}\sigma_a \qquad (14)$$

식(14)은 공정 평균 변화의 크기를 $\omega = c \cdot \sigma_z$이라고 할 때, $c=1$인 경우를 검출 대상의 평균 변화 크기로 선정한 것이다. 본 연구에서는 식(14)의 참고값(k_z)을 적용한 CUSUM 관리도와 식(13)의 참고값(k_∞)을 적용한 CUSUM 관리도에 대하여 모의실험을 통해 잔차의 평균 변화에 따른 CUSUM 관리도의 성능에 차이가 있는지에 대해 살펴보기로 한다. 특히, Runger et al.가 제안한 식(13)의 참고값이 AR(1)이 아닌 다른 모형에서도 타당한 것인지를 분석한다.새로운 참고값 제안의 타당성에 대해 논하고자 한다.

V. Simulation

모의실험은 모수가 (ϕ, θ)인 ARMA(1, 1) 과정을 따르는 공정 자료 Z_t에 대해 모수(ϕ, θ)를 달리하여 자료를 생성하고, Z_t의 표준편차를 계산하여 식(14)의 k_z와 식(13)의 k_∞등의 참고값을 결정한다. 그리고 ARL_0가 370이 되는 관리한계(h)를 결정하고 공정 평균의 변화 크기에 따라 ARL_1이 어떻게 나타나는지를 비교하여 CUSUM 관리도의 성능을 비교하였다. 이 때, 얻어지는 CUSUM의 통계량은 다음과 같다.

$$S_t^+ = \max[0, S_{t-1}^+ + e_t - k]$$
$$S_t^- = \max[0, S_{t-1}^- - e_t - k]$$

Table 1. $(\phi = 0.4, \theta = 0.7)$ in ARMA(1,1), ARL_1 using in k_z and k_∞

(ϕ, θ)	k_z	h_z	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.4$ $\theta = 0.7$	0.4554 0.5050 0.5565	5.022 4.644 4.303	12.3 12.2 12.6	5.9 5.7 5.7	4.1 3.9 3.9	3.2 3.1 3.0	2.3 2.2 2.1
(ϕ, θ)	k_∞	h_∞	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.4$ $\theta = 0.7$	0.9818 1.0318 1.0818	2.565 2.439 2.320	16.2 17.1 17.8	5.6 5.6 5.8	3.5 3.4 3.5	2.6 2.5 2.5	1.7 1.7 1.7

여기서, e_t는 시점 t에서의 잔차이며 $S_0^+ = 0$이고 $S_0^- = 0$으로 설정한다.

모의실험은 각각 5000번 수행하였으며, h의 산출은 ARL_0가 370이 되도록 Hawkins & Owell에서 제공한 프로그램을 이용하여 결정하였다[12].

1. Simulation results

ARMA(1, 1) 과정을 따르는 공정 자료에 평균 변화가 발생한 경우에, 도출된 잔차의 평균이 크게 세 가지 유형으로 변하는 것을 알 수 있다. 각 유형의 대표적인 모형에 대하여 식(13)과 식(14)의 참고값을 결정하고 원 공정의 평균이 $0.5\sigma_z$에서 $3.0\sigma_z$까지 다양하게 변하는 경우에 대하여 모의실험을 통해 ARL_1을 구하고 CUSUM 관리도의 성능을 비교하고자 한다.

(1) 잔차의 평균이 증가하는 형태($\theta > 0, |\phi| < \theta$)

잔차의 평균이 증가하는 경우에는 <표1>과 같이 관리상태에서의 공정 평균(μ_0)에서 1.0 이하의 공정 평균 변화($\mu_1 \le \mu_0 + 1\sigma_z$)가 발생했을 때에는 k_∞를 이용한 CUSUM 관리도의 ARL_1이 k_z를 이용한 경우보다 작게 나타났다. 예를 들어, $0.5\sigma_z$의 평균 변화가 발생한 경우에 k_∞를 이용한 ARL_1은 17.1이지만 k_z를 이용한 ARL_1은 12.2로 작게 나타났다. 즉, 잔차의 평균이 증가하는 형태에서 작은 공정 변화에 대해 관심이 있는 경우에는 (k_z, h_z)의 CUSUM 관리도 설계모수를 사용하는 것이 (k_∞, h_∞)의 경우보다 민감도를 좀 더 높일 수 있음을 확인할 수 있다.

(2) 잔차의 평균이 감소하는 형태($\theta > 0, |\phi| > \theta$)

잔차의 평균이 감소하는 형태인 경우에는 <표2>와 같이 공정에 $1.5\sigma_z$ 이상의 평균 변화가 발생했을 때, k_z를 이용한 CUSUM 관리도가 k_∞를 이용한 것보다 ARL_1이 작게 나타났다. 예를 들어, $2.0\sigma_z$ 의 평균 변화가 발생한 경우, k_z를 이용한 ARL_1이 5.7인데 비해서 k_∞를 ARL_1이용한 경우에는 7.1로 나타났다. 따라서 잔차의 평균이 감소하는 형태에서는 작은 공정 평균의 변화에 대해 관심이 있는 경우에는 k_∞를 이용한 CUSUM 관리도 설계를, 큰 공정 평균의 변화에 대해 관심이 있는 경우에는 k_z를 이용하여 CUSUM 관리도를 설계하는 것이 좋은 것으로 나타났다.

(3) 잔차의 평균이 진동하는 형태($\theta < 0$)

잔차의 평균이 진동하는 형태에서는 공정 자료의 평균이 $2.0\sigma_z$ 이상으로 비교적 큰 폭의 공정평균이 변하는 경우에는 k_z를 이용한 CUSUM 관리도의 ARL_1이 낮게 나타났다.

하지만 $1.0\sigma_z$ 이하의 작은 공정 평균의 변화에서는 상대적으로 k_z를 이 용한 CUSUM 관리도의 ARL_1 값은 <표3>에서 보는 것 같이 성능이 나쁘게 나타났다.

2. Performance comparison with other control charts

앞의 모의실험을 통해, ARMA(1,1) 과정을 따르는 공정에 대하여 잔차를 CUSUM 관리도에 적용하는 경우에는 ARMA(1,1) 과정의 모수(ϕ, θ)와 관심 있는 평균 변화의 크기(δ)를 고려하여 참고값을 결정하는 것이 바람직하다는 것을 알 수 있었다. <표4>에서 보는 것 공정관리에 따른 참고값 관리도별 차이를 보여준다

CUSUM 관리도를 이용한 결과가 공정 평균의 변화에 관계없이 다른 관리도들 보다 ARL_1이 작게 나온 것을 알

Table 2. $(\phi = 0.7, \theta = 0.3)$ in ARMA(1,1), ARL_1 using in $k_z$and k_∞

(ϕ, θ)	k_z	h_z	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.7$ $\theta = 0.3$	0.4840 0.5340 0.5840	4.903 4.518 4.185	87.2 92.8 97.5	22.4 23.5 24.8	10.1 10.1 10.2	5.8 5.7 5.6	3.0 2.9 2.7
(ϕ, θ)	k_∞	h_∞	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.7$ $\theta = 0.3$	0.2551 0.3051 0.3551	7.904 6.981 6.249	61.1 66.0 70.0	20.7 20.2 20.3	11.6 10.9 10.4	7.7 7.1 6.6	4.5 4.0 3.6

Table 3. $(\phi = 0.6, \theta = -0.8)$ in ARMA(1,1), ARL_1 using in $k_z$and k_∞

(ϕ, θ)	k_z	h_z	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.6$ $\theta = -0.8$	0.9578 1.0078 1.0578	2.629 2.496 2.374	209.9 221.9 219.6	72.0 76.6 80.4	20.2 21.4 22.2	4.3 4.4 4.3	1.0 1.0 1.0
(ϕ, θ)	k_∞	h_∞	ARL_1 for δ in $w = \delta \cdot \sigma_z$				
			0.5	1.0	1.5	2.0	3.0
$\phi = 0.6$ $\theta = -0.8$	0.1740 0.2240 0.2740	10.062 8.608 7.526	88.5 93.2 100.2	32.8 32.3 32.6	18.3 17.1 16.5	12.5 11.2 10.2	6.9 5.8 4.8

Table 4. Control chart performance comparison

$\phi = -0.45$ $\theta = 0.44$ / Shift	SCC	X	EWMA	CUSUM (σ_∞)	CUSUM (σ_z)
0.0	370	392	392	369.	366.
0.5	8.7	152.2	5.9	3.0	3.02
1.0	2.6	44.1	2.9	3.0	2.72
1.5	1.8	45.2	2.0	1.12	1.18
2.0	1.5	6.2	1.8	1.00	1.00
2.5	1.2	2.6	1.3	1.00	1.00
3.0	1.1	1.4	1.2	1.00	1.00

수 있다. 따라서, 위 모형의 경우에서는 자기상관이 존재하는 공정의 품질 특성치를 시계열 모형에 적합하여 구한 잔차를 모니터링 하는 방법 중에서 CUSUM 관리도를 적용하는 것이 가장 민감도가 높음을 확인할 수 있다.

VI. Conclusions

과거 Walter A. Shewhart에 의해 제안된 관리도는 자료가 서로 독립인 공정에 적합한 관리 방법으로써 당시 산업 구조가 부품 단위의 생산에 머무는 산업(part industry)에 적합한 방법이라고 할 수 있다. 반면에 현재의 산업은 반도체와 같이 장치에 의한 프로세스에 의존하는 산업(process industry)으로써 산업의 구조가 과거와는 사뭇 달라졌다. 이러한 프로세스 중심의 산업에서 발생하는 공정 자료는 자료들 사이에 독립성 보장되지 못하고 자기상관이 존재하는 형태의 종속 구조를 내재하고 있다. 자기상관이 존재하는 공정 자료에 대한 관리방법으로는 관측치를 대상으로 한 관리 방법과 관측치를 시계열 모형에 적합하여 얻은 잔차를 관리도에 적용하여 공정을 모니터링 하는 두 가지의 방법이 소개되었는데, 이 두 가지 접근 방법은 서로 간에 많은 차이를 보이고 있다.

만일 모형 추정이 올바르게 될 수 있는 경우라면 추정된 모형으로부터 구한 잔차를 CUSUM 관리도에 적용하여 작은 공정의 평균 변화를 검출하는 것이 타당한 것으로 판단된다. 잔차에 대해 CUSUM 관리도를 적용하는 경우에 있어서는 공정 자료가 갖는 시계열 모형의 형태와 관심 있는 공정 평균 변화의 폭을 고려하여 CUSUM 관리도의 설계 과정에서 필요한 참고값이 적절히 선택되어 사용되어야 함을 모의실험을 통해 확인하였다.

REFERENCES

[1] Layth C. Alwan and Harry V. Roberts, , “Time-Series Modeling for Statistical Process Control,” Journal of Business & Economic Statistics, Vol. 6, pp 87-95 January 1988

[2] George C. Runger, and Thomas R. Willemain. “Batch Mean Control Charts for Autocorrelated Data,” IIE Transactions, Vol. 28, No. 6, pp. 483-487 June 1996

[3] Chung. Chen and Lon-Mu. Liu “Joint Estimation of Model Parameters and Outlier Effects in Time Series,” Journal of the American Statistical Association, vol 88, pp284-297 March 1993

[4] O. O. Atienza,. L. C. Tang and Ang, B. W. Ang “A SPC Procedure for Detecting Level Shifts of Autocorrelated Processes,” Journal of Quality Technology, Vol. 30, No. 4, pp 340-351 Octorber 1998

[5] Chao-Wen. Lu and Reynolds, Marion R. Reynolds “EWMA Control Charts for Monitoring the Mean of Autocorrelated Processes,” Journal of Quality Technology, vol 31 issue 2 pp

166-188 April 1999

[6] Chao-Wen. Lu and Marion R. Reynolds Jr "CUSUM Charts for Monitoring An Autocorrelated Process," Journal of Quality Technology, Vol. 33, No. 3, pp. 316-334 July 2001

[7] Lim, Tae Jin "A Selectivety Cumulative Sum chartchart" Journal of the Korean Society for Quality management, Vol. 33, No. 3, pp. 126-134 2005

[8] Kim, J.G and Um .S,J, Choi, S.W "Resarch Results and trends analysis on CUSUM control chart" https:// www.earticle.net/Article/153737 2010

[9] S. Jack. Hu and Chinmo. Roan "Change Patterns of Time Series-Based Control Charts", Journal of Quality Technology, Vol. 28, No. 3, pp. 302-312 July 1996

[10] Jones, M.A and Steiner,S.H "Assessing the effect of estimation error on the risk-adjusted CUSUM chart performance" Internatinal Journal for Quality in Health Care, vol 24, issue 2 pp176-181 April 2012

[11] Montgomery, Douglas C. "The Use of Statistical Process Control and Design of Experiments in Product and Process Improvements," IIE Transactions, Vol. 24, issue 5 pp. 4-17 May 1992

[12] Montgomery, Douglas C. Statistical quality control: A modern introduction, 7th Ed,., John Wiley & Sons, Inc., Hoboken .NJ 2013

[13] Runger, George C, Willemain, Thomas R. and Prabhu, Sharad "Average Run Length for CUSUM Control Charts Applied to Residuals", Communications Statistics - Theory and Method, Vol. 24, No. 1, pp. 273-282 March 1995

[14] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. "Time Series Analysis, Forecasting and Control, 3rd Ed." Prentice Hall, Englewood Gliffs, NJ 1994

Authors

Sang-Pyo Jun received the B.S,M,S. and Ph.D degree in Statistics from Inha University, korea, in 1985,1987 and 2000, respectively. Dr. Jun joined the faculty of the collage of General Education at Namseoul Unversity, Chonen Korea, in 2005. He is currenty a Professor in the collage of General Education at Namseoul Unversity, Chonen Korea. He is interests include linar modeling, deep learning, SPC and statistics