

소표본 자기상관 자료의 분산 추정을 위한 최적 부분군 크기에 대한 연구

이종선*[†], 이재준*, 배순희*

* 인하대학교 자연과학대학 통계학과

To study of optimal subgroup size for estimating variance on autocorrelated small samples

Jong Seon Lee, Jae Jun Lee, Soon Hee Bae

Department of Statistics, Inha University

Key Words: Autocorrelation, Average Run Length, AR(1), Optimal Subgroup Size

Abstract

To conduct statistical process control needs the assumption that the process data are independent. However, most of chemical processes, like a semi-conduct processes do not satisfy the assumption because of autocorrelation. It causes abnormal out of control signal in the process control and misleading process capability. In this study, we introduce that Shore's method to solve the problem and to find the optimal subgroup size to estimate variance for AR(1) model. Especially, we focus on finding an actual subgroup size for small samples using simulation. It may be very useful for statistical process control to analyze process capability and to make a Shewhart chart properly.

1. 서론

최근 IT의 발전에 따른 컴퓨터 저장 용량의 확대에 자동화된 공정에서의 공정관리는 표본추출을 통한 샘플링 검사보다는 생산되는 모든 제품을 측정하여 관리하는 전수검사의 형태가 보편화 되는 추세다. 특히 반도체나 석유화학 공정과 같은 연속공정(Continuous Process)에서 이런 관리방법이 보편화되어 있다. 그런데 이러한 공정의 데이터 특징은 자료들이 서로 자기상관(Autocorrelation)을 지니고 있다는 것이다. 자료들이 자기상관이 존재함에도 독립을 가정하여 σ 를 추정하고 일반적인 슈와르트 관리도를 사용하게 되면, σ 가 과소추정(underestimated)된다. 이러한 문제는 슈와르트 관리도 적용시 관리한계선(Control Limits)의 폭이 줄어들어 빈번한 관리이탈이나 잘못된 공정능력지수 등이

산출되게 되는 것이다.

이러한 자기상관의 문제를 해결하는 방안으로 부분군 평균을 이용하는 것에 대해, Runger and Willemain(1995), Shore(1997)등 많은 연구가 이루어졌다.

본 논문에서는 자기상관이 존재하는 1차 자기회귀모형, 즉 AR(1) 자료들에 대해 자기상관 해소를 위한 최적의 부분군의 크기는 얼마인지를 보이고자 한다. 또한, 이를 통해 Shore(1999)가 제안한 부분군 평균으로 실제 공정의 산포를 추정하여 공정능력분석에 활용하는 방법을 소개하고 공정능력분석에 필요한 σ 추정 방법을 제시하고자 한다.

2. 자기상관 자료의 최적 부분군 크기

2.1 1차 자기회귀 모형: AR(1)

[†] 교신저자 jslee@stat.inha.ac.kr

1차 자기회귀 모형은 흔히 AR(1)으로 표기하며 다음과 같이 나타낼 수 있다.

$$X_t = \phi X_{t-1} + a_t \quad (1)$$

여기서 $a_t \sim NID(0, \sigma^2)$ 이 된다. 따라서 시점 t 에서의 자료는 시점 $t-1$ 이하의 자료들에 대해서 점진적으로 영향을 받게 된다. 이와 같이 각 시점의 자료가 다른 시점의 자료에 영향을 받는 것을 자기상관(Autocorrelation)이 존재한다고 말한다.

2.2 자기상관 자료의 평균 런 길이(ARL)

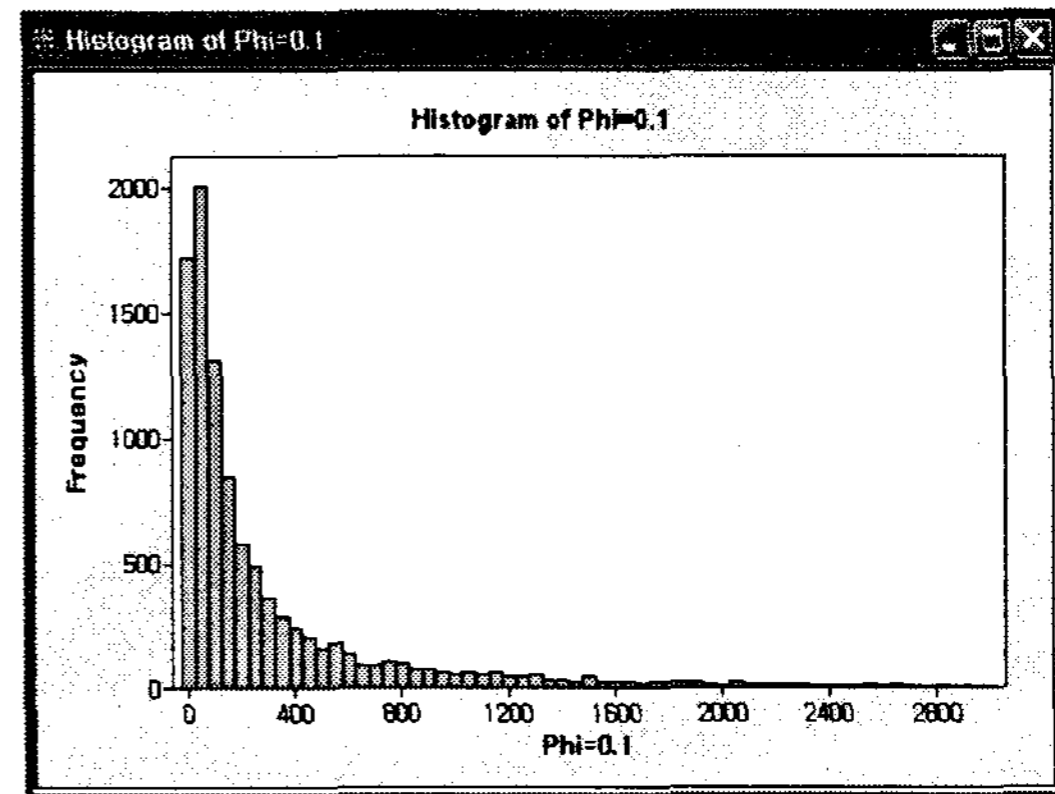
자기상관이 존재하는 개별 자료에 대해 슈와르트 관리도를 적용하여 공정관리를 하는 경우, 일정한 크기의 부분군을 구성하여 그 부분군 평균에 대한 1차 자기상관계수(ρ_1)가 0.1 이하가 되도록 하면 서로 독립인 형태의 자료가 된다고 Runger and Willemain(1995)는 제안하고 있다. 서로 독립인 자료를 사용하는 슈와르트 관리도에서 제1종오류를 고려한 관리상태에서의 평균 런 길이(In-control Average Run Length; ARL_0)는 다음과 같이 구해진다.

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

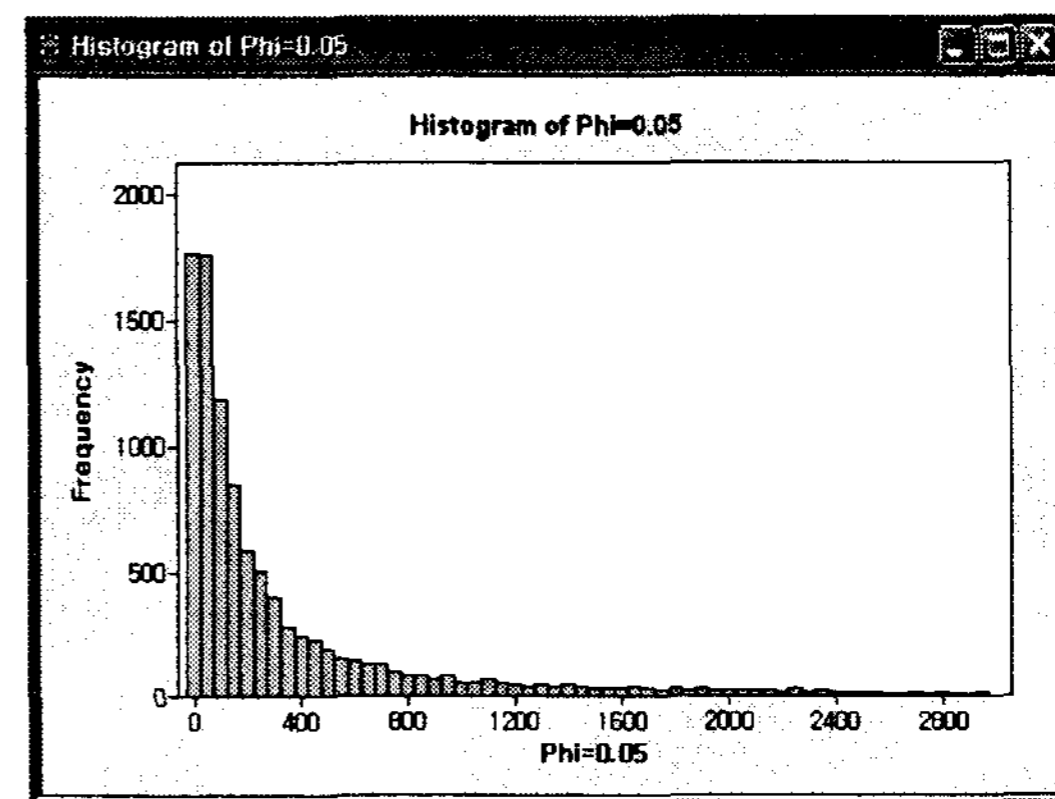
슈와르트 관리도는 일반적으로 정규분포를 가정한 3σ 관리한계를 사용하므로 $\alpha = 0.0027$ 이 된다. 따라서 슈와르트 관리도에서의 ARL_0 은 약 370정도가 되는 것을 알 수 있다. 따라서 부분군 간에 자기상관이 없어지도록 부분군이 구성되었다면 부분군의 평균을 이용한 \bar{X} 관리도 역시 $ARL_0 = 370$ 이 되어야 한다.

<그림 1>과 <그림 2>는 자기상관이 약한, 즉 식(1)의 ϕ 가 0.1 혹은 0.05 정도인 AR(1) 자료들을 생성하고, 그 자료들에 대해 두 개씩을 하나의 부분군으로 묶어 \bar{R}/d_2 를 이용해 σ 를 추정한 후에 개별 자료들에 대한 ARL을 측정한 결과를 보면 1차 자기상관계수가 0.1 이하가 되더라도 ARL_0 가 370보다 작게 나타나는 것을 알 수

있다. $\phi = 0.1$ 인 경우는 ARL의 평균이 322.28로 나타났고, $\phi = 0.05$ 인 경우에는 358.58로 나타났다. 즉, 부분군 내에서 자료가 충분히 독립인 상태가 되지 않았으므로 σ 가 실제보다 작게 추정되었기 때문에 ARL_0 이 370이하로 나타나게 되었음을 알 수 있다. 따라서 자기상관의 크기가 작은 경우라 하더라도 완전히 독립된 자료의 결과를 기대하기는 어렵다.



<그림 1> $\phi = 0.1$ 인 AR(1) 자료의 ARL_0



<그림 2> $\phi = 0.05$ 인 AR(1) 자료의 ARL_0

개별 자료의 자기상관 문제를 해결하기 위해 부분군의 평균을 사용하게 되면 부분군 간에는 자기상관이 해소되는 것을 다음 식을 통해 확인할 수 있다. 예를 들어, 식 (1)의 AR(1) 모형에 대해 부분군의 크기가 b 가 되도록 부분군을 구성한 후에 서로 다른 시차 i 와 시차 j 에서의 부분군 평균(\bar{X}_i, \bar{X}_j) 사이의 상관계수는 다음과 같다.

$$\text{corr}(\bar{X}_i, \bar{X}_j) = \frac{\phi^{(j-i-1)b+1}(1-\phi^b)^2}{(b+(2-b)\phi-2\phi^b)(1-\phi)} \quad (2)$$

식(2)에서 ϕ 는 0과 1사이의 값을 가지므로 부분군의 크기(b)가 커질수록 상관계수는 작아지는 것을 알 수 있다. 따라서 Runger and Willemain (1995)이 언급한 것처럼 부분군의 크기가 커질수록 부분군 평균 간에는 서로 독립인 자료가 되는 것을 알 수 있다. ρ_1 이 0.1 이하가 되어 서로 독립을 가정할 수 있다고 본다면 부분군의 크기를 얼마로 잡아야 하는지가 중요한 문제이다.

부분군의 평균 간에 ρ_1 이 0.1 이하가 되는 부분군의 크기는 AR(1) 모형의 모수인 ϕ 에 따라 다르게 된다. 이것을 알아보기 위하여 AR(1)을 따르는 자료를 만들어 ϕ 가 0.2부터 0.8까지 변하도록 하고, 이 때, 부분군 평균의 1차 자기상관계수가 0.1 이하가 되는 부분군 크기를 구한 것이 <표 1>이다.

<표 1> AR(1) 자료의 최적 부분군 크기

| ϕ | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|--------|-----|-----|-----|-----|-----|-----|-----|
| $b(2)$ | 3 | 4 | 6 | 8 | 11 | 17 | 26 |

위에서 구한 부분군의 크기를 이용하여 부분군을 구성하고, 부분군 평균에 대한 표준오차($\sigma_{\bar{x}}$)를 추정하여 다음과 같이 관리한계를 산출하였다.

$$UCL = \bar{\bar{x}} + 3\hat{\sigma}_{\bar{x}} \quad (3)$$

$$LCL = \bar{\bar{x}} - 3\hat{\sigma}_{\bar{x}}$$

식(3)에서 $\bar{\bar{x}}$ 는 각 부분군 평균들에 대한 총평균이고, $\hat{\sigma}_{\bar{x}}$ 는 두 개의 부분군 평균을 하나로 묶어 \bar{R}/d_2 을 이용하여 산출했다. 여기서 \bar{R} 는 두 개의 부분군 범위들의 평균이다. 현실적으로 공정에서는 충분히 많은 자료를 이용하여 관리한계를 산출하는 것이 어렵다고 보고 본 연구에서는 소표본을 가정하기 위하여 관리한계 산출을 위한 부분군의 개수는 10개로 제한했다. 즉, 10개의 부분군을 이용하여 관리한계를 산출하고

ARL₀을 구한 것이다. 이러한 방법을 사용한 이유는 소표본의 자료를 이용한 부분군 평균에 의한 관리한계 산출 방법이 적절한 것인지를 평가하기 위함이다. 이와 같은 방법으로 식(3)을 활용하여 관리한계를 산출한 후에 3000개의 부분군을 대상으로 \bar{X} 관리도에서 처음 관리이탈이 발생하는 시점을 찾아 ARL₀을 구한 결과가 다음 <표 2>와 같이 나타났다.

<표 2> 최적부분군 자료를 이용한 ARL₀

| 모수 | 부분군 크기(b) | ARL ₀ |
|--------------|---------------|------------------|
| $\phi = 0.2$ | 3 | 331.10 |
| $\phi = 0.3$ | 4 | 338.02 |
| $\phi = 0.4$ | 6 | 334.93 |
| $\phi = 0.5$ | 8 | 339.60 |
| $\phi = 0.6$ | 11 | 333.44 |
| $\phi = 0.7$ | 17 | 337.30 |
| $\phi = 0.8$ | 26 | 333.95 |

10000번의 모의실험을 통해 얻은 <표 2>의 결과를 보면 부분군 평균 간의 ρ_1 이 0.1 이하가 되는 최적 부분군 크기를 이용하여 부분군을 구성했지만 ARL₀은 370보다 작게 나타났음을 알 수 있다.

<표 2>와 같은 결과가 나오게 된 이유는 부분군 평균 간의 1차 자기상관계수가 0.1 이하가 되는 최적 부분군 크기를 산출함에 있어서 평균 부분군 크기를 사용했기 때문이라고 보인다.

ϕ 가 달라짐에 따라 부분군 평균의 ρ_1 이 0.1 이하가 되는 부분군의 크기는 일정한 분포를 갖게 되는데, <표 2>의 모의실험 과정에서 얻은 부분군 크기에 대한 통계량은 다음과 같다.

<표 3> 부분군 크기에 대한 기술통계량

| Descriptive Statistics: Phi=0.2, Phi=0.3, Phi=0.4, Phi=0.5, Phi=0.6, ... | | | | | | |
|--|--------|---------|--------|---------|--------|---------|
| Variable | Mean | SE Mean | StDev | Minimum | Median | Maximum |
| Phi=0.2 | 3.0600 | 0.0155 | 0.4906 | 2.0000 | 3.0000 | 7.0000 |
| Phi=0.3 | 4.5760 | 0.0290 | 0.9171 | 3.0000 | 4.0000 | 15.0000 |
| Phi=0.4 | 6.3710 | 0.0444 | 1.4026 | 4.0000 | 6.0000 | 15.0000 |
| Phi=0.5 | 8.8670 | 0.0812 | 2.5686 | 5.0000 | 8.0000 | 48.0000 |
| Phi=0.6 | 11.952 | 0.111 | 3.496 | 7.000 | 11.000 | 42.000 |
| Phi=0.7 | 17.049 | 0.223 | 7.049 | 9.000 | 15.000 | 123.000 |
| Phi=0.8 | 26.392 | 0.345 | 10.925 | 15.000 | 24.000 | 161.000 |
| Phi=0.9 | 50.477 | 0.639 | 20.186 | 25.000 | 45.000 | 192.000 |

<표 3>의 결과를 통해 보면 <표 2>에서 구한 최적 부분군 크기는 부분군 간의 ρ_1 이 0.1이 되기 위한 평균적인 부분군 크기임을 알 수 있다. 따라서 실제 부분군 평균 간에 독립인 상태가 되기 위한 부분군의 크기는 <표 2>에서 제기한 결과보다 커야만 된다는 것을 알 수 있다.

<표 2>의 모의실험 과정에서 얻은 부분군 크기의 자료들에 대한 90백분위수, 95백분위수를 구한 것이 <표 4>와 <표 5>에 나와 있다.

<표 4> 최적 부분군 크기의 90백분위수와 ARL

| ϕ | b | ARL ₀ | p(%) |
|--------|-----|------------------|------|
| 0.2 | 4 | 355.8 | 9910 |
| 0.3 | 6 | 384.1 | 9841 |
| 0.4 | 8 | 378.4 | 9627 |
| 0.5 | 12 | 379.1 | 9817 |
| 0.6 | 16 | 379.7 | 9705 |
| 0.7 | 22 | 385.5 | 9434 |
| 0.8 | 37 | 366.5 | 9643 |

<표 5> 최적 부분군 크기의 95백분위수와 ARL

| ϕ | b | ARL ₀ | p(%) |
|--------|-----|------------------|------|
| 0.2 | 4 | 362.5 | 9916 |
| 0.3 | 6 | 379.5 | 9835 |
| 0.4 | 9 | 386.2 | 9893 |
| 0.5 | 13 | 375.8 | 9924 |
| 0.6 | 18 | 377.4 | 9902 |
| 0.7 | 26 | 382.4 | 9875 |
| 0.8 | 42 | 371.8 | 9919 |

<표 4>와 <표 5>의 결과를 보면 90백분위수를

사용한 경우에는 최적 부분군 크기를 이용한 경우 약 94% ~ 99% 정도의 모의실험 자료의 부분군 평균의 1차 자기상관계수(ρ_1)가 0.1 이하가 된다. 95백분위수를 사용한 경우에는 99% 이상의 모의실험 자료가 포함되는 것으로 나타났다. 부분군 평균에 대한 ARL₀ 역시 90백분위수와 99백분위수의 최적 부분군 크기를 활용한 경우 ARL₀가 370에 근접한 것으로 나타났다. 특히 <표 4>보다는 <표 5>가 더 좋은 결과를 나타냈음을 알 수 있는데 부분군의 크기가 클수록 ARL₀이 370에 더 가까워짐을 알 수 있다. 따라서 부분군 평균을 이용한 공정관리에서 부분군 평균 간에 자기상관을 해소하기 위한 최적의 부분군 부분군의 크기는, AR(1)을 따르는 자료인 경우에는 각 ϕ 별로 구해진 부분군의 90백분위수 혹은 95백분위수로 정하는 것이 바람직하다.

3. 자기상관 자료의 σ 추정

3.1 부분군의 평균을 이용한 산포 추정

자기상관이 존재하는 공정에서 자기상관을 무시하고 독립을 가정하여 일반적인 슈와르트 관리도를 적용하게 되면 σ 가 과소추정 되는 문제가 발생하는 것을 앞 절에서 밝힌 바 있다. 그런데 이러한 자기상관 자료의 모형을 알고 있거나 충분히 자료가 많이 있는 경우라면 σ 추정이 문제가 되지 않는다. 하지만 주어진 자료의 모형도 모르고, 소량의 표본만이 주어진 경우에는 σ 추정이 어렵게 된다. 이러한 σ 추정의 문제는 현실적으로 품질관리 분야에서 부딪히게 되는데 신제품의 사양산 과정에서 공정능력분석을 실시하는 경우 충분히 많은 자료를 확보하여 σ 를 추정하는 것은 어렵다. 더욱이 자료가 서로 독립인

경우에도 소표본을 이용한 σ 추정의 정확도가 떨어지겠지만 자기상관이 존재하는 경우라면 그 문제는 더욱 심각하다. 이러한 자기상관이 존재하는 AR(1) 자료들에 대한 σ 추정 방법에 대해서 Shore(1999)는 다음과 같은 방법을 제시하였다.

<그림 3>처럼 ($b \times m$)개의 자료에 대해 부분

군 평균들이 서로 독립이 되는 부분군의 크기 (b)를 정하고, 이러한 부분군을 m 개를 취한다.

<그림 3> 부분군의 구성

$$\begin{aligned} (x_1, x_2 \cdots x_b) &\rightarrow \bar{x}_1 \\ (x_{b+1}, x_{b+2} \cdots x_{2b}) &\rightarrow \bar{x}_2 \\ &\vdots \\ (x_{(m-1)b+1}, x_{(m-1)b+2} \cdots x_{(m-1)2b}) &\rightarrow \bar{x}_m \end{aligned}$$

m 개의 각 부분군 평균, \bar{x} 을 2개씩 묶게 되면 부분군 평균 간에는 서로 독립인 상태가 되므로 범위(R)를 이용한 $(\bar{R}/d_2)^2$ 를 이용하여 부분군 평균에 대한 분산, $Var(\bar{X})$ 를 추정하게 된다. 이렇게 추정된 $Var(\bar{X})$ 는 식(4)에 나와 있는 AR(1) 자료의 부분군 평균에 대한 분산과 같게 된다. 따라서 Shore(1997)가 제시한 방법에 따라 부분군 평균의 분산, $Var(\bar{X})$ 을 정리하여 식(5)와 같이 σ_x 에 대해 정리하면 실제 추정하고자 하는 공정 산포인 σ_x 를 추정할 수 있게 된다.

$$Var(\bar{X}) = \left(\frac{\sigma_x^2}{b} \right) \left[1 + \frac{2}{b} \sum_k^{b-1} (b-k)\rho_k \right] \quad (4)$$

$$\hat{\sigma}_x = \sqrt{\frac{b \cdot Var(\bar{X})}{1 + \frac{2}{b} \sum_k^{b-1} (b-k)\rho_k}} \quad (5)$$

여기서 ρ_k 는 시차(Lag) k 에서의 자기상관계수이다.

오차항이 백색잡음(White Noise)인 AR(1)의 자료에서 실제 산포는 다음과 같은 식으로 표현된다.

$$\sigma_x = \sqrt{\frac{\sigma_a^2}{1-\phi^2}} \quad (6)$$

3.2 모의실험: Shore의 방법을 통해 추정한 산포와 실제 산포의 비교

식(6)에서 정의된 결과와 Shore의 방법에 의한 결과를 비교하기 위한 모의실험을 다음과 같이 실시하였다.

모의실험은 먼저 AR(1)을 따르는 길이가 $(m \times b)$ 인 자료를 만든다. 여기서 m 은 부분군의 개수이며, b 는 <표 5>에서 구한 최적 부분군 크기를 의미한다. 본 모의실험은 소표본을 가정하기 위하여 부분군의 개수(m)를 6과 10으로 나누어 수행하였다. 그 이유는 자료의 길이가 충분히 길다면 정상 시계열 모형 AR(1)에서는 표본의 표준편차, $s = \sqrt{(x_i - \bar{x})^2 / (n-1)}$ 을 이용하여 σ_x 를 추정하는 훨씬 효과적이기 때문이다.

Shore(1999)가 제시한 식(5)을 이용해서 σ_x 를 추정하기 위해서는 부분군 평균들이 서로 독립이 되는 부분군의 크기(b)가 필요하고, 또한 어느 시점까지의 자기상관계수(ρ_k)를 사용할 것인가가 매우 중요하다. 식(5)에서는 $(b-1)$ 개의 자기상관계수(ρ_k)를 사용한다고 표현되어 있지만 대개 통계적으로 유의한(Significant) 시차(Lag)까지의 자기상관계수(ρ_k)를 사용한다고 알려져 있다. ϕ 가 0.8인 경우는 유의한 시차가 10 정도인데 소표본을 이용한 <표 6>과 <표 7>의 결과를 보면 유의한 시차까지의 자기상관계수(ρ_k)만을 사용하게 되면 실제 참 산포, σ_x 와 차이가 커진다. 특히 AR(1) 모형에서의 모수인 ϕ 가 커질수록 더 많은 시차를 이용해서 추정하는 것이 참값과 근접함을 볼 수 있다. 부분군의 개수, $m=6$ 인 <표 6>의 결과 보다 $m=10$ 인 <표 7>을 보면 사용하는 시차가 많이 줄어들음을 알 수 있다. 따라서 부분군의 개수(m)가 충분히 많다면 통계적으로 유의한 시차만을 사용하는 것이 옳은 것으로 판단되지만 소표본의 자료에서는 좀 더 많은 시차까지의 자기상관계수(ρ_k)을 이용하여 σ_x 를 추정하는 것이 바람직해 보인다.

<표 6> $m = 6$ 일 때, Shore 방법에 의한 σ_x 추정치 비교

| ϕ | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|------------|------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| σ_x | 1.0206 | 1.0483 | 1.0911 | 1.1547 | 1.2500 | 1.4003 | 1.6667 |
| s | 0.9978 | 1.0258 | 1.0698 | 1.1320 | 1.2241 | 1.3687 | 1.6302 |
| Lag(k) | Shore 방법에 의한 추정치 | | | | | | |
| 2 | 1.0097 | 1.0565 | 1.1256 | 1.2266 | 1.4030 | 1.7128 | 2.3817 |
| 3 | <u>1.0063</u> | <u>1.0443</u> | 1.1019 | 1.1867 | 1.3312 | 1.5829 | 2.1364 |
| 4 | | 1.0406 | <u>1.0916</u> | 1.1677 | 1.2972 | 1.5154 | 1.9937 |
| 5 | | 1.0397 | 1.0868 | <u>1.1575</u> | 1.2782 | 1.4768 | 1.9033 |
| 6 | | | 1.0847 | 1.1517 | 1.2665 | 1.4528 | 1.8433 |
| 7 | | | 1.0838 | 1.1484 | 1.2588 | 1.4369 | 1.8021 |
| 8 | | | 1.0835 | 1.1466 | 1.2537 | 1.4257 | 1.7730 |
| 9 | | | | 1.1456 | <u>1.2504</u> | 1.4175 | 1.7516 |
| 10 | | | | 1.1451 | 1.2481 | 1.4114 | 1.7353 |
| 11 | | | | 1.1448 | 1.2467 | 1.4069 | 1.7227 |
| 12 | | | | 1.1448 | 1.2458 | 1.4034 | 1.7126 |
| 13 | | | | | 1.2453 | <u>1.4007</u> | 1.7045 |
| 14 | | | | | 1.2449 | 1.3988 | 1.6979 |
| 15 | | | | | 1.2447 | 1.3973 | 1.6924 |
| 16 | | | | | 1.2446 | 1.3963 | 1.6878 |
| 17 | | | | | 1.2446 | 1.3955 | 1.6840 |
| 18 | | | | | | 1.3949 | 1.6807 |
| 19 | | | | | | 1.3944 | 1.6780 |
| 20 | | | | | | 1.3941 | 1.6757 |
| 21 | | | | | | 1.3939 | 1.6739 |
| 22 | | | | | | 1.3937 | 1.6723 |
| 23 | | | | | | 1.3936 | 1.6710 |
| 24 | | | | | | 1.3936 | 1.6699 |
| 25 | | | | | | 1.3935 | 1.6689 |
| 26 | | | | | | | 1.6682 |
| 27 | | | | | | | 1.6676 |
| 28 | | | | | | | 1.6670 |
| 29 | | | | | | | <u>1.6666</u> |
| 30 | | | | | | | 1.6662 |
| 31 | | | | | | | 1.6660 |
| 32 | | | | | | | 1.6657 |
| 33 | | | | | | | 1.6655 |
| 33 | | | | | | | 1.6654 |
| 34 | | | | | | | 1.6653 |
| 35 | | | | | | | 1.6652 |
| 36 | | | | | | | 1.6651 |
| 37 | | | | | | | 1.6651 |
| 38 | | | | | | | 1.6651 |
| 39 | | | | | | | 1.6650 |
| 40 | | | | | | | 1.6650 |

<표 7> $m = 10$ 일 때, Shore 방법에 의한 σ_x 추정치 비교

| ϕ | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|------------|------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| σ_x | 1.0206 | 1.0483 | 1.0911 | 1.1547 | 1.2500 | 1.4003 | 1.6667 |
| s | 1.0071 | 1.0345 | 1.0785 | 1.1420 | 1.2334 | 1.3794 | 1.6427 |
| Lag(k) | Shore 방법에 의한 추정치 | | | | | | |
| 2 | <u>1.0073</u> | <u>1.0498</u> | 1.1206 | 1.2233 | 1.3865 | 1.6944 | 2.3685 |
| 3 | 1.0030 | 1.0374 | 1.0974 | 1.1819 | 1.3122 | 1.5616 | 2.1220 |
| 4 | | 1.0332 | <u>1.0870</u> | 1.1626 | 1.2776 | 1.4912 | 1.9777 |
| 5 | | 1.0320 | 1.0821 | <u>1.1519</u> | 1.2588 | 1.4512 | 1.8855 |
| 6 | | | 1.0797 | 1.1456 | <u>1.2471</u> | 1.4267 | 1.8236 |
| 7 | | | 1.0787 | 1.1418 | 1.2394 | 1.4108 | 1.7807 |
| 8 | | | 1.0784 | 1.1397 | 1.2341 | <u>1.3996</u> | 1.7503 |
| 9 | | | | 1.1384 | 1.2305 | 1.3914 | 1.7282 |
| 10 | | | | 1.1377 | 1.2280 | 1.3853 | 1.7116 |
| 11 | | | | 1.1374 | 1.2263 | 1.3806 | 1.6987 |
| 12 | | | | 1.1373 | 1.2252 | 1.3771 | 1.6886 |
| 13 | | | | | 1.2245 | 1.3744 | 1.6804 |
| 14 | | | | | 1.2240 | 1.3723 | 1.6737 |
| 15 | | | | | 1.2237 | 1.3707 | <u>1.6680</u> |
| 16 | | | | | 1.2235 | 1.3694 | 1.6633 |
| 17 | | | | | 1.2235 | 1.3685 | 1.6593 |
| 18 | | | | | | 1.3678 | 1.6560 |
| 19 | | | | | | 1.3673 | 1.6531 |
| 20 | | | | | | 1.3669 | 1.6507 |
| 21 | | | | | | 1.3667 | 1.6486 |
| 22 | | | | | | 1.3665 | 1.6468 |
| 23 | | | | | | 1.3664 | 1.6453 |
| 24 | | | | | | 1.3663 | 1.6440 |
| 25 | | | | | | 1.3663 | 1.6430 |
| 26 | | | | | | | 1.6421 |
| 27 | | | | | | | 1.6413 |
| 28 | | | | | | | 1.6407 |
| 29 | | | | | | | 1.6401 |
| 30 | | | | | | | 1.6397 |
| 31 | | | | | | | 1.6393 |
| 32 | | | | | | | 1.6390 |
| 33 | | | | | | | 1.6388 |
| 34 | | | | | | | 1.6386 |
| 35 | | | | | | | 1.6384 |
| 36 | | | | | | | 1.6383 |
| 37 | | | | | | | 1.6382 |
| 38 | | | | | | | 1.6381 |
| 39 | | | | | | | 1.6381 |
| 40 | | | | | | | 1.6380 |
| 41 | | | | | | | 1.6380 |

4. 결 론

앞선 모의실험의 결과처럼 자기상관이 존재하는 1차 자기회귀 모형의 자료들을 슈와르트 관리도에 적용하여 공정관리를 하는 경우에 부분군을 구성하여 자기상관을 해소하는 방법을 사용할 수 있으나, 이 때 적당한 부분군의 크기를 잡는 것이 매우 중요하다. 즉, 부분군 평균 간에 독립이 되기 위한 1차 자기상관계수(ρ_1)가 0.1 이하가 되는 부분군의 크기를 결정하는데 있어서 추정된 부분군 크기 자료의 90백분위수 혹은 95백분위수를 사용하게 되면 슈와르트 관리도에서 요구하는 $ARL_0 = 370$ 을 만족하게 된다. 또한 공정능력분석 필요한 σ_x 추정을 위해서 제시된 Shore(1999)의 방법을 활용하기 위해서는 앞서 구한 최적 부분군 크기와 더불어 적절한 시점 k 까지의 자기상관계수(ρ_k)를 결정하는 것이 필요하다.

본 연구에서는 AR(1)을 따르는 자기상관 자료들에 대한 최적 부분군 크기에 대해 논의했는데 향후 다른 형태의 자기상관 자료들에 대한 최적 부분군 크기를 결정하는 보다 강건한(Robust) 방법이 연구되어야 할 것이다.

참고 문헌

- [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] Alwan, L. C.(1992), "Effects of Autocorrelation on Control Chart Performance", *Communications in Statistics-Theory and Methods*, Vol. 21, pp. 1025-1049.
- [3] Haim Shore(1997), "Process Capability Analysis When Data are Autocorrelated", *Quality Engineering*, Vol. 9, pp. 615-626.
- [4] Kang, K., and Schmeiser, Bruce.(1987), "Properties of Batch Means From Stationary ARMA Time Series", *Oper. Res.* Vol.6, pp 19-24.
- [5] Montgomery, D. C. and Mastrangelo, C. M.(1991), "Some Statistical Process Control Methods for Autocorrelated Data", *Journal of Quality Technology*, Vol. 23, pp. 179-193.
- [6] Montgomery, D. C.(2005), "Introduction to Statistical Quality Control", 5th ed., New York: Wiley.
- [7] Runger, G. C. and Willemain, T. R.(1995), "Model-based and Model-free Control of Autocorrelated Process", *Journal of Quality Technology*, Vol. 27, pp. 283-292.
- [8] Runger, G. C. and Willemain, T. R.(1996), "Batch-Means Control Charts for Autocorrelated Data", *IIE Transactions*, Vol. 28, pp. 483-487.
- [9] Schmid, W. and Schone, A.(1997), "Some Properties of the EWMA control chart in the presence of autocorrelation", *The Annals of Statistics*, Vol. 25, pp. 1277-1283.
- [10] Zhang, N. F.(1998), "A Statistical Control Chart for Stationary Process Data", *Technometrics*, Vol. 40, pp. 24-38.