

## A Study on UBM Method Detecting Mean Shift in Autocorrelated Process Control

Sang-Pyo Jun\*

\*Professor, College of General Education, Namseoul University, Cheonan, Korea

### [Abstract]

In today's process-oriented industries, such as semiconductor and petrochemical processes, autocorrelation exists between observed data. As a management method for the process where autocorrelation exists, a method of using the observations is to construct a batch so that the batch mean approaches to independence, or to apply the EWMA (Exponentially Weighted Moving Average) statistic of the observed value to the EWMA control chart.

In this paper, we propose a method to determine the batch size of UBM (Unweighted Batch Mean), which is commonly used as a management method for observations, and a method to determine the optimal batch size based on ARL (Average Run Length) We propose a method to estimate the standard deviation of the process. We propose an improved control chart for processes in which autocorrelation exists.

▶ **Key words:** Autocorrelation, CUSUM control chart, EWMA control chart, Unweighted Batch Mean, Average Run Length

### [요 약]

오늘날 반도체나 석유 화학 공정과 같이 프로세스 중심의 산업에서는 관측된 자료들 사이에 자기 상관관계가 존재한다. 자기상관계가 존재하는 공정에 대한 관리 방법으로는 관측치를 이용하여 배치 평균이 독립에 가까워지도록 배치를 구성하여 관리하거나, 관측치의 EWMA (지수 가중치 이동 평균) 통계량을 EWMA 관리도에 적용하는 방법등이 주로 사용되고 있다.

본 논문에서는 관찰치에 대한 관리 방법으로 일반적으로 사용되는 UBM의 배치 크기를 결정하는 방법을 소개하고, ARL(평균 실행 길이)을 기반으로 최적 배치 크기를 정하는 방법과 그러한 배치 구성에서 공정의 표준 편차를 추정하는 방법을 제안 한다. 자기상관계가 존재하는 공정에 대한 개선된 관리도를 제안하고자 한다.

▶ **주제어:** 자기상관, CUSUM 관리도, EWMA 관리도, UBM, ARL

## I. Introduction

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 검출 되지 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 발생하는 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 “슈하르트 관리도” 혹은 “3-시그마 관리도” 라고 불리는데 공정에서 관측된 자료들이 서로 독립인 가정에서 적용될 수 있다. 하지만 자기상관이 존재하는 공정 자료에 대해 슈하르트 관리도는 잘못된 경보가 발생할 가능성이 높다. 이러한 자기상관이 존재하는 공정 자료에 대해 슈하르트 관리도를 적용하는 방법에 대해 다양한 연구가 되어 왔다[1-5].

이와 같이 자기상관이 존재하는 공정에 대한 공정관리 방법들의 성능에 어떠한 차이가 있는지를 비교하고, 현장 엔지니어들이 모형에 관계없이 쉽게 사용할 수 있는 UBM(unweighted batch mean)을 이용한 공정관리 방법에서 필요한 뱃치의 크기와 이를 이용한 공정의 표준편차 추정 방법을 제안하고자 한다.

## II. Process control of autocorrelation data

### 1. Process control using observations

자기상관이 존재하는 공정에 대하여 관측치를 직접 이용하여 공정을 관리하는 방법으로는 UBM을 이용한 방법, EWMA를 이용한 방법, 누적합(CUSUM)을 이용한 방법 등이 있다.

### 1.1 Method using UBM

Runger & Willemain에 의해 제안된 이 방법은 공정 자료가 어떤 시계열 모형을 따르는지와 무관하게 시차(lag)가 커짐에 따라 뱃치 평균들 사이에 자기상관이 줄어드는 성질을 이용한 방법이다[2]. 따라서 자기상관이 큰 공정에서는 뱃치 평균들 사이의 상관관계가 거의 0에 가깝도록 뱃치의 크기를 충분히 크게 설정해야 하기 때문에 자료가 충분히 많이 주어지는 경우에 적용될 수 있는 방법으로 다음과 같이 표현될 수 있다.

$$\bar{x} = \frac{1}{b} \sum_{i=1}^b x_{b(j-1)+i}, \quad j = 1, 2, \dots, \quad (1)$$

UBM을 이용한 방법에서 중요한 문제는 자기상관이 없어질 정도의 적절한 뱃치 크기( $b$ )를 구하는 것이다. Kang & Schmeiser는 UBM 간에 시차가 1인 자기상관계수,  $\rho_1 (= Cov[\bar{x}_t, \bar{x}_{t-1}] / V[\bar{x}_t])$ 가 0.1 이하가 만족될 때, 뱃치 평균이 거의 독립에 가깝다고 가정하고, AR(1) 과정의 다양한 모수값  $\phi$ 에 대해 뱃치 평균 사이의 1차 자기상관계수가 처음으로 0.1 이하가 되는 최소의 뱃치 크기를 표 1와 같이 제시했다[6].

UBM에 대한 관리도는 뱃치 평균의 기대값( $\mu_x$ )과 뱃치 평균들의 표준편차( $\sigma_x$ )를 이하에 위아래의 관리한계는 식 (2)와 같이 표현된다.

$$\begin{aligned} UCL &= \mu_x + 3\sigma_x \\ LCL &= \mu_x - 3\sigma_x \end{aligned} \quad (2)$$

UBM을 이용한 방법에서 시차가 1인 자기상관계수( $\rho_1$ )가 0.1 이하가 되는 최소의 뱃치 크기로 UBM을 설정하면 자기상관이 큰 공정의 경우, 뱃치 내의 자료 수가 너무 많아질 수 있다. 이런 경우에 이상원인(assignable causes)에 의해 공정변화가 발생하더라도 이를 검출하는 데까지 시간이 너무 많이 소요되는 효율성의 문제가 발생할 수 있다. 반면에 뱃치 크기가 너무 작게 설정될 때에는 작은 공정의 변화에 대해 검출력이 떨어지는 단점이 발생한다

### 1.2 Method using EWMA

Montgomery는 관측치  $X_t$ 의 EWMA(Exponentially Weighted Moving Average) 통계량을 이용한 공정 평균에 대한 공정관리 방법을 제안하였다[7]. 공정의 품질 특성치  $X_t$ 의 EWMA  $Z_t$ 는 다음식 (3)과 같이 표현되는데,

Table 1. UBM minimum batch size for AR(1)

$\phi$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
$b$	1	2	3	4	6	8	12	17	27	58	118	596

$$Z_t = \lambda X_t + (1 - \lambda)Z_{t-1} \quad t = 0, 1, 2, \dots \quad (3)$$

여기서  $0 < \lambda \leq 1$ 로서  $Z_t$ 의 계산에 적용되는 과거 관측치의 가중치(weight)를 의미한다. ( $t$ 는 기간)

자료가 독립인 경우에 EWMA관리도에 의한 관리 한계는 식(4)와 같이 가중치  $\lambda$ 에 의존하여 표준편차의 관련식으로 결정된다.

$$\begin{cases} UCL = \mu_0 + L\sigma \sqrt{\lambda/(1-\lambda)} \\ LCL = \mu_0 - L\sigma \sqrt{\lambda/(1-\lambda)} \end{cases} \quad (4)$$

공정이 ARIMA(0,1,1), 즉 IMA(1,1) 과정을 따르는 경우에 EWMA는 MMSEF(Minimum Mean Square Error Forecast)가 되고, 따라서 그로부터 구한 잔차는 평균이 0이고 표준편차가  $\sigma$ 인 서로 독립인 백색 잡음 과정(white noise process)이 된다. 이러한 특성을 반영하여 일반적인 슈하르트 관리도의 적용이 가능하게 된다. 하지만 IMA(1,1) 과정을 따르지 않는 공정 자료들에 대해서는 가중치를 적당히 잘 선택하면 잔차가 독립에 근사할 것으로 기대되어 자기상관이 존재하는 공정 자료에 대해서도 적용할 수 있다고 제안되었다. 다만, 관리한계는 다음과 같은  $(1 - \alpha)$ 의 확률한계(probability limits)를 이용하여 식(5)과 같이 설정한다.

$$P_r[-Z_{\alpha/2}\sigma \leq e_t \leq Z_{\alpha/2}\sigma] = 1 - \alpha \quad (5)$$

여기서,  $Z_{\alpha/2}$ 는 표준정규분포의 상위에 해당하는 값이다.

식(5)을 잔차( $e_t$ ) 아닌 관측치( $X_t$ )에 의한 확률한계로 바꾸면 다음과 같이 관리한계가 유도 된다.

$$\begin{cases} UCL = \hat{X}_{t-1}(1) + Z_{\alpha/2}\sigma_e \\ LCL = \hat{X}_{t-1}(1) - Z_{\alpha/2}\sigma_e \end{cases} \quad (6)$$

여기서,  $\hat{X}_{t-1}(1)$ 는 시점  $t-1$ 에서 구한 1-시점 후의 예측 값으로서, EWMA 관리도에서는 식(3)의 EWMA 통계량인  $Z_t$ 를 사용한다.

EWMA 관리도를 잔차에 적용하는 경우, 잔차가 평균 0,

표준편차가  $\sigma_e$ 인 서로 독립을 만족하는 과정이라 가정하고, 관리한계는 식(4)로부터 식(7)과 같이 표현된다.

$$\begin{cases} UCL = L\sigma_e + L\sigma \sqrt{\lambda/(1-\lambda)} \\ LCL = L\sigma_e - L\sigma \sqrt{\lambda/(1-\lambda)} \end{cases} \quad (7)$$

EWMA 통계량을 이용한 방법에서는 식(3)의  $\lambda$ 을 선택하는 것이 중요하다. Montgomery는 자기상관이 존재하는 공정 자료에 대한  $\lambda$ 의 선택 방법으로 1-시점 후 예측 오차의 제곱합(1-step ahead MSE)이 최소가 되는  $\lambda$ 을 선택하도록 제안하였다. 대개 자기상관이 존재하는 공정 자료에 대해 EWMA 관리도를 적용하는 경우에  $\lambda$ 는 0.1에서 0.2 사이에서 정하는 것이 바람직한 것으로 권고되고 있다. 따라서 가장 최근의 관측치  $X_t$ 에 0.1에서 0.2 사이의 가중치가 부여되고, 과거 관측치에 훨씬 큰 비중이 주어지기 때문에, EWMA 관리도를 이용한 관리방법은 공정 평균이 급격하게 변하는 경우에 그 변화를 조기에 반영하는 정도가 작아서 그러한 변화를 민감하게 반응하지 못하는 단점이 있다.

### 1.3 Method using cumulative sum (CUSUM)

CUSUM 관리도를 이용한 자기상관 자료의 모니터링 방법에 대한 최근 연구는 Runger et al. 그리고 Lu & Reynolds 등이 있다. 개별 관측치에 적용하는 X-관리도를 포함한 일반적인 슈하르트 관리도는 작은 공정의 변화 변화에는 검출력이 낮다는 단점이 있기 때문에 이를 보완하기 위한 목적으로 CUSUM 관리도를 사용하는 것이 제안되었다.

일반적으로 누적합관리도는 식(8)과 같이 모평균( $\mu_0$ )으로 조정된 관측치( $x_t - \mu_0$ )에서 그 크기가 참고값( $k$ )을 초과하는 변화량을 누적하는 방식으로 공정의 평균 변화를 모니터링 하는 것이다. 누적합관리도에서는 공정 평균이 증가하는 것과 감소하는 두 가지 경우를 모두 검출하기 위해 식(8)과 같이 각각  $S_t^+$ ( $t$  시점까지의 증가 누적합)와  $S_t^-$ ( $t$ 시점까지의 감소누적합) 두가지 통계량을 사용한다.

$$\begin{aligned} S_t^+ &= \max\left[0, S_{t-1}^+ + (x_t - \mu_0 - k)\right] \\ S_t^- &= \max\left[0, S_{t-1}^- - (x_t - \mu_0 - k)\right] \end{aligned} \quad (8)$$

누적합관리도 설계 시에 고려해야 할 중요한 두 가지 요소는 참고값( $k$ )과 관리한계( $h$ )를 결정하는 것이다. 자료가 서로 독립인 경우에 검출하고자 하는 공정의 평균 변화 크기를 표준편차의 배수로 표현하게  $\omega = |\mu_1 - \mu_2|/\sigma$  되면 이 된다. 여기에서, 변경된 공정 평균은  $\mu_1 = \mu_0 + \omega\sigma$ 로서 참고값은 평균 변화 크기의 절반인  $k = |\mu_1 - \mu_2|/2$ 로 설정됨을 알 수 있다. 따라서 참고값은  $k = (\omega/2)\sigma$ 로 설정된다.

이러한  $k$ 에 대해 관리한계  $h$ 는 미리 정한 관리상태 평균런길이 ( $ARL_0$ )의 수준에 따라 결정 된다.

#### 1.4 Process control performance comparison

자료가 서로 독립인 경우, 슈하르트 관리도를 적용할 때 발생하는 제1종오류( $\alpha$ )를 고려한 관리상태 평균-런-길이(In-control Average Run Length:  $ARL_0$ )는 다음과 같이 정의된다.

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

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

만약 관리도에 적용되는 자료가 정규분포를 따른다고 가정하고, 관리한계의 폭을  $\pm 3\sigma_x$ 로 잡는다면  $\alpha = 0.0027$ 이 되므로  $ARL_0$ 는 약 370 정도가 된다. 즉, 공정에 아무런 이상원인에 의한 영향이 없는 경우에도 평균적으로 370번 마다 한 번쯤은 관리한계를 벗어나는 관리이탈(out-of-control)현상이 발생한다는 것이다.

$ARL_0$ 는 여러 가지 관리도의 성능을 비교해 각 관리도의 관리한계를 공정 평균의 변화가 없는 경우에 대하여 가정하여  $ARL_0$ 가 모두 같도록 설정하는데 적용될 수 있는데, 슈하르트 관리도의 경우를 고려하여 일반적으로 약 370 정도가 되도록 한다.

관리도의 성능은 공정에 이상원인이 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  $ARL_1$ 은 다음과 같이 정의된다.

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

여기서,  $\beta = \Pr[LCL < Z_t < UCL | \mu(= \mu_0 + \delta\sigma)]$ 로 나타나는 제 2종 오류. 즉, 공정 평균이  $\mu_0$ 에서  $\mu_1(= \mu_0 + \delta\sigma, \delta \neq 0)$ 으로 변경되었을 때, 관리한계를 이탈하는 데 까지 소요되는 런의 길이를 의미한다. 만약,  $1-\beta$ 가 1에 가까워 진다면  $ARL_1$ 은 1에 가까운 값이 나오게 되는데, 이것은 공정에 이상원인이 발생하게 되면 평균적으로 그 다음 시점에서 관리도에 관리이탈로 검출된다는 것이다[8, 9].

### III. UBM control chart design

#### 1. Proposal background

자기상관이 존재하는 공정 자료에 대한 관리 방법으로 관측치를 관리하는 방법과 잔차를 관리하는 방법인 UBM은 공정 자료의 모형에 무관하게 사용할 수 있다. UBM을 이용한 방법에서는 UBM 상호 간의 자기상관이 거의 0에 가깝도록 배치 크기를 결정하는 것이 필요하다. 또 다른 문제로서, UBM 자료를 적용한 관리도의 관리한계 결정에 필수적인 공정의 표준편차  $\sigma_x$ 의 추정이 필요하다. 관측 자료의 수가 충분치 못하여 시계열 모형의 추정이 어려운 경우에 대해 UBM을 이용한 공정 모니터링이 적용 가능할 수 있게 하는 배치 크기 및 공정 표준편차 추정 방법을 제안하고, 이를 통해 ARL을 기반으로 하는 최적의 배치 크기의 개선된 관리기법을 제안하고자 한다.

##### 1.1 UBM control chart design

자기상관이 존재하는 공정자료를 슈하르트 관리도에 적용하는 경우에 UBM들 간에 독립이 만족될 수 있도록 배치를 구성하는 것이 중요하다.

이렇게 제안된 배치 크기를 적용하여 식 과 같이 관리한계를 설정하여 UBM에 대한 슈하르트 관리도를 적용하게 된다.

$$UCL = \mu_0 + 3\frac{\sigma_x}{\sqrt{b}}$$

$$LCL = \mu_0 - 3\frac{\sigma_x}{\sqrt{b}}$$

여기서,  $b$ 는 배치의 크기이다.

식에서 모수를 모르는 경우에  $\hat{\mu}_0 = \bar{\bar{x}}$ ,  $\hat{\sigma}_x = \bar{R}/d_2$ , 로 추정한다.  $\sigma_x$  추정에 사용되는 는 인접한 두 UBM을 하나의 부분군(subgroup)으로 구성한 후에 부분군 내의 범위, ( $R_i = |\bar{x}_{2i-1} - \bar{x}_2|$ ,  $i = 1, 2, \dots, k$ )의 평균을 구한다.

Runger & Willemain이 제시한 표2 의 뱃치 크기가 타당한 지를 검토하기 위하여 식에서 제시한 관리한계로부터 산출된  $ARL_0$ 가 370을 만족하는지 알아보면 다음과 같다.

Table 2. Minmum batch size for  $ARL_0$

$\phi$	뱃치 크기( $b$ )	$ARL_0$
0.2	3	333.1
0.3	4	338.0
0.4	6	334.9
0.5	8	339.6
0.6	11	333.4
0.7	17	337.3
0.8	26	333.9

표 2.에서 보여지듯이 Runger & Willemain가 제시한 최소 뱃치 크기를 이용한  $\bar{X}$ 관리도를 적용한 결과,  $ARL_0$ 는 전반적으로 370보다는 작게 나타났다. 따라서 부분군의 수가 10개 이하인 경우에는 공정에 이상원인이 발생하지 않았음에도 불구하고 관리이탈 오류 경보(false alarm)가 기준보다 많이 발생하게 된다.  $ARL_0$ 가 370을 만족시킬 수 있는 뱃치의 크기( $b$ )를 어느 정도로 결정해야 하는지에 대해 제안하고자 한다.

2. Determination of the optimal batch size

UBM 간의 독립성을 확보할 수 있도록  $ARL_0$ 의 값이 370을 가질 수 있도록 뱃치의 크기를 결정하기 위한 모의 실험 방법을 요약하면 다음과 같다. 모의실험은 Visual Basic을 이용하여 프로그래밍하였다.

STEP 1. 일반적으로 사용하는 몬테칼로 모의실험 방법을 이용하여 모수가  $\phi$ 인 AR(1) 과정을 따르는 자료를 생성하여 UBM을 구성하고, UBM의  $\rho_1$ 이 0.1 이하가 되는 뱃치 크기를 구한다. 이러한 모의실험을 5,000회 실시하여 히스토그램으로 뱃치 크기의 분포를 구한다.

STEP 2. 위 STEP 1의 과정을 모수 가 0.2로부터 0.9까지 0.1씩 증가시키는 방식으로 각 모수값에 대해 뱃치 크기의 분포를 구한다.

이러한 모의실험의 결과를 요약한 표들을 다음과 같다

표2에서 언급한 최소 뱃치 크기는 표3의 모의실험을 통해 얻은 뱃치 크기의 통계량을 통해 보면 평균 임을 알 수 있다. 또한 이것은 표3과 표4로부터 확인할 수 있듯이 모의실험 결과를 통해 얻은 뱃치 크기의 분포에서 약 75백분위수에 해당하는 뱃치 크기 임을 알 수 있다. 하지만 표4와 표5에서 지적한 것처럼 전체의 75%가 포함되는, 평균적으로  $\rho_1$ 이 0.1이하가 되는 뱃치 크기 즉, 약 75백분위수

에 해당하는 뱃치 크기를 최적의 뱃치 크기로 사용하게 되면  $ARL_0$ 가 370보다 작게 나타나게 되므로 최적의 뱃치 크기를 결정하기 위해서는 새로운 기준이 필요하다.

모의실험 결과로 제시된 표5 에서 확인할 수 있듯이, 뱃치의 분포에서 75백분위수의  $ARL_0$ 는 370보다 전반적으로 작기 때문에 뱃치 크기 분포의 90백분위수와 95백분위수의 효용성을 검토해 보기로 한다.

표5는 각 AR(1)의 모형에 대하여 모의실험으로부터 측정된 최초 뱃치 크기의 90백분위수와 95백분위수 크기값과 그 뱃치 크기에서의  $ARL_0$ 를 계산한 결과이다.

표5의 결과를 보면 모의실험으로 결정된 뱃치 크기가 표1 에 제시된 Runger & Willemain의 최소 뱃치 크기로 추정됨을 알 수 있다[2]. 구체적으로 살펴보면, 뱃치 크기를 작게 잡은 상태에서 관리한계를 산출하게 되면, 뱃치 평균들 사이에 자기상관이 완전히 사라지지 않아 표준편차가 과소 추정되어 실제보다 관리한계가 좁게 설정되는 문제가 발생할 수 있다. 따라서, 공정 평균의 변화에 대한 관리도의 성능 비교 시, UBM을 적용한 관리도는 iid공정에서 기대되는 것보다  $ARL_1$ 이 작게 나오는 현상이 발생되고, 그로 인해 성능이 과대 평가될 가능성이 높다. 따라서 표5에서 보여지듯이 자기상관이 큰 공정 자료의 경우에는 Runger & Willemain이 제안한 최소 뱃치 크기에 비해 약 1.5배 큰 뱃치 크기를 갖도록 UBM을 구성하는 것이 바람직하다.

3. Estimation of process standard deviation using UBM

UBM을 슈하르트 관리도에 적용하기 위한 관리한계는 식2에 제시되어 있다. 여기서 공정 자료에 대한 표준편차( $\sigma_x$ )가 올바르게 추정되어야만 공정 변화를 검출할 수 있는 관리한계가 제대로 설정되게 된다. 특히, 관리도를 이용한 공정관리뿐 아니라 공정능력 평가를 위한 혹은  $C_p$ 와  $C_{pk}$  같은 공정능력지수를 구하는 경우에 있어서도 올바른 공정 표준편차( $\sigma_x$ )의 추정은 매우 중요하다.

Shore는 AR(1)과정을 따르는 공정 자료에 대하여 UBM의 분산과 공정의 분산( $\sigma_x^2$ ) 사이의 관계식을 이용하여 공정의 표준편차( $\sigma_x$ ) 추정 방법을 제시하였다[10]. 앞에서 구한 최적 뱃치 크기를 이용하여 공정 표준편차( $\sigma_x$ )를 추정하는 방법을 제시하고 추정치의 타당성에 대해 분석하고자 한다.

AR(1) 과정을 따르는 공정 자료에 대해, 뱃치 크기가  $b$ 인 UBM을  $m$ 개 구성하고 인접한 두 UBM을 하나의 부분군으로 구성하는 방식으로  $s(= m/2)$ 개의 부분군을 작성

Table 3. Statistics for batch sizes b

$\phi$	Statistic of $b$					
	Mean	SE(Mean)	SD	Min	Median	Max
0.2	3.1	0.022	1.312	2.0	3.0	15.0
0.3	4.3	0.023	1.823	2.0	4.0	26.0
0.4	6.0	0.041	2.525	2.0	6.0	51.0
0.5	8.2	0.080	3.211	3.0	8.0	51.0
0.6	11.5	0.120	4.021	3.0	10.0	52.0
0.7	15.4	0.221	5.118	7.0	14.0	52.0
0.8	24.0	0.345	8.602	10.0	21.0	123.0
0.9	49.0	0.633	16.03	20.0	44.0	139.0

Table 4. Batch size percentile of table3

percentage	$\phi$ Batch size							
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
75%	3	5	7	10	13	16	26	54
90%	4	6	8	12	16	20	35	73
95%	4	6	9	12	17	25	40	85
99%	5	7	10	16	22	36	63	120

Table 5. Batch size percentile

$\phi$	90 percentage		95 percentage	
	Batch size $b$	$ARL_0$	Batch size $b$	$ARL_0$
0.2	4	353.8	4	362.5
0.3	6	379.1	6	379.5
0.4	8	380.4	9	386.2
0.5	12	374.1	13	375.8
0.6	16	379.0	18	374.4
0.7	22	385.5	26	381.4
0.8	37	366.5	42	371.0

하면 UBM의 분산은 정규분포의 가정하에  $V(\bar{X}) = (\bar{R}/d_2)^2$ 가 된다. 이렇게 추정된 분산은 UBM에 대한 분산을 나타내는 식(9)와 같이 된다.

$$V(\bar{X}) = \left(\frac{\sigma_x^2}{b}\right) \left[1 + \frac{2}{b} \sum_{k=2}^{b-1} (b-k)\rho_k\right] \tag{9}$$

여기서,  $\rho_k$ 는 시차가  $k$ 인 자기상관계수이다. 식9 을  $\sigma_x$ 에 대해서 정리하면 식(10)과 같이 된다.

$$\sigma_x = \sqrt{\frac{b \text{Var}(\bar{X})}{1 + \frac{2}{b} \sum_{k=2}^{b-1} (b-k)\rho_k}} \tag{10}$$

만일 공정의 품질 특성치가 AR(1) 과정을 따른다면 이론적인 공정의 표준편차 ( $\sigma_x$ )는 식(11)와 같이 표현된다.

$$\sigma_x = \sqrt{\frac{\sigma_a^2}{1 - \phi^2}} \tag{11}$$

식(10)에 의하면 공정의 표준편차( $\sigma_x$ )의 추정에 중요한 요소로 식(10)에 정의된  $\sigma_x$ 의 식은 사용된  $\rho_k$ 를 어느 시차까지 사용하느냐가 중요한 문제가 된다. 식(10)에서는  $(b-k)$ 개의 시차를 사용하도록 되어 있으나, 자기상관의 정도가 커지면 뱃치 크기( $b$ )도 상대적으로 많이 늘어나기 때문에 실제로는 통계적 유의성이 보장되는 시차까지만 사용한다.

### 3.1 Estimated result

앞에서 언급한 것처럼 부분군의 수( $s$ )가 그리 많지 않은 ( $s \leq 10$ ) 경우에 대한  $\sigma_x$  추정치의 적절성을 검토하고자  $s = 6$ 과  $s = 20$ 의 두 부분군의 수를 대상으로 분석하였다.

AR(1) 모형의 경우,  $\phi$ 값이 0.8인 경우에 유의한 시차가 10정도인데 반해서, AR(1) 과정을 따르는 공정의 참표준

편차( $\sigma_x$ )와 가장 가까운 경우는  $s = 6$ 일 때,  $k = 29$ 이고,  $s = 10$ 일 때,  $k = 15$ 이다. 따라서, 부분군의 수에 따라서 공정의 표준편차 추정에 적용되는 시차가 달라지는 것을 알 수 있다. 부분군이 많지 않을 때에는 충분히 많은 시차를 이용하여 공정 표준편차( $\sigma_x$ )를 추정하는 것이 바람직하고, 부분군이 충분하다면 유의한 시차까지만 사용해도 될 것이다.

표본 데이터를 이용한 실제 공정 표준편차( $\sigma_x$ )의 추정 표준편차는 다음과 같다.

$$\hat{\sigma}_x^{aH} = \sqrt{\sum (x_i - \bar{x})^2 / (bs - 1)}$$

식에서  $\phi = 0.8$ 인 AR(1) 과정의 경우에  $s = 6$ 일 때,  $\hat{\sigma}_x^{aH} = 1.6302$ 이고,  $s = 10$ 일 때에는  $\hat{\sigma}_x^{aH} = 1.6427$ 로 나타났다. 이 값은 참값인 1.6667과 비교해서 더 많은 차이가 나는 것을 알 수 있다. 이는 부분군의 수( $s$ )가 적을 때는 전체 데이터를 이용한  $\hat{\sigma}_x^{aH}$ 이 부적합하다는 것을 보여준다.

앞의 두 가지 모의실험을 통해 UBM을 구성하는데 필요한 배치 크기( $b$ )와 공정 표준편차( $\sigma_x$ ) 추정에 필요한 시차를 알아보았다. 이와 같은 분산 추정은 특히, 제품 생산 수명이 짧거나 예비(Pilot) 생산과 같이 관리한계 설정을 위해 많은 자료를 수집할 수 없는 경우에 유용하다. 대개 iid자료에 대한 관리한계 설정 시에 20 ~ 25개의 부분군을 이용하여 관리한계를 설정하는 것이 제시되어 있지만 Montgomery, [7]), 자료의 수가 충분치 않은 자기상관 자료에 대해서는 UBM에 대한 식의 슈하르트 관리도를 적용하는 관리한계를 산출하는 것이 타당한 것으로 볼 수 없다 [11, 12, 13].

#### IV. Conclusion

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.

특히, 자료가 충분치 않은 경우에는 올바른 시계열 모형을 추정하기 어렵기 때문에 UBM을 구성하고, 이를 슈하르트 관리도에 적용하는 것이 효율적일 수 있다. 다만, UBM을 이용하는 경우에 있어서 배치의 크기는 UBM 간에 시차가 1인 자기상관 계수가 0.1 이하가 되도록 하여 UBM 간에 자기상관이 거의 소멸될 수 있도록 추정한 배치 크기의 분포에서 95백분위수에 해당하는 배치 크기로 설정하는 것이 올바른 관리한계를 설정하고 공정의 표준편차를 효율적으로 추정하는데 유용하다고 할 수 있다.

향후에는 이러한 공정관리 방법들이 공정 조정(process adjustment)이 이뤄지는 경우에도 효과적으로 적용될 수 있는지에 대한 연구가 필요하다.

#### 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 02 January 1988 Published online July 2012..
- [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 Published online May 2007
- [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, pp 284-297 March 1993 Published online Dec 2012
- [4] O. O. Atienza,. L. C. Tang and Ang, B. W. Ang (1998), "A SPC Procedure for Detecting Level Shifts of Autocorrelated Processes," Journal of Quality Technology, Vol. 30, No. 4, pp 340-351 October 1998 Published online Feb 2018
- [5] 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 May 1995
- [6] 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 Published online Feb 2018
- [7] Montgomery, Douglas C. Statistical quality control: A modern introduction, 7th Ed., John Wiley & Sons, Inc., Hoboken .NJ 2013

- [8] Sven Knoth “ Accurate ARL computation for EWMA-S2 control charts” *Statistics and computing*, 15, pp 341-352 September 2004
- [9] Sven Knoth “ Accurate ARL computation for EWMA control charts monitoring simultaneously normal mean and variance ” *Sequential Analysis* 26(3), pp 251-264 June 2007
- [10] Haim Shore (1997), “ Process Capability Analysis When Data are Autocorrelated,” *Quality Engineering*, 4(4) pp. 615-626 January 1997
- [11] 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 Published online: 21 Feb 2018
- [12] Min. Zhang, Yiming. Peng,, Anna. Schuh, Fadel. M Megahed and W..H. Woodall “ Geometric charts with estimated control limits” *Qualaity and Reliability Engineering International*, vol 29 pp 209-223 March 2013
- [13] Min. Zhang, Fadel. M Megahed and W..H. Woodall “ Exponential CUSUM charts with estimated control limits” *Qualaity and Reliability Engineering International*, vol 30 pp 275-286 March 2014

## 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

University, Chonen Korea, in 2005. He is currently a Professor in the collage of General Education at Namseoul University, Chonen Korea. He is interests include linar modeling, deep learning, SPC and statistics