

분산성분모형 관리도의 설계와 효율[†]

조찬양¹ · 박창순²

^{1,2}중앙대학교 통계학과

접수 2017년 7월 5일, 수정 2017년 8월 22일, 게재확정 2017년 8월 31일

요약

단순확률모형을 고려하는 표준관리도에서는 표본간 분산을 고려하지 않고 공정분산을 추정한다. 표본간 분산이 존재하는 경우에는, 공정분산이 과소추정된다. 공정분산이 과소추정되면 좁아진 관리한계로 인해 관리도의 민감도는 향상되지만 과도한 오경보율을 발생시킨다. 이 논문에서는 공정모형으로 분산성분모형, 즉 변동의 원인을 표본내 분산과 표본간 분산으로 구분하는 확률모형을 고려한다. 관리한계는 표본내 분산과 표본간 분산을 모두 사용하여 설정하고 그에 따른 평균런길이를 통하여 효율을 살펴 보았다. 관리형태는 가장 널리 사용되는 \bar{X} , EWMA, CUSUM 관리도를 고려하였다. 관리한계 설정에서 표본내 분산만을 사용한 경우 (Case I)와 표본간 분산도 함께 사용한 경우 (Case II)를 통해 관리도의 효율을 비교하였다. 또한, 공정 모수가 주어진 경우와 추정된 두 경우에 대해서도 관리도의 효율을 비교하였다. 그 결과, 표본간 분산이 증가할 때 Case I의 오경보율은 급격히 증가한 반면 Case II의 경우에는 동일하게 유지됨을 알 수 있었다.

주요용어: 단순확률모형, 오경보율, 평균런길이, 표본간 분산, 표본내 분산.

1. 서론

관리도는 통계적 공정관리를 위한 수단으로서 공정품질의 상태를 판단하는 데 사용된다. 공정에서는 관리도를 적용할 때 시간의 흐름에 따라 표본을 추출하고 추출된 표본으로 공정의 상태를 판단하는데, 일반적으로 모든 표본집단이 같은 분포를 따른다고 가정한다. 이러한 가정을 따르는 관리도를 표준관리도 (standard control chart)라 한다. 표준 관리도에서는 표본내 분산 (within-sample variance)만을 고려하여 관리한계, 평균런길이 등을 구하지만 이 경우 표본간 분산 (between-sample variance)을 무시하게 된다. 이 때 표본간 분산이 현저히 작은 경우 문제가 발생하지 않으나, 그렇지 않은 경우 공정분산이 과소추정 (under estimate)되어 관리한계가 좁아지므로 평균런길이가 작아지게 된다. 이와 같은 경우 관리도의 민감도는 향상되지만 오경보가 과도하게 발생되므로 바람직하지 못한 현상이라 할 수 있다.

따라서 변동의 원인이 표본내 뿐만 아니라 표본간에도 발생하는 확률모형, 즉 분산성분모형에 대한 연구가 되어져 왔다. Park (1997)은 분산성분모형의 평균을 관리하기 위해 \bar{X} , EWMA 관리도를 연구하였고, Mullins (2003)는 화학실험에서 표본간 분산이 존재할 때 관리한계를 계산하는 방법을 연구하였다. Woodall과 Thomas (1995), Laubscher (1996)은 표본간 분산을 포함하는 관리도모형을 일반화 하였으며, Tsai와 Hsieh (2009)는 분산성분을 관리하기 위한 관리도를 제안하였다. 최근 Pascual과 Park (2017)은 분산성분이 존재할 때 웨이블공정 (Weibull process)의 백분율, 평균을 관리하기 위한 관리

[†] 이 논문은 2015년도 중앙대학교 연구장학기금 지원에 의한 것임.

¹ (06974) 서울특별시 동작구 흑석로 83, 중앙대학교 통계학과, 석사과정.

² 교신저자: (06974) 서울특별시 동작구 흑석로 83, 중앙대학교 통계학과, 교수. Email:cspark@cau.ac.kr

도를 제안하였다. 분산성분모형에 관한 최근 국내 연구로는 사영을 이용한 분산성분 분석 방법 (Choi, 2014; Choi, 2015)을 들 수 있다.

이 논문에서는 공정에서 표본간 분산이 있는 경우, 이를 고려한 공정표본의 분산을 추정하여 관리한계를 설정하고, 그에 따른 효율을 살펴보고자 한다. 공정의 평균을 관리하는 세 가지 대표적 관리유형, 즉 \bar{X} , 지수가중이동평균 (exponentially weighted moving average; EWMA) 관리도, 누적합관리도 (cumulative sum; CUSUM) 관리도를 활용하였다.

Park과 Lee (2014)에 의하면, 표준관리도의 절차는 아래와 같다. 표준관리도에서는 반응값이 아래와 같은 확률 모형을 따르는 것으로 가정한다.

$$X_{ij} = \mu + \sigma_W \cdot \epsilon_{ij}, \quad i = 1, 2, \dots, g, \quad j = 1, 2, \dots, n. \quad (1.1)$$

단 μ 는 평균, σ_W 는 표본내 표준편차, n 은 표본크기, g 는 표본군의 수를 나타내며, $\epsilon_{ij} \sim iid N(0, 1)$ 이다. 식 (1.1)로부터 공정품질 특성치 X 가 $N(\mu, \sigma_W^2/n)$ 을 따를 때, 공정평균 \bar{X} 는 $N(\mu, \sigma_W^2/n)$ 를 따르게 된다. 표준관리도에서의 평균 변화량은 σ_W/\sqrt{n} 의 배수로 나타내어 $\mu = \mu_0 + \delta\sigma_W/\sqrt{n}$ 이라 정의한다.

1.1. \bar{X} 관리도

\bar{X} 관리도는 공정평균을 관리하고자 하는 관리도로서 Shewhart 관리도 중 가장 널리 사용되며, 관리한계는 아래 식과 같이 설정한다.

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}}, \quad i = 1, 2, \dots. \quad (1.2)$$

단, 표본평균 \bar{X}_i 는 관리통계량, μ_0 는 목표치, c 는 관리한계계수이며 관리한계선은 μ_0 를 중심으로 표준오차의 관리한계계수배만큼 크거나 같은 값을 의미한다. \bar{X} 관리도에서의 절차는 관리한계선이 표시된 도표 상에 관리통계량 값을 적어나가는 일련의 과정을 의미한다. 이 때, 평균런길이는 검정력의 역수이므로 \bar{X} 관리도의 평균런길이는 아래와 같이 표현된다 (부록 A.1 참조). 단, $\delta = (\mu - \mu_0)/(\sigma_W/\sqrt{n})$ 이다.

$$ARL_{\bar{X}}(c, \delta) = \frac{1}{1 - \Phi(c - \delta) + \Phi(-c - \delta)}. \quad (1.3)$$

1.2. EWMA 관리도

EWMA 관리도는 시간에 따라 가중치를 달리 설정하는 관리도로서, 현시점에서 과거로 갈수록 가중치를 기하적으로 작게 주는 방법이다. EWMA 관리도의 관리한계는 아래와 같이 설정한다.

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}} \cdot \sqrt{\frac{\lambda}{2 - \lambda}} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}} \cdot \sqrt{\frac{\lambda}{2 - \lambda}}, \quad i = 1, 2, \dots. \quad (1.4)$$

단, c 는 관리한계계수이며 λ 는 가중치이다. EWMA 관리도에서의 절차는 가중치를 설정한다는 점을 제외하고는 \bar{X} 관리도와 같으나, EWMA 관리도의 평균런길이에 대한 함수의 정확한 표현은 알려져 있지 않다. 이 논문에서는 통계패키지 R에서 “SPC” library의 `xewma.arl`을 활용하여 계산하였다. 이를 통해 구한 평균런길이는 다음과 같이 표현한다.

$$ARLEWMA(\lambda, c, \delta) = xewma.arl(\lambda, c, \delta). \quad (1.5)$$

1.3. CUSUM 관리도

CUSUM관리도는 현재의 관측값과 함께 과거의 관측값을 공정의 상태판단에 사용하는 방법으로, 관리모수가 평균일 때 CUSUM관리도의 관리한계는 아래와 같이 설정한다.

$$\max_{0 \leq r \leq i} \left| \left\{ \sum_{l=i-r+1}^i \frac{\bar{X}_l - \mu_0}{\sigma_W / \sqrt{n}} - k \right\} \right| \geq h, \quad i = 1, 2, \dots \quad (1.6)$$

단, k 는 참고값, h 는 결정구간이다. CUSUM 관리도의 평균런길이에 대한 정확한 표현은 알려져 있지 않다. 이 논문에서는 통계패키지 R에서 “SPC” library의 `xcusum.arl`을 활용하여 계산하였다. 이를 통해 구한 평균런길이는 아래와 같이 정의한다.

$$ARLCUSUM(k, h, \delta) = xcusum.arl(k, h, \delta). \quad (1.7)$$

2. 분산성분모형

표준관리도가 표본내 분산만을 사용하여 관리도의 모수를 가정하는데 반해, 분산성분모형은 표본간 분산을 추가하여 고려함으로써 공정분산의 과소추정을 방지할 수 있다. 공정품질특성치를 X 라 하고 공정에서 제 i 번 표본의 j 번째 관측값을 X_{ij} 라 할 때 표본내 분산과 표본간 분산을 고려한 관측값의 분산성분모형은 아래와 같이 표현된다.

$$X_{ij} = \mu + \sigma_B \cdot w_i + \sigma_W \cdot \epsilon_{ij}, \quad i = 1, 2, \dots, g, \quad j = 1, 2, \dots, n. \quad (2.1)$$

단, σ_B 는 표본간 표준편차, σ_W 는 표본내 표준편차, n 은 표본크기, g 는 표본군의 수, w_i 는 i 번째 표본에 해당하는 오차, ϵ_{ij} 는 i 번째 표본의 j 번째 품질특성치에 해당하는 오차를 나타낸다. 이 때 w_i, ϵ_{ij} 는 독립동일분포 (independent identically distributed; iid) $N(0, 1)$ 를 따른다고 가정한다. 공정평균을 관리하는데 사용되는 통계량으로 표본평균 \bar{X} 를 사용하고, \bar{X} 관리도의 평균과 표준편차는 식 (2.1)로부터 $\mu_{\bar{X}} = \mu, \sigma_{\bar{X}} = \sqrt{\sigma_B^2 + \sigma_W^2/n}$ 이 되어 아래 분포를 따르게 된다.

$$\bar{X}_i \sim N(\mu, \sigma_B^2 + \sigma_W^2/n). \quad (2.2)$$

관리한계를 설정할 때 관리통계량 \bar{X} 의 평균의 관리모수 $\mu_{\bar{X},0}$ 와 표준편차의 관리모수 $\sigma_{\bar{X},0}$ 를 사용한다. 이 때, 평균의 관리값 μ_0 가 주어진 경우에는 $\mu_{\bar{X},0} = \mu_0$ 를 사용한다. 반면에 $\sigma_{\bar{X},0}$ 는 표본내 분산만을 고려할 때, 표본내 분산과 표본간 분산을 모두 고려할 때에 따라 다르게 설정한다. 이 때, 표본내 분산만 고려하는 경우는 Case I, 표본내 분산과 함께 표본간 분산을 이용하는 경우를 Case II라 정의한다.

Case I :

$$\sigma_{\bar{X},0} = \sigma_W / \sqrt{n}. \quad (2.3)$$

Case II :

$$\sigma_{\bar{X},0} = \sqrt{\sigma_{B,0}^2 + \sigma_W^2/n}. \quad (2.4)$$

단, $\sigma_{B,0}$ 는 관리상태일 때 표본간 표준편차의 값을 의미한다. 표본내 표준편차는 공정평균에 무관하게 일정한 값 σ_W 를 갖는다고 가정한다. 반면, 표본간 표준편차는 공정평균이 관리값 μ_0 부터 변화가 발생할 때 선형적으로 커진다고 가정하여 아래와 같이 설정한다.

$$\sigma_B = \frac{\sigma_W}{\sqrt{n}}(a + b\delta). \quad (2.5)$$

단, δ 는 공정평균의 변화량이다. 이 때, a 는 공정이 관리상태일 때 표본내 표준편차에 대한 표본간 표준편차의 비를 나타내며, b 는 공정이 이상상태일 때 $(\sigma_W/\sqrt{n})\delta$ 의 1단위 증가에 따른 표본간 표준편차의 증가량을 나타낸다. 공정이 관리상태일 때는 $\delta = 0$ 이므로 $\sigma_{B,0} = a \cdot \sigma_W/\sqrt{n}$ 가 된다. 이 때, δ 는 관리상태일 때 \bar{X} 의 표준편차의 배수로서 $\mu = \mu_0 + \delta\sqrt{\sigma_{B,0}^2 + \sigma_W^2/n}$ 로 표시한다. 따라서 δ 는 아래 식과 같다.

$$\delta = \frac{\mu - \mu_0}{\sqrt{\sigma_{B,0}^2 + \sigma_W^2/n}} = \frac{\mu - \mu_0}{(\sigma_W/\sqrt{n}) \cdot \sqrt{1 + a^2}}. \quad (2.6)$$

따라서 \bar{X} 의 표준편차는 다음과 같다.

$$\sigma_{\bar{X}} = \frac{\sigma_W}{\sqrt{n}}\sqrt{1 + (a + b\delta)^2}. \quad (2.7)$$

Case I 일 때는 $\sigma_{\bar{X},0} = \sigma_W/\sqrt{n}$ 이며, Case II의 경우에는 $\sigma_{\bar{X},0} = \sigma_W/\sqrt{n}(\sqrt{1 + a^2})$ 으로 표현된다.

2.1. \bar{X} 관리도

공정평균을 관리하는 \bar{X} 관리도의 관리한계는 아래와 같이 일반화하여 표현할 수 있다.

$$\mu_0 - c \cdot \sigma_{\bar{X},0} < \bar{X}_i < \mu_0 + c \cdot \sigma_{\bar{X},0}. \quad (2.8)$$

Case I, Case II 일 때의 $\sigma_{\bar{X},0}$ 가 각각 σ_W/\sqrt{n} , $(\sigma_W/\sqrt{n}) \cdot \sqrt{1 + a^2}$ 으로 상이하므로 다음과 같이 관리한계를 설정한다.

Case I :

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}}. \quad (2.9)$$

Case II :

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}}\sqrt{1 + a^2} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}}\sqrt{1 + a^2}. \quad (2.10)$$

\bar{X} 관리도의 평균런길이는 검정력의 역수이므로, 위의 식 (2.9), 식 (2.10)을 활용하여 평균런길이를 구한 결과는 다음과 같다 (부록 A.2 참조).

Case I :

$$ARL_{\bar{X}}\left(\frac{c}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}}\right). \quad (2.11)$$

Case II :

$$ARL_{\bar{X}}\left(\frac{\sqrt{1+a^2}c}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}}\right). \quad (2.12)$$

동일한 a, b, δ 조건 내에서 Case I 에 비해 Case II의 관리한계계수 값이 $\sqrt{1+a^2}$ 배 크다. 따라서 표본간 분산을 고려하지 않은 Case I 은 상대적으로 관리한계가 좁게 사용되었다고 할 수 있다.

2.2. EWMA 관리도

공정평균을 관리하는 EWMA 관리도의 관리한계는 아래와 같이 일반화하여 표현할 수 있다.

$$\mu_0 - c \cdot \sigma_{\bar{X},0} \cdot \sqrt{\frac{\lambda}{2-\lambda}} < \bar{X}_i < \mu_0 + c \cdot \sigma_{\bar{X},0} \cdot \sqrt{\frac{\lambda}{2-\lambda}}. \quad (2.13)$$

Case I , Case II 에 대한 관리한계는 다음과 같이 설정한다.

Case I :

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}} \cdot \sqrt{\frac{\lambda}{2-\lambda}} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}} \cdot \sqrt{\frac{\lambda}{2-\lambda}}. \quad (2.14)$$

Case II :

$$\mu_0 - c \cdot \frac{\sigma_W}{\sqrt{n}} \sqrt{1+a^2} \sqrt{\frac{\lambda}{2-\lambda}} < \bar{X}_i < \mu_0 + c \cdot \frac{\sigma_W}{\sqrt{n}} \sqrt{1+a^2} \sqrt{\frac{\lambda}{2-\lambda}}. \quad (2.15)$$

평균런길이는 다음과 같다 (부록 A.3 참조).

Case I :

$$ARL_{EWMA}(\lambda, \frac{c}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}}). \quad (2.16)$$

Case II :

$$ARL_{EWMA}(\lambda, \frac{\sqrt{1+a^2}c}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}}). \quad (2.17)$$

\bar{X} 관리도와 마찬가지로 표본간 분산을 고려하지 않은 Case I 은 상대적으로 관리한계를 좁게 사용하였음을 알 수 있다.

2.3. CUSUM 관리도

공정평균을 관리하는 CUSUM 관리도의 관리한계는 아래와 같이 일반화하여 표현할 수 있다.

$$\max_{0 \leq r \leq i} \sum_{l=i-r+1}^i \left[\frac{\bar{X}_l - \mu_0}{\sigma_{\bar{X},0}} - k \right] \geq h. \quad (2.18)$$

Case I, Case II에 대한 관리한계는 다음과 같이 설정한다.

Case I :

$$\max_{0 \leq r \leq i} \sum_{l=i-r+1}^i \left[\frac{\bar{X}_l - \mu_0}{\sigma_W/\sqrt{n}} - k \right] \geq h. \quad (2.19)$$

Case II :

$$\max_{0 \leq r \leq i} \sum_{l=i-r+1}^i \left[\frac{\bar{X}_l - \mu_0}{(\sigma_W/\sqrt{n}) \cdot \sqrt{1+a^2}} - k \right] \geq h. \quad (2.20)$$

평균런길이는 다음과 같다 (부록 A.4 참조).

Case I :

$$ARL_{CUSUM} \left(\frac{k}{\sqrt{1+(a+b\delta)^2}}, \frac{h}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}} \right). \quad (2.21)$$

Case II :

$$ARL_{CUSUM} \left(\frac{\sqrt{1+a^2}k}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}h}{\sqrt{1+(a+b\delta)^2}}, \frac{\sqrt{1+a^2}\delta}{\sqrt{1+(a+b\delta)^2}} \right). \quad (2.22)$$

Case II의 관리한계에 비해 Case I의 관리한계가 좁게 설정되었음을 알 수 있다.

표본간 분산을 나타내는 상수 a, b 에 따른 \bar{X} , EWMA, CUSUM 관리도의 평균런길이를 계산한 결과가 각각 Table 2.1, Table 2.2, Table 2.3이다. 이 때 관리도의 공정한 비교를 위해 a, b, \bar{X} 가 모두 0인 관리상태의 평균런길이 ARL_0 가 500이 되도록 하였다. 이를 만족하는 모수는 \bar{X} 관리도에서 $c = 3.09$, EWMA 관리도에서 $c = 2.96, \lambda = 0.2$, 그리고 CUSUM 관리도에서 $h = 5.07, k = 0.5$ 이다. CUSUM 관리도에서 r 를 사용하는 것은 탐지하고자 하는 평균 변화는 1배의 공정평균의 표준편차임을 의미한다.

오경보율은 \bar{X} 관리도에서 관리상태일 때의 평균런길이의 역수임이 알려져 있다. EWMA와 CUSUM 관리도의 오경보율은 관리상태의 평균런길이의 역수와 정확하게 일치하지 않지만, 비교의 편리성을 위해 아래와 같이 정의한다.

$$P(\text{false signal}) = \frac{1}{E_0(L)}. \quad (2.23)$$

단, $E_0(L)$ 은 평균런길이이다. Case I : \bar{X} 관리도에서 오경보율 ($\delta = 0$ 일 때 평균런길이의 역수)은 a 가 각각 0, 0.5, 1, 2일 때 0.0020, 0.0057, 0.0289, 0.1670 으로 a 값이 커짐에 따라 오경보율 또한 증가하는 것으로 나타났다. 한편, EWMA 관리도에서 오경보율 또한 a 가 각각 0, 0.5, 1, 2일 때 0.0020, 0.0048,

0.0183, 0.0849이며, CUSUM 관리도에서의 오경보율은 0.0020, 0.0051, 0.0205, 0.0952으로, \bar{X} 와 마찬가지로 a 가 커짐에 따라 오경보율도 증가하는 것으로 나타났다. 반면, $\delta > 0$ 인 이상상태의 평균런길이 값에서는 동일한 a 값 내에서 b 의 값이 큰 차이를 나타내지 않았고, 더욱이 \bar{X} 관리도의 $\{a = 0, \delta = 4\}$ 구간 같이 b 값이 커짐에도 역으로 평균런길이 값이 커지는 구간이 존재하였다.

Case II: 평균런길이를 계산한 결과를 살펴보면 $\delta = 0$ 일 때의 평균런길이 값이 동일하다. 즉, Case I의 경우 a 값에 따라 오경보율이 달라지는데 반해 Case II일 때 오경보율이 \bar{X} , EWMA, CUSUM 관리도 모두 0.0020으로 동일하다.

\bar{X} , EWMA, CUSUM 관리도에서 a, b 값의 조합에 대한 평균런길이의 변화형태를 나타낸 그림이 Figure 2.1, Figure 2.2, Figure 2.3이다.

Table 2.1 ARLs of the \bar{X} charts

a	b	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	500.00	54.59	7.26	2.16	1.22	500.00	54.59	7.26	2.16	1.22
	0.5	500.00	32.36	4.53	2.08	1.52	500.00	32.36	4.53	2.08	1.52
	1.0	500.00	13.96	3.08	1.94	1.59	500.00	13.96	3.08	1.94	1.59
	2.0	500.00	4.79	1.98	1.53	1.36	500.00	4.79	1.98	1.53	1.36
0.5	0.0	175.13	25.67	4.50	1.69	1.12	500.00	54.59	7.26	2.16	1.22
	0.5	175.13	12.04	3.13	1.82	1.43	500.00	20.07	4.00	2.07	1.54
	1.0	175.13	6.81	2.50	1.76	1.50	500.00	9.71	2.92	1.92	1.59
	2.0	175.13	3.44	1.82	1.47	1.33	500.00	4.21	1.98	1.55	1.38
1	0.0	34.63	8.42	2.34	1.26	1.04	500.00	54.59	7.26	2.16	1.22
	0.5	34.63	5.48	2.19	1.50	1.26	500.00	19.53	4.07	2.08	1.52
	1.0	34.63	4.02	2.01	1.54	1.36	500.00	10.21	3.08	1.98	1.60
	2.0	34.63	2.67	1.66	1.40	1.28	500.00	4.79	2.17	1.66	1.45
2	0.0	5.99	2.78	1.37	1.06	1.00	500.00	54.59	7.26	2.16	1.22
	0.5	5.99	2.50	1.48	1.19	1.08	500.00	24.02	4.53	2.09	1.45
	1.0	5.99	2.28	1.51	1.27	1.17	500.00	13.96	3.57	2.05	1.57
	2.0	5.99	1.94	1.44	1.27	1.20	500.00	7.06	2.67	1.87	1.57

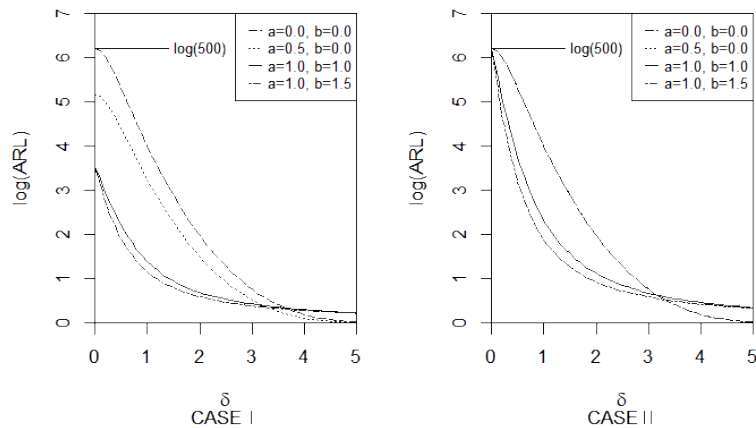


Figure 2.1 ARL curve of the \bar{X} charts for some selected value of (a, b)

Table 2.2 ARLs of the EWMA charts

<i>a</i>	<i>b</i>	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	500.00	10.54	3.74	2.38	1.86	500.00	10.54	3.74	2.38	1.86
	0.5	500.00	10.12	3.81	2.50	1.93	500.00	10.12	3.81	2.50	1.93
	1.0	500.00	9.24	3.82	2.59	2.06	500.00	9.24	3.82	2.59	2.06
	2.0	500.00	7.08	2.97	2.04	1.68	500.00	7.08	2.97	2.04	1.68
0.5	0.0	207.60	8.47	3.30	2.16	1.68	500.00	10.54	3.74	2.38	1.86
	0.5	207.60	7.96	3.40	2.28	1.79	500.00	9.66	3.83	2.52	1.95
	1.0	207.60	7.33	3.38	2.35	1.90	500.00	8.74	3.80	2.59	2.06
	2.0	207.60	5.55	2.64	1.91	1.61	500.00	6.61	2.97	2.08	1.72
1	0.0	54.60	5.85	2.60	1.76	1.31	500.00	10.54	3.74	2.38	1.86
	0.5	54.60	5.66	2.70	1.88	1.51	500.00	9.63	3.83	2.51	1.93
	1.0	54.60	5.34	2.72	1.98	1.65	500.00	8.82	3.82	2.59	2.05
	2.0	54.60	4.26	2.30	1.75	1.52	500.00	7.08	3.26	2.27	1.85
2	0.0	11.79	3.42	1.74	1.22	1.04	500.00	10.54	3.74	2.38	1.86
	0.5	11.79	3.38	1.84	1.37	1.18	500.00	9.84	3.81	2.47	1.89
	1.0	11.79	3.25	1.90	1.48	1.30	500.00	9.24	3.84	2.55	1.98
	2.0	11.79	2.85	1.82	1.50	1.35	500.00	8.15	3.71	2.55	2.04

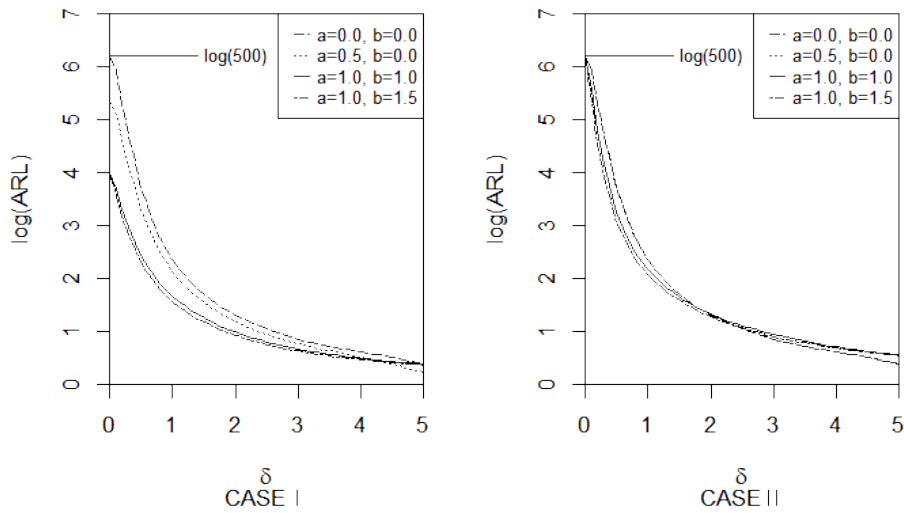


Figure 2.2 ARL curve of the EWMA charts for some selected value of (*a*, *b*)

Table 2.3 ARLs of the CUSUM charts

a	b	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	500.00	10.52	4.06	2.60	2.03	500.00	10.52	4.06	2.60	2.03
	0.5	500.00	16.46	6.21	3.80	2.77	500.00	10.30	4.12	2.72	2.10
	1.0	500.00	14.84	5.25	3.21	2.40	500.00	9.60	4.05	2.72	2.14
	2.0	500.00	8.15	3.11	2.12	1.74	500.00	7.01	2.96	2.07	1.72
0.5	0.0	196.12	13.71	5.09	3.17	2.37	500.00	10.52	4.06	2.60	2.03
	0.5	196.12	12.93	5.19	3.28	2.45	500.00	9.97	4.13	2.74	2.12
	1.0	196.12	10.22	4.29	2.80	2.18	500.00	9.04	3.99	2.71	2.14
	2.0	196.12	5.95	2.75	1.99	1.67	500.00	6.47	2.96	2.11	1.75
1	0.0	48.67	9.60	4.12	2.67	2.05	500.00	10.52	4.06	2.60	2.03
	0.5	48.67	8.29	3.86	2.56	1.97	500.00	9.95	4.13	2.73	2.10
	1.0	48.67	6.78	3.34	2.33	1.88	500.00	9.13	4.05	2.75	2.16
	2.0	48.67	4.51	2.42	1.84	1.58	500.00	7.01	3.28	2.31	1.89
2	0.0	10.50	4.78	2.49	1.72	1.32	500.00	10.52	4.06	2.60	2.03
	0.5	10.50	4.29	2.39	1.71	1.39	500.00	10.12	4.12	2.70	2.07
	1.0	10.50	3.81	2.25	1.71	1.45	500.00	9.60	4.13	2.76	2.14
	2.0	10.50	3.03	1.95	1.58	1.41	500.00	8.31	3.84	2.64	2.11

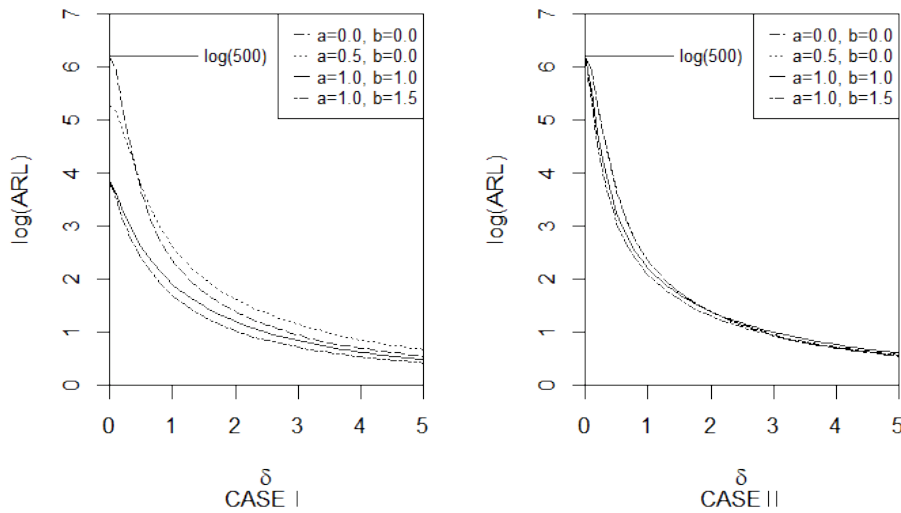


Figure 2.3 ARL curve of the CUSUM charts for some selected value of (a, b)

3. 모수의 추정

제 2절에서는 공정평균과 표준편차를 아는 경우의 관리도의 효율에 대해 논하였으나, 실제 공정에서는 공정모수를 모르는 상태, 즉 공정모수의 관리값이 주어지지 않은 상태에서 모수를 추정하여 관리한계를 설정한 다음 공정관리를 시행하게 된다. 이 절에서는 공정평균과 공정표준편차를 추정할 때의 관리도의 평균런길이를 계산하는 방법에 대해 알아보고 또한 그때의 평균런길이를 통해 관리도의 효율을

살펴보고자 한다. 이를 위해서 관리상태의 공정평균과 표준편차는 제 1국면 (phase I) 자료를 사용하여 추정한다.

3.1. 공정평균과 공정표준편차의 추정과 분포

표본간 분산에 의해 결정되는 상수 a 는 과거의 경험에 의해 결정되거나, 매번 제 1국면 자료에 의해 추정될 수 있다. 제 1국면 자료를 사용할 때는 랜덤효과 (random effect) 검정을 위한 일원분산분석 (one-way ANOVA)를 사용하여 추정한다 (Wetherill and Brown, 1991). 이 논문에서는 상수 a 가 과거의 경험에 의해 결정되는 것으로 간주하여 알려진 값으로 가정한다. 공정평균의 추정치 $\hat{\mu}_{\bar{X},0}$ 는 제 1국면의 총평균 \bar{X} 는 다음과 같은 확률모형을 따른다.

$$\bar{X} = \mu + \sigma_B \cdot \bar{w} + \sigma_W \bar{\epsilon}. \quad (3.1)$$

단, $\bar{\epsilon} = \sum_{i=1}^m \sum_{j=1}^n \epsilon_{ij}/nm$, $\bar{w} = \sum_{i=1}^m w_i/m$ 이다. 이 때, m 은 제 1국면의 표본수, n 은 표본의 크기이다. 식 (3.1)에 의해 표본평균 \bar{X} 는 아래와 같은 분포를 따른다.

$$\bar{X} \sim N\left(\mu, \frac{\sigma_W^2/n + \sigma_B^2}{m}\right). \quad (3.2)$$

관리상태일 때 \bar{X} 의 누적분포함수를 $F_{\bar{X}}$ 라 할 때, 그 역함수 $F_{\bar{X}}^{-1}$ 은 아래와 같이 표현된다.

$$F_{\bar{X}}^{-1}(x) = \Phi^{-1}(x) \sqrt{\frac{\sigma_W^2/n + \sigma_{B,0}^2}{m}} + \mu_0. \quad (3.3)$$

관리상태일 때 \bar{X} 의 분산은 Case I 일 때 $\sigma_W^2/(n \cdot m)$, Case II 일 때 $(\sigma_W^2/(n \cdot m))\sqrt{1+a^2}$ 이므로 누적분포함수의 형태가 상이하게 표현되며 식은 다음과 같다.

Case I :

$$F_{\bar{X}}^{-1}(x) = \Phi^{-1}(x) \sqrt{\frac{\sigma_W^2/n}{m}} + \mu_0. \quad (3.4)$$

Case II :

$$F_{\bar{X}}^{-1}(x) = \Phi^{-1}(x) \sqrt{\frac{\sigma_W^2/n \cdot (1+a^2)}{m}} + \mu_0. \quad (3.5)$$

공정표준편차는 Case I, Case II 모두 σ_W 의 함수 형태로 표현되고, σ_W 의 추정량은 $\hat{\sigma}_W = \bar{S}/c_4(n)$ 으로 사용한다.

단, $\bar{S} = \sum_{i=1}^m S_i/m$, $i = 1, 2, \dots, m$, $S_i = \sqrt{\sum_{j=1}^n (X_{ij} - \bar{X}_i)/(n-1)}$, $c_4(n)$ 은 표본크기 n 에 따라 결정되는 상수이다. 따라서 Case I, Case II에 따른 공정표준편차의 추정치는 아래 식과 같이 표현된다.

Case I :

$$\hat{\sigma}_{\bar{X},0} = \frac{\hat{\sigma}_W}{\sqrt{n}} = \frac{\bar{S}}{c_4(n) \cdot \sqrt{n}}. \quad (3.6)$$

Case II :

$$\hat{\sigma}_{\bar{X},0} = \frac{\hat{\sigma}_W}{\sqrt{n}} \sqrt{1+a^2} = \frac{\bar{S}}{c_4(n) \cdot \sqrt{n}} \cdot \sqrt{1+a^2}. \quad (3.7)$$

카이제곱분포의 특성, 즉 $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$ 를 이용하면, 식 (3.6), 식 (3.7)에 해당하는 공정표준편차의 분포를 추정할 수 있다. 이제 m 개의 서로 독립인 $\chi^2(n-1)$ 을 따르는 확률변수를 $\mathbf{Z} = (z_1, z_2, \dots, z_m)$ 라 하면 $\hat{\sigma}_{\bar{X},0}$ 의 분포는 아래와 같다.

Case I :

$$\hat{\sigma}_{\bar{X},0} = \frac{\bar{S}}{c_4(n) \cdot \sqrt{n}} \sim \frac{\sigma_W}{c_4(n) \cdot \sqrt{n \cdot (n-1)}} \sum_{i=1}^m \frac{\sqrt{z_i}}{m}. \quad (3.8)$$

Case II :

$$\hat{\sigma}_{\bar{X},0} = \frac{\bar{S}}{c_4(n) \cdot \sqrt{n}} \sim \frac{\sigma_W \cdot \sqrt{1+a^2}}{c_4(n) \cdot \sqrt{n \cdot (n-1)}} \sum_{i=1}^m \frac{\sqrt{z_i}}{m}. \quad (3.9)$$

식 (3.8), (3.9)에서 $\hat{\sigma}_{\bar{X},0} = g(\mathbf{Z})$ 라 하자. 이 경우 $g(\mathbf{Z})$ 는 그 분포의 형태가 알려지지 않고 또한 분포 함수가 닫힌 형태 (closed form)로 나타나지 않는다. 따라서 카이제곱 분포로부터 모의실험을 통해 난수를 생성하고 그 난수의 함수값을 이용하여 분포함수를 추정할 수 있다. 구체적인 알고리즘은 아래와 같다. $g(\mathbf{Z})$ 의 표본분포함수는 $\hat{F}_{g(\mathbf{z})}(x) = \{\# \text{ of } g(\mathbf{z}) \leq x\} / N$ 으로 정의한다.

- (1) $\chi^2(n-1)$ 분포로부터 N 개의 난수벡터를 생성한다.
- (2) $D = (g(\mathbf{Z}_1), \dots, g(\mathbf{Z}_N))$ 을 계산한다.
- (3) D 를 사용하여 $g(\mathbf{Z})$ 의 표본분포함수 $\hat{F}_{g(\mathbf{z})}(x)$ 를 계산한다.
- (4) (3)의 결과에 연속성 수정을 하여 표본분포함수 $F_{g(\mathbf{z})}^*(x)$ 를 추정한다.

표본분포함수 $\hat{F}_{g(\mathbf{z})}(x)$ 를 모든 실수구간에서 정의되도록 연속성 수정을 하여 표현하면 아래와 같다 (Park, 2017).

$$F_{g(\mathbf{z})}^*(x) = \begin{cases} \frac{1}{2N} e^{x-\min(D)}, & x < \min(D), \\ \hat{F}_{g(\mathbf{z})}(x) - \frac{1}{2N}, & \min(D) \leq x \leq \max(D), \\ \frac{2N-1}{2N} + \frac{1-e^{-(x-\max(D))}}{2N}, & x > \max(D). \end{cases} \quad (3.10)$$

3.2. 평균런길이의 계산

공정관리도의 런길이를 L 이라 하자. 공정모수 $\mu_{\bar{X},0}$ 와 $\sigma_{\bar{X},0}$ 가 $\hat{\mu}_{\bar{X},0} = u$, $\hat{\sigma}_{\bar{X},0} = v$ 로 추정될 때 조건부평균런길이를 $E(L|u, v)$ 이라 하면 평균런길이는 아래와 같이 표현된다.

$$E(L) = \int_0^\infty \int_{-\infty}^\infty E(L|u, v) f_u(u) f_v(v) dudv. \quad (3.11)$$

단, f_u 는 $\hat{\mu}_{\bar{X},0}$ 의 확률밀도함수, f_v 는 $\hat{\sigma}_{\bar{X},0}$ 의 확률밀도함수이다. 이 적분식은 $\hat{\mu}_{\bar{X},0}$ 와 $\hat{\sigma}_{\bar{X},0}$ 의 누적분포 함수를 사용하여 아래식과 같이 근사된다 (Park, 2017).

$$E(L) \approx \sum_{j=1}^l \sum_{i=1}^l E(L|F_u^{-1}(x), F_v^{-1}(y)) \left(\frac{1}{l^2}\right). \tag{3.12}$$

단, l 은 양의 정수, F_u 는 $\hat{\mu}_{\bar{X},0}$ 의 분포함수, F_v 는 $\hat{\sigma}_{\bar{X},0}$ 의 분포함수이다. 식 (3.12)에 의한 평균런길이는 세 종류의 관리형태에 대해 계산되었다.

계산된 평균런길이는 Table 3.1부터 Table 3.3에 나타나 있다. 이 결과들은 모의실험 방법을 이용하여 추정된 결과이기 때문에 추정치의 값에 오차가 포함되어 있으며, 편리상 반복수 $N = 10,000$, $n = 5$, $m = 30$, $\sigma_W/\sqrt{n} = 1$, $l = 20$ 인 경우로 한정하였다. 모수 추정에 따른 평균런길이의 변화를 알아보기 위해 제 2절에서와 동일한 모수들을 사용하여 계산하였다. 즉 \bar{X} 관리도에서 $c = 3.09$, EWMA 관리도에서 $c = 2.96$, $\lambda = 0.2$, CUSUM 관리도에서 $h = 5.07$, $k = 0.5$ 를 사용하였다.

\bar{X} 관리도에서 Case I에서 $b = 0$ 일 때 오경보율이 a 가 각각 0, 0.5, 1, 2일 때 0.0016, 0.0048, 0.0269, 0.1642으로 a 값이 커짐에 따라 오경보율 또한 증가하는 것으로 나타났다. 한편, EWMA 관리도에서 오경보율 또한 a 가 각각 0, 0.5, 1, 2일 때 0.0017, 0.0043, 0.0174, 0.0838 이며, CUSUM 관리도에서의 오경보율은 0.0016, 0.0045, 0.0195, 0.094으로, \bar{X} 와 마찬가지로 a 가 커짐에 따라 오경보율도 증가하는 것으로 나타났다.

Case II일 때를 살펴보면 a , b 의 값과 관계없이 δ 가 0일 때의 평균런길이 값이 거의 동일한 값을 나타냈다. 즉, 관리상태일 때 a 값과 오경보율이 비례관계에 있었던 Case I 과 달리 Case II 일 때는 a , b 값과 상관없이 관리상태일 때의 오경보율이 동일한 값을 갖는다.

Table 3.1 ARLs of the estimated \bar{X} charts

a	b	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	639.25	62.73	7.76	2.22	1.23	639.30	62.70	7.76	2.22	1.23
	0.5	640.49	35.64	4.65	2.10	1.52	639.88	35.61	4.65	2.10	1.52
	1.0	639.03	14.62	3.11	1.94	1.59	639.87	14.63	3.11	1.94	1.59
	2.0	639.22	4.85	1.99	1.53	1.36	639.31	4.85	1.99	1.53	1.36
0.5	0.0	206.79	28.07	4.68	1.71	1.13	639.51	62.78	7.77	2.22	1.23
	0.5	206.61	12.58	3.17	1.83	1.43	640.93	21.41	4.07	2.08	1.55
	1.0	206.62	6.96	2.52	1.77	1.50	638.67	10.01	2.94	1.93	1.59
	2.0	206.34	3.46	1.82	1.47	1.33	639.37	4.26	1.99	1.55	1.38
1	0.0	37.23	8.74	2.38	1.27	1.04	638.33	62.72	7.76	2.22	1.23
	0.5	37.26	5.58	2.20	1.50	1.26	638.83	20.80	4.14	2.09	1.52
	1.0	37.23	4.06	2.02	1.55	1.36	639.65	10.55	3.11	1.98	1.60
	2.0	37.25	2.68	1.66	1.40	1.29	639.10	4.86	2.18	1.66	1.45
2	0.0	6.09	2.80	1.37	1.06	1.01	639.61	62.77	7.76	2.22	1.23
	0.5	6.09	2.52	1.48	1.19	1.08	639.04	25.92	4.65	2.11	1.46
	1.0	6.09	2.28	1.51	1.27	1.17	639.51	14.62	3.62	2.06	1.58
	2.0	6.09	1.94	1.44	1.27	1.20	638.98	7.21	2.68	1.87	1.57

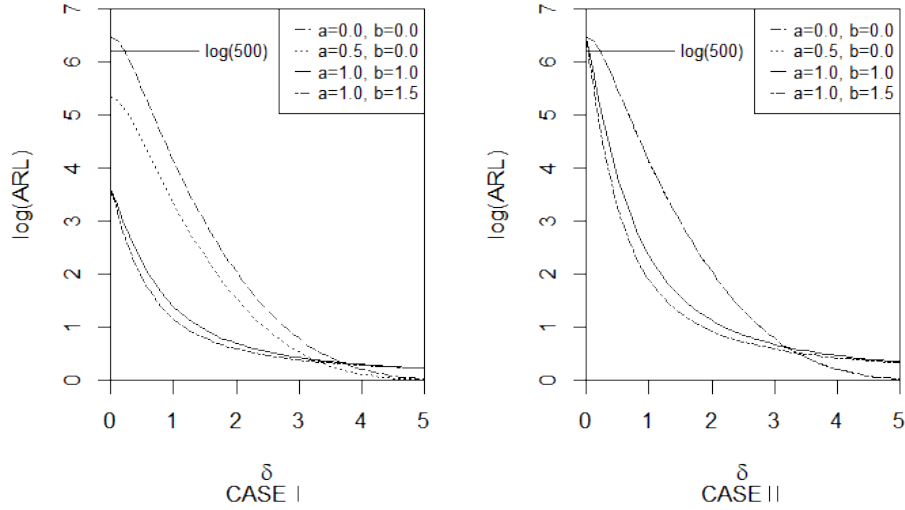


Figure 3.1 ARL curve of the estimated \bar{X} charts for some selected value of (a, b)

Table 3.2 ARLs of the estimated EWMA charts

a	b	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	600.13	10.67	3.75	2.38	1.86	599.96	10.67	3.75	2.38	1.86
	0.5	600.25	10.23	3.82	2.50	1.93	600.21	10.22	3.82	2.50	1.93
	1.0	600.43	9.31	3.83	2.59	2.06	600.44	9.31	3.83	2.59	2.06
	2.0	600.44	7.11	2.98	2.04	1.68	599.73	7.11	2.98	2.04	1.68
0.5	0.0	233.88	5.51	2.40	1.66	1.18	600.46	10.67	3.75	2.38	1.86
	0.5	233.52	5.44	2.50	1.73	1.39	599.73	9.74	3.84	2.52	1.95
	1.0	233.58	5.31	2.59	1.87	1.57	599.84	8.78	3.80	2.59	2.06
	2.0	233.87	4.65	2.36	1.77	1.52	600.26	6.63	2.98	2.08	1.72
1	0.0	57.40	3.82	1.86	1.23	1.02	599.84	10.67	3.75	2.38	1.86
	0.5	57.43	3.85	1.93	1.39	1.18	600.57	9.71	3.83	2.51	1.93
	1.0	57.41	3.83	2.03	1.54	1.33	600.36	8.87	3.83	2.59	2.06
	2.0	57.38	3.52	2.00	1.58	1.40	600.45	7.11	3.27	2.28	1.85
2	0.0	11.94	2.55	1.34	1.04	1.00	599.43	10.67	3.75	2.38	1.86
	0.5	11.94	2.59	1.44	1.14	1.04	599.55	9.93	3.82	2.48	1.89
	1.0	11.94	2.59	1.54	1.24	1.13	599.72	9.31	3.85	2.55	1.99
	2.0	11.94	2.47	1.60	1.35	1.24	600.25	8.18	3.71	2.55	2.04

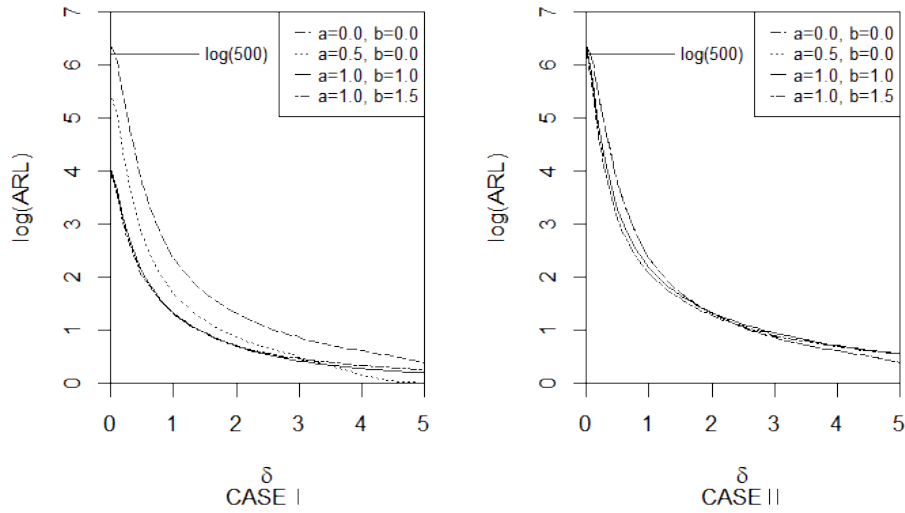


Figure 3.2 ARL curve of the estimated EWMA charts for some selected value of (a, b)

Table 3.3 ARLs of the estimated CUSUM charts

a	b	Case I					Case II				
		δ					δ				
		0	1	2	3	4	0	1	2	3	4
0	0.0	617.83	10.59	4.06	2.61	2.03	617.04	10.58	4.06	2.61	2.03
	0.5	617.63	10.36	4.13	2.73	2.10	618.56	10.36	4.13	2.72	2.10
	1.0	618.44	9.65	4.05	2.72	2.15	617.75	9.65	4.05	2.72	2.15
	2.0	617.53	7.04	2.97	2.07	1.72	617.97	7.04	2.97	2.07	1.72
0.5	0.0	224.23	8.77	3.59	2.36	1.86	617.59	10.59	4.06	2.61	2.03
	0.5	224.22	8.36	3.68	2.48	1.94	618.05	10.03	4.14	2.74	2.12
	1.0	224.26	7.61	3.54	2.45	1.98	617.13	9.08	3.99	2.71	2.15
	2.0	224.32	5.45	2.65	1.94	1.65	617.32	6.49	2.97	2.11	1.76
1	0.0	51.34	6.23	2.84	1.94	1.48	617.73	10.58	4.06	2.61	2.03
	0.5	51.38	5.97	2.92	2.03	1.62	617.28	10.01	4.14	2.73	2.10
	1.0	51.34	5.49	2.85	2.06	1.71	616.81	9.17	4.05	2.75	2.17
	2.0	51.32	4.20	2.32	1.79	1.56	617.01	7.04	3.28	2.31	1.89
2	0.0	10.64	3.65	1.90	1.32	1.07	617.47	10.58	4.06	2.61	2.03
	0.5	10.64	3.54	1.97	1.45	1.23	616.94	10.18	4.13	2.71	2.07
	1.0	10.64	3.34	1.99	1.55	1.34	617.78	9.65	4.13	2.76	2.14
	2.0	10.64	2.86	1.86	1.53	1.38	617.86	8.34	3.84	2.65	2.12

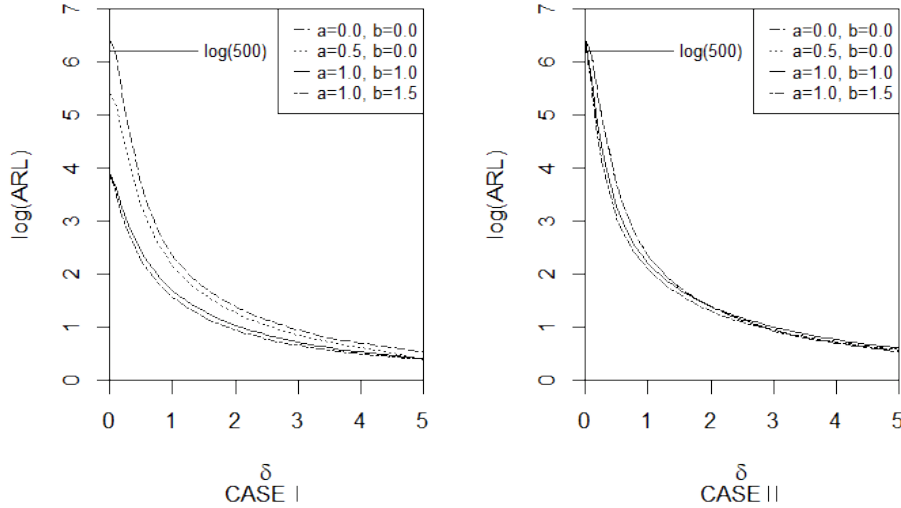


Figure 3.3 ARL curve of the estimated CUSUM charts for some selected value of (a, b)

4. 오경보율의 비교

이 절에서는 제 1종 오류에 해당하는 오경보율을 비교했다. 표본간 분산이 존재하는 표본표준편차의 추정방법 (Case I 과 Case II)에 따라 오경보율이 어떤 변화를 나타내는 지 살펴보도록 한다.

아래 Figure 4.1은 $\{a = 1, b = 1\}$ 일 때 \bar{X} 관리도의 평균런길이를 Case I 과 Case II로 구분하여 비교한 그림이고, Figure 4.2는 공정모수를 추정할 경우 중 $\{a = 1, b = 1\}$ 일 때 \bar{X} 관리도의 평균런길이를 Case I 과 Case II로 구분하여 비교한 그림이다.

Case I 에서 a 의 값이 상승함에 따라 오경보율이 감소하는 추세를 보이고 있음에 반해 Case II에서는 a 값과는 무관하게 오경보율이 고정 값을 나타냄을 관찰할 수 있다. a 가 0이 아니라는 것은 관리상태일 때 표본간 분산이 존재한다는 의미이나, Case I 은 표본간 분산을 고려하지 않기 때문에 표본간 분산의 값이 증가할수록 오경보를 더 빨리 감지하게 된다. 결론적으로 a 가 0과 가까운 경우에는 편리성을 담보할 수 있는 단순확률모형을 사용해도 무방하나 a 값이 0이 아닌 경우에는 필수적으로 Case II, 즉 분산성분모형을 사용하는 것이 올바르다. 이와 같은 현상은 공정모수를 추정하여 관리한계를 설정하는 Figure 4.2의 경우에서도 동일하다.

\bar{X} 관리도 외 EWMA와 CUSUM 관리도를 사용하여 Case I , Case II의 오경보율을 비교한 결과가 아래 그림과 같다. EWMA와 CUSUM관리도의 오경보율은 식 (2.23)으로 계산하였다. 조건은 \bar{X} 관리도와 동일하게 $\{\delta = 0, b = 0\}$ 인 경우로 한정하였다. EWMA와 CUSUM 관리도에서 Case II의 오경보율이 Case II의 경우 0에 근접하는 형태를 보임에 반해 Case I 은 a 값이 증가함에 따라 오경보율도 증가하는 양상을 보였다. 이는 \bar{X} 관리도와 동일하다.

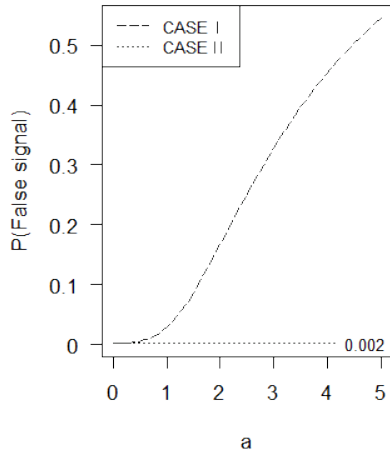


Figure 4.1 False signal rate of \bar{X} chart when parameters are known for $\{a = 1, b = 1\}$

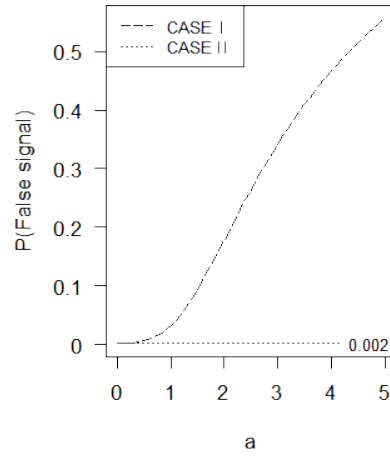


Figure 4.2 False signal rate of \bar{X} chart when parameters are estimated for $\{a = 1, b = 1\}$

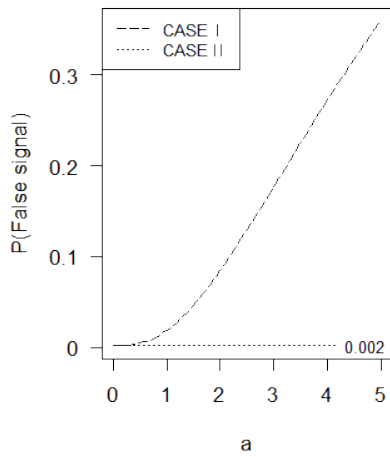


Figure 4.3 False signal rate of EWMA chart when parameters are known for $\{a = 1, b = 1\}$

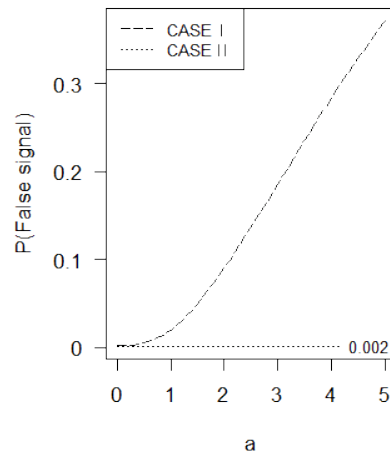


Figure 4.4 False signal rate of EWMA chart when parameters are estimated for $\{a = 1, b = 1\}$

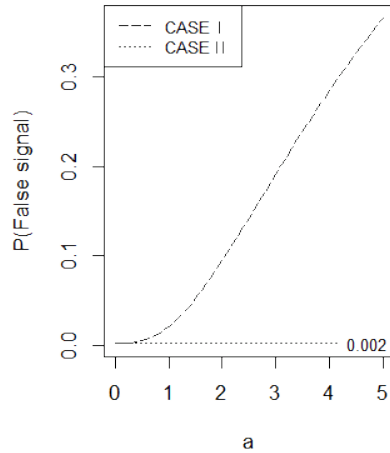


Figure 4.5 False signal rate of CUSUM chart when parameters are known for $\{a = 1, b = 1\}$

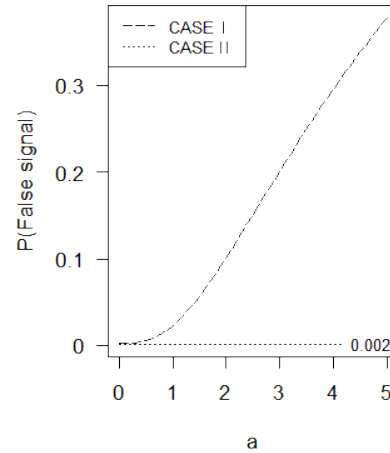


Figure 4.6 False signal rate of CUSUM chart when parameters are estimated for $\{a = 1, b = 1\}$

5. 결론

표본추출 시점에 따라 제품품질의 평균이 변화하는 공정에서는 표본내 분산 외에도 표본간 분산을 고려하여 공정모수를 추정해야 한다. 하지만 대부분의 공정에서는 편리성의 이유로 표본내 분산만을 사용하여 관리도를 작성하게 되고, 이 때 분산을 과소추정하게 되므로 실제 이상상태가 아님에도 공정에 문제가 있다고 판단하게 된다. 특히, 표본내 분산 대비 표본간 분산의 크기가 클 수록 이상신호를 더 빨리 감지하게 되는데, 이러한 문제가 발생하면 무엇보다 ‘공정의 이상상태를 탐지하고 이를 관리한다’는 관리도의 목적을 상실하게 되고, 일종의 오류가 발생하게 된 꼴이므로 공정 관리자와의 신뢰관계에 치명적인 악영향을 끼칠 수 있다.

이 논문에서는 앞서 언급된 문제를 해결하기 위해 분산의 형태에 따라 관리도를 Case I, Case II로 나누고 \bar{X} , EWMA, CUSUM 관리도를 이용하여 관리한계를 설정한 후 이들간 평균런길이를 비교하였다. 또한, 실제 공정에서는 공정평균과 공정표준편차를 모르는 상태에서 이를 추정하여 관리값으로 사용한다는 사실에 기초하여 참 평균과 표준편차를 모른다는 전제하에 공정모수를 추정하여 관리한계를 설정하고 평균런길이를 비교하였다.

그 결과, 표본간 분산이 존재할 때, 분산성분모형을 사용한 관리도가 표준관리도에 비해 평균런길이가 길고, 오경보율은 낮았다. 또한 공정모수를 아는 경우와 추정하여 사용하는 경우의 평균런길이 및 오경보율의 차이는 미미했다. 이를 종합하여, 표본간 분산이 존재할 가능성이 항상 존재하는 실제 공정에서 공정관리절차에 앞서 표본간 분산의 존재여부를 검정하고, 공정분산을 올바르게 추정하는 것이 다른 어떤 복잡한 관리도를 적용시키는 것 보다 중요하다고 판단할 수 있다.

References

- Choi, J. S. (2014). Projection analysis for two-way variance components. *Journal of the Korean & Information Science Society*, **25**, 547-554.
- Choi, J. S. (2015). The analysis of random effects model by projections. *Journal of the Korean & Information Science Society*, **26**, 31-39.

- Laubscher, N. F. (1996). A variance components model for statistical process control. *South African Statistical Journal*, **30**, 27-47.
- Mullins, E. (2003). *Statistics for the quality control chemistry laboratory*, The Royal Statistical Society of Chemistry, Cambridge.
- Park, C. (1997). Design of xbar and ewma charts in a variance components model. *Communications in Statistics : Theory and Methods*, **27**, 659-672.
- Park, C. and Lee, J. (2014). *Statistical process control*, Freeacademy, Seoul.
- Park, C. (2016). Average run length calculation of the ewma control chart using the first passage time of the markov process. *The Korean Journal of Applied Statistics*, **30**, 1-12.
- Pascual, F. G. and Park, C. (2017). Control charts for monitoring the mean and percentiles of weibull processes with variance components. *submitted to International Journal of Production Research*.
- Tsai, T. and Hsieh, Y. (2009). Simulated shewhart control chart for monitoring variance components. *International Journal of Reliability, Quality and Safety Engineering*, **16**, 1-22.
- Wetherill, G. B. and Brown, D. W. (1991). *Statistical process control: theory and practice*, Springer, Chapman and Hall.
- Woodall, W. H. and Thomas, E. V. (1995). Statistical process control with several components of common cause variability. *IIE Transactions*, **27**, 757-764.

Design and efficiency of the variance component model control chart[†]

Chan Yang Cho¹ · Changsoon Park²

¹²Department of Statistics, Chung-Ang University

Received 5 July 2017, revised 22 August 2017, accepted 31 August 2017

Abstract

In the standard control chart assuming a simple random model, we estimate the process variance without considering the between-sample variance. If the between-sample exists in the process, the process variance is under-estimated. When the process variance is under-estimated, the narrower control limits result in the excessive false alarm rate although the sensitivity of the control chart is improved. In this paper, using the variance component model to incorporate the between-sample variance, we set the control limits using both the within- and between-sample variances, and evaluate the efficiency of the control chart in terms of the average run length (ARL). Considering the most widely used control chart types such as \bar{X} , EWMA and CUSUM control charts, we compared the differences between two cases, Case I and Case II, where the between-sample variance is ignored and considered, respectively. We also considered the two cases when the process parameters are given and estimated. The results showed that the false alarm rate of Case I increased sharply as the between-sample variance increases, while that of Case II remains the same regardless of the size of the between-sample variance, as expected.

Keywords: Average run length, between-sample variance, false alarm rate, simple random model, within-sample variance.

[†] This research was supported by the Chung-Ang University Research Scholarship Grants in 2015.

¹ Graduate student, Department of Statistics, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, Seoul 06974, Korea.

² Corresponding author: Professor, Department of Applied Statistics, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, Seoul 06974, Korea. E-mail: cspark@cau.ac.kr