Journal of the Korean Data & Information Science Society
2010, **21**(6), 999–1008 한국데이터정보과학회지

가변추출간격을 갖는 다변량 슈하르트 관리도[†]

조교영[1]

[1]경북대학교 통계학과

접수 2010년 5월 10일, 수정 2010년 9월 30일, 게재확정 2010년 10월 5일

요 약

공정을 모니터링 하기 위한 전통적인 관리도는 표본들 사이의 일정한 추출간격에서 일정한 수의 표본을 취하여 만들어 지는 고정추출율 관리도이다. 본 연구의 목표는 표준적인 고정추출율을 갖는 다변량 관리도에 비하여 성능이 우수한 가변추출간격을 갖는 다변량 관리도를 개발하는데 있다. 대부분의 다변량 관리도에 대한 연구는 공정의 평균벡터를 모니터링 하는데 초점이 맞추어져 있다. 그러나 본 논문에서는 공정의 평균벡터와 분산-공분산을 동시에 모니터링 하기 위한 다변량 관리도를 연구한다. 가변추출간격을 갖는 다변량 슈하르트 관리도에 대하여 연구 하고자 한다.

주요용어: 가변추출간격, 다변량 관리도, 정상상태 평균경보시간, 통계적 공정관리, 평균경보시간.

1. 서론

공정에 모니터링을 적용할 때 많은 공정의 상태가 여러 변수들에 의하여 특성화 된다. 이러한 상황에서는 다변량 공정관리도를 사용 하는 것이 적절하다. 본 연구에서는 관심있는 공정이 p개의 연속변수에 의하여 특성화 되고, 변수들의 결합분포는 평균벡터가 μ이고 공분산행렬이 Σ인 다변량 정규분포인 경우를 고려한다. 공정 모니터링의 목적은 평균벡터 μ와 공분산행렬 Σ를 변화시키는 이상 원인을 탐지하는데 있다.

최초의 다변량 관리도는 Hotelling (1947)이 개발한 슈하르트 (Shewhart) 형태의 관리도이다. 다변량 모니터링과 다변량 슈하르트 형태의 관리도에 대한 일반적인 논의는 Alt (1985), Wierda (1994), Lowry와 Montgomery (1995), Montgomery (2005)에 있다.

다변량관리도에 대한 대부분의 연구는 평균벡터를 모니터링 하는 문제에 집중되어 있다. 또한 공분산행렬을 모니터링 하기 위한 다변량관리도에 대한 연구도 있다. 본 논문은 평균벡터와 공분산행렬을 동시에 모니터링 하기 위한 연구이다.

공정을 모니터링 하기 위한 전통적인 접근방법은 표본간의 추출간격과 표본크기가 고정된 표본을 추출한다. 그러나. 공정의 변화를 탐지하는 능력은 전통적인 고정추출율 (Fixed Sampling Rate: FSR) 관리도 대신에 가변추출율 (Variable Sampling Rate: VSR) 관리도를 사용함으로써 향상되어질 수 있다. 가변추출율 관리도의 기본적인 아이디어는 공정에서 변화의 징조가 있을 때 표본 추출율이 증가하고 변화의 징조가 없을 때 추출율이 감소하는 것이다. 공정의 변화에 대한 징조가 충분히 강하면 가변추출율 관리도는 표준 고정추출율 관리도와 같은 방식으로 경보를 울린다.

† 이 논문은 2009년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2009-0076310).
1 (702-701) 대구광역시 북구 산격동 1370번지, 경북대학교 통계학과, 교수. E-mail: gycho@knu.ac.kr

가변추출간격 (Variable Sampling Interval: VSI) 관리도는 표본추출간격이 자료의 함수로서 변하기 때문에 가변추출율 관리도의 특수한 형태이다. 특히, 공정에서 문제의 징조가 있으면 다음 표본 추출시 짧은 추출간격을 사용하고 문제의 징조가 없으면 긴 추출간격을 사용한다. 고정추출율 관리도의 또 다른 형태는 표본크기가 자료의 함수로써 변하는 가변표본크기 (Variable Sample Size: VSS) 관리도이다.

가변추출율 관리도는 대부분 단변량인 경우에 개발되어져 있다. Tagaras (1998)의 리뷰 논문은 1997년까지 발표된 논문을 포함하고 있다. 평균과 분산을 동시에 모니터링 하는 문제를 다룬 최근의 단변량 가변추출간격 관리도에 대한 연구는 Reynolds와 Stoumbos (2001), Stoumbos와 Reynolds (2005), Zhang과 Wu (2006)이다.

가변추출율 다변량 관리도에 대한 연구로는 다변량 가변추출간격 관리도에 대한 Aparisi와 Haro (2001), 다변량 가변표본크기 관리도에 대한 Aparisi (1996), 가변추출간격과 가변추출크기를 동시에 사용한 다변량 슈하르트 관리도에 대한 Aparisi와 Haro (2003)가 있다.

본 논문의 목적은 평균벡터와 공분산행렬을 동시에 모니터링하는 문제를 위한 다변량 가변추출간격 관리도를 개발하는데 있다. 여기서는 다변량 가변추출간격 슈하르트 관리도를 연구한다.

공정 모니터링의 목적은 공정 모수들의 작은 이동 (shift) 뿐만 아니라 큰 이동도 탐지하는데 있다고 가정하자. 다변량 셋팅에서 모수 변화의 방향을 생각하는 것은 필수적이다. 평균벡터 μ에 대해서는 주 관심이 특별한 이동 방향에 있는 것이 아니므로 모든 방향에서 성과 (performance) 평균을 고려한다. 공분산 Σ에서 이동인 경우 변동 (variability)의 증가에 초점을 둔다. 여기서 고려한 유사한 관리도가 변동의 감소를 탐지하는데 사용될 수 있다 (Reynolds와 Cho, 2006). 그러나 본 논문에서는 변동의 감소를 탐지하는 문제는 다루지 않는다.

2. 다변량 슈하르트 관리도

p개 변수들의 평균벡터를 μ, 공분산행렬을 Σ, 표준편차들 (공분산행렬 Σ의 대각원소의 제곱근)의 벡터를 $\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_p)'$라 하자. 또한 μ_0, Σ_0, σ_0을 각각 μ, Σ, σ의 관리상태의 값이라 하자. 여기서 관리상태에서의 모수값들은 알려졌거나 추정과 관련된 오차는 무시할 수 있을 정도로 정확하게 추정된다고 가정한다.

공정은 각 표본추출 시점에서 취해진 크기가 $n \geq 1$인 관측벡터들을 표본으로 사용함으로써 모니터링 된다고 가정한다. 사전 연구에서, 가변추출간격 관리도는 단지 두 개의 가능한 추출간격을 사용하는 것이 효율적임을 보였다 (Reynolds 등, 1998; Reynolds, 1995; Stoumbos 등, 2001). 두 개의 가능한 추출간격 사용은 가변추출간격 관리도의 적용을 단순화 한다. 따라서 두 개의 추출간격을 갖는 가변추출간격 관리도를 고려한다.

d_1과 d_2 $(0 < d_1 < d_2)$를 두 개의 가능한 추출간격이라 하고 d를 공정의 관리 상태에서 평균추출간격이라 하자. 가변추출간격 관리도와 고정추출율 관리도를 비교할 때 d를 고정추출율 관리도의 추출간격으로 사용한다. 예를 들어 현재 사용 중인 고정추출율 관리도가 매 $d = 4$시간 마다 표본 크기 $n = 4$에 기반 한다고 하자. 그러면 가변추출간격 관리도는 추출간격으로 $d_1 = 1$ 혹은 $d_2 = 7$을 사용할 것이다. 여기서 가변추출간격 관리도는 관리상태 하에서 평균 추출 간격 $d = 4$이 되도록 설계되어 진다.

표본 추출 시점 k에서, X_{kij}를 변수 i $(i = 1, 2, \cdots, p)$에 대한 j $(j = 1, 2, \cdots, n)$번째 관측이라 하고, 대응되는 표준화된 관측을 다음과 같이 두자.

$$Z_{kij} = (X_{kij} - \mu_{0i})/\sigma_{0i}$$

여기서 μ_{0i}은 벡터 μ_0의 i번째 원소이고, σ_{0i}는 벡터 σ_0의 i번째 원소이다. 표본 추출 시점 k에서, 관측 벡터 j에 대한 표준화된 관측벡터를 다음과 같이 두자.

$$z_{kj} = (Z_{k1j}, Z_{k2j}, \cdots, Z_{kpj}), \quad j = 1, 2, \cdots, n. \tag{2.1}$$

Σ_Z를 z_{kj}의 공분산 행렬, Σ_{Z0}를 관리상태 하에서 Σ_Z의 값이라 하자. 공정의 관리상태 하에서 Z_{kij}가 표준정규분포를 가질 때 Σ_{Z0}는 원래 관측에 대한 관리상태 하에서 상관계수행렬이 된다.

$\bar{X}_{ki} = \sum_{j=1}^{n} X_{kij}/n$를 표본 추출 시점 k에서 변수 i에 대한 표본 평균이라 두고, 표준화된 표본 평균을 다음과 같이 정의 한다.

$$Z_{ki} = \sqrt{n}(\bar{X}_{ki} - \mu_{0i})/\sigma_{0i}, \quad i = 1, 2, \cdots, p.$$

공정이 관리상태 하에서 Z_{ki}는 표준정규분포를 갖는다. 따라서 Σ_{Z0}는 $(Z_{k1}, Z_{k2}, \cdots, Z_{kp})$의 관리상태 하에서 상관계수행렬이 된다.

Reynolds와 Cho (2006)에서 사용한 바와 같이, 슈하르트 관리도를 간단히 기호 "S"로 쓴다. 평균벡터 μ를 모니터링 하기 위하여 Hotelling (1947)에 의하여 제안된 슈하르트 형태의 관리도는 원래 Σ_0이 알려지지 않은 경우에 대하여 개발되었다. 그러나 만약 Σ_0이 알려 졌다고 가정하면 이 관리도는 다음과 같은 통계량에 기초한 관리도와 대등하다.

$$S_k^Z = (Z_{k1}, Z_{k2}, \cdots, Z_{kp})\Sigma_{Z0}^{-1}(Z_{k1}, Z_{k2}, \cdots, Z_{kp})'. \tag{2.2}$$

여기서 위 첨자 "Z"는 관리도가 Z_{ki}에 근거함을 나타낸다. 만약 $S_k^Z > h$이면 관리도는 경보를 울린다. 여기서 h는 관리상한 (upper control limit)이다. 이러한 관리도를 "SZ"관리도라 하자 ("S"는 Shewhart, "Z"는 Z_{ki}에 근거). 산업에서 관리도의 적용에 대한 것은 Mason과 Young (2002)을 참고하기 바란다.

SZ관리도의 가변추출간격 버전은, 만약 $S_k^Z \le g$이면 긴추출간격 d_2를 사용하고, 만약 $g < S_k^Z \le h$이면 짧은추출간격 d_1를 사용한다. 여기서 $0 < g \le h$이다. 관리상태 하에서 S_k^Z의 분포는 자유도 p를 갖는 카이제곱 분포이다. 따라서 g와 h는 카이제곱 분포의 분위수를 사용하여 결정 되어질 수 있다.

공정의 변동을 모니터링 하는 문제에 대하여, 본 논문에서는 목표로부터 관측까지의 편차의 제곱에 근거한 관리도에 초점을 맞춘다. 이런 형태의 관리도는 변동을 모니터링 하는 문제 혹은 평균과 변동 둘 다를 모니터링 하는 하나의 관리도 문제에 대하여 이미 조사되어져 있다 (Reynolds와 Ghosh, 1981; Domangue와 Patch, 1991; Macgregor와 harris, 1993; Shamma와 Amin, 1993).

표본 추출 시점 k에서, 변수 i에 대한 목표로부터 편차의 표준화 제곱을 다음과 같이 두자.

$$Z_{ki}^2 = \sum_{j=1}^{n} Z_{kij}^2/n, \quad i = 1, 2, \cdots, p. \tag{2.3}$$

Reynolds와 Cho (2006)는 관리상태 하에서 통계량 (2.3)의 공분산행렬이 $2\ \Sigma_{Z0}^{(2)}/n$임을 보였다. 여기서 $\Sigma_{Z0}(2)$의 원소는 공분산행렬 Σ_{Z0}에 대응하는 원소의 제곱인 행렬을 나타낸다.

$$S_k^{Z^2} = n(2)^{-1}(Z_{k1}^2, Z_{k2}^2, \cdots, Z_{kp}^2)(\Sigma_{Z0}^{(2)})^{-1}(Z_{k1}^2, Z_{k2}^2, \cdots, Z_{kp}^2)'. \tag{2.4}$$

여기서 위 첨자 "Z^2"는 관리도가 $Z_{ki}{}^2$에 근거함을 나타낸다. 만약 $S_k^{Z^2} > h$이면 관리도는 경보를 울린다. 여기서 h는 관리상한 (upper control limit)이다. 가변추출간격 버전은, 만약 $S_k^{Z^2} \le g$이면 긴추출간격 d_2를 사용하고, 만약 $g < S_k^{Z^2} \le h$이면 짧은추출간격 d_1를 사용한다. 여기서 $0 < g \le h$이다. 이러한 관리도를 "SZ^2"관리도라 하자

SZ관리도는 평균벡터 μ의 변화를 탐지하기 위하여 디자인 되었으나 표본크기 $n = 1$일 때 분산벡터 σ에서 큰 증가를 탐지하는데도 효율적임이 입증되었다. 만약 슈하르트 관리도가 사용되면, $n = 1$인 경우는 SZ관리도 자체가 분산벡터 σ에서 큰 증가를 탐지하는데 사용되어 질 수 있다. 그러나, $n > 1$일 때 SZ관리도는 분산벡터 σ에서 증가를 탐지하는데 효율적이지 않을 것이다. 그래서 분산벡터 σ에서 증가를 탐지하는데 다른 관리도가 사용되어 져야만 한다. 그것이 현재의 샘플링시점에서 편차들의 제곱합에 근거한 슈하르트 SZ^2관리도이다.

SZ관리도가 Σ의 이동에 민감하다 할지라도 보통 평균벡터 μ를 모니터링하기위한 관리도로 언급된다. 비슷하게 목표로부터 편차제곱에 근거한 관리도는 평균벡터 μ의 이동에 민감할지라도 공분산행렬 Σ를 모니터링하기위한 관리도로 언급된다. 목표로부터 편차제곱에 근거한 관리도는 Σ에서 변화를 탐지하는데 매우 효과적이나, 평균벡터 μ의 큰 이동에도 매우 민감하다는 중요한 이점을 가지고 있다.

3. 관리도의 성능 측도

서로 다른 관리도들의 비교는 공정모수의 여러 가지 이동을 탐지하는데 걸리는 평균시간으로 한다. 서로 다른 관리도의 평균 탐지 시간의 공정한 비교를 위하여 각 관리도가 시간당 같은 오경보율 (false alarm rate)과 같은 추출율 (sampling rate)을 가져야 한다. 따라서 평균탐지시간, 오경보율과 추출율에 대한 측도가 필요하다.

평균경보시간 (Average Time to Signal: ATS)은 공정모니터링의 시작부터 경보가 발생할 때까지의 평균시간이다. 평균경보시간은 시간당 오경보율에 대한 하나의 측도이다. 특히, 재생과정 (renewal process)으로 모형화 될 수 있는 관리도에 대하여, 관리상태에서 1/ATS는 시간당 극한 평균오경보율 (long run average false alarm rate per unit time)이다. 비교할 관리도들이 시간당 같은 오경보율에 대한 요건은 그들이 관리상태 하에서 같은 ATS를 갖는 것이다.

가변추출 관리도에서 어떤 주어진 구간에서 표본의 수는 확률변수이다. 따라서 서로 다른 관리도들의 공정한 비교를 위해 평균추출율을 고려해야 한다. 표본크기 n이 서로 다른 관리도의 비교를 위하여 관측수로서 나타내는 추출율의 측도가 필요하다. 평균경보관측수 (Average Number of Observations to Signal: ANOS)를 공정모니터링의 시작부터 경보가 발생할 때까지의 평균관측수로 정의한다. 재생과정 (renewal process)으로 모형화 될 수 있는 관리도에 대하여, 관리상태 하에서 ANOS/ATS는 시간당 극한 평균추출율 (long run average sampling rate per unit time)이다. 만약 비교할 관리도들이 관리상태 하에서 같은 ATS와 같은 ANOS를 갖는다면 그들은 관리상태 하에서 같은 평균추출율을 가지게 된다.

공정모수에서 이동 (shift)이 있을 때, ATS는 공정모니터링이 출발할 때 이동이 발생한다고 가정하면 이동에 대한 탐지시간의 측도로 적절하다. 그러나 실제로 어떤 이동은 모니터링출발 이후 미래 랜덤시간에 발생할 것이다. 이런 경우 탐지시간에 대한 적절한 측도는 이동이 발생한 시간의 랜덤시점부터 관리도가 경보를 울리는 시간의 평균길이이다.

본 논문에서, 관리도들은 관리상태 하에서 같은 ATS와 같은 ANOS값을 가지도록 설정되어 있다. 그러면 SSATS는 공정모수에서 이동을 탐지할 수 있는 측도로 사용되어질 수 있다. ATS와 SSATS를 계산하기 위하여 1,000,000번의 런의 시뮬레이션이 사용되었다. SSATS는 관리상태 하에서 400번의 관측벡터에 대한 관리도 작동을 시뮬레이션 한 후 평균벡터 μ와 표준편차벡터 σ에서 이동하에서 시뮬레이션을 함으로써 구한다.

4. 비교를 위한 모수선택

본 논문에서 수치결과는 표본크기 $n=1$과 $n=4$에 대한 것이다. $n=1$인 경우 평균추출구간은 시간 단위이다. 주어진 결과를 어떤 시간 단위에 적용할지라도 편의상 한 시간으로 한다. $n=4$인 경우 평균추출구간은 4시간 단위이다. 평균추출구간과 관련하여 표본크기 $n=1$과 $n=4$는 시간당 하나의 관측 평균추출비에 해당하므로 두 표본크기에 공정한 비교가 가능하다.

VSI 관리도 평가에서 여러 개의 d_1과 d_2를 선택한다. VSI 관리도에 대한 연구결과 d_1은 가능한 작은 것이 바람직하다. 그러나 d_2의 선택은 작은 이동 혹은 큰 이동에 따라 달라진다. 특히, d보다 조금 큰 d_2의 선택은 큰 이동을 탐지하는데 효과적이고, 큰 d_2의 선택은 작은 이동 혹은 중간크기의 이동을 탐지하는데 더욱 효과적이다. 본 논문에서, 큰 이동 탐지를 위하여 $d_1=0.1d$와 $d_2=1.25d$, 작은 이동 탐지를 위하여 $d_1=0.1d$와 $d_2=1.90d$를 고려한다. 또한, 실제 적용시 신속한 표본추출이 용이하지 않는 경우 $d_1=0.25d$와 $d_2=1.25d$그리고 $d_2=1.90d$를 고려한다. $n=1$과 $d=1.0$일 때 $(d_1,d_2)=(0.10,1.25),(0.10,1.90),\ (0.25,1.25),\ (0.25,1.90)$를 고려한다. $n=4$과 $d=4$일 때 $(d_1,d_2)=(0.40,5.00),(0.40,7.60),\ (1.00,5.00),\ (1.00,7.60)$에 해당한다.

공정모니터링은 시간표기 0에서 출발한다고 가정하나 실제 이 시간에는 어떤 관측도 이루어지지 않는다. 따라서, VSI 관리도를 설정할 때 첫 번째 표본까지의 시간 d_0를 명시하여야 한다. 어떤 적용에서는 출발 이후 표본이 신속히 취해질 수 있게 $d_0=d_1$를 한다. 그러나, 여기서 주어진 수치 결과는 $d_0=d$라는 가정에서 계산한다.

본 논문에서 비교할 모든 관리도는 관리 상태 하에서 ATS=800시간으로 맞추어져 있다. 두 관리도가 조합으로 같이 사용될 때, 관리 상태에서 각 관리도의 ATS는 800보다 커야 조합 관리도의 ATS=800이 된다. 이 경우 두 관리도가 관리 상태에서 같은 ATS (대략 1,550시간)를 가지도록 관리한계를 조절 하면 조합관리도는 관리상태 하에서 ATS=800이 된다.

두 VSI 관리도가 조합으로 사용될 때, 두 관리도 중 어느 하나가 짧은 추출간격을 사용한다면 짧은 추출간격을 사용하여야 한다. 즉, 두 통계량 중 최소한 하나가 각 각의 g값 위에 있다면 다음 추출간격은 d_1이 사용되나, 두 통계량이 각 각의 g값 보다 밑에 있으면 d_2가 사용된다.

여기서 4변수 $(p=4)$의 모든 쌍이 같은 상관계수 $(\rho=0$ 혹은 $\rho=0.9)$를 가질 때 수치결과를 제시한다. ρ의 두 값은 두 개의 극단적인 경우를 고려하였다.

평균벡터 μ에서 이동은 다음과 같이 정의된 표준화된 평균 이동 벡터 $\nu=(\nu_1,\nu_2,\cdots,\nu_p)'$로 나타낸다. 여기서 $\nu_i=(\mu_i-\mu_{0i})/\sigma_{0i},\ i=1,2,\cdots,p$이다. SZ관리도 같은 관리도의 이상상태의 특성은 ν의 이동방향에 의존하는 것이 아니라 비중심모수 $\delta=\sqrt{\nu'\Sigma_{Z0}^{-}1\nu}$에 의존한다. 따라서 이동 크기의 지표로 δ를 사용하는 것이 편리하다. 그러나 편차제곱관리도 혹은 편차제곱관리도를 포함하는 어떤 조합관리도의 특성은 이동방향에 의존한다. 주어진 δ의 값에 Reynolds와 Cho (2006)는 여러 가지의 특정한 이동방향에 대하여 SSATS값을 주고 또한 모든 이동방향에서 평균한 평균 SSATS도 주고 있다. 여기서는 단지 평균 SSATS를 고려한다. 특히, $\sum_{i=1}^{p}\beta_i^2=1$하에서 $\beta=(\beta_1,\beta_2,\cdots,\beta_p)$는 p차원공간에서 단위구면 위에서 랜덤하게 선택된 점을 나타낸다. 그때 비중심모수 δ의 특별한 값에 대응하는 방향 β에서 이동은 $\nu=\delta\beta/\sqrt{\beta'\Sigma_{Z0}^{-}1\beta}$로부터 얻어진다.

Σ에서 이동에 대한 수치결과는 실제로 σ에서 이동이고 변수들간의 상관계수는 변화지 않는다고 가정한다. σ에서 이동은 $\gamma=(\gamma_1,\gamma_2,\cdots,\gamma_p)'$로 나타낸다. 여기서 $\gamma_i=\sigma_i/\sigma_{0i},i=1,2,\cdots,p$이다. 여기서 고려한 관리도의 SSATS는 γ의 이동방향에 의존하므로 평균 SSATS 값들이 주어져 있다. 여기서 평균은 하나 혹은 그 이상의 표준편차에서 증가에 해당하는 모든 방향에 대하여 취한다. 만약 $\psi=1+\sqrt{(\gamma-1)'(\gamma-1)}=1+(\sum_{i=1}^{p}(\gamma_i-1)^2)^{1/2}$으로 σ의 이동크기를 측정하면, $\sigma=\sigma_0$인 관리상태는 $\gamma=1$와 $\psi=1$에 해당한다. ψ의 특정한 값에 해당하는 방향 $1+\beta$에서 1로부터 γ에서 이동은

$\gamma = 1 + (\psi - 1)\beta$로 둠으로써 얻어진다. 평균 SSATS는 단지 분산의 증가에 대한 계산이므로 랜덤방향 발생에서 비음원소를 갖는 β의 값을 사용한다. 본 논문에서 σ의 증가이동을 고려하였으나, 단순히 증가 이동을 σ에서 이동으로 언급한다.

5. 표에 대한 설명

수치결과는 3개의 표에 나타나 있다. 우선, 표의 구조를 설명하고 표로부터 얻어진 결론을 논의한다. 표는 개별관리도와 두 관리도의 조합관리도에 대한 값들을 나타낸다. SSATS를 포함하는 행을 [1], [2], [3]로 나타낸다. μ에서의 랜덤이동은 δ로, 랜덤방향에서 σ의 이동은 ψ로 나타낸다. 표에 있는 값들의 첫 번째 열은 관리상태인 $\delta = 0$와 $\psi = 1.0$에 대응 되므로 이 경우에 대한 ATS값들이 주어져 있다.

표 5.1-표 5.2는 $n = 1$과 $d = 1$, 표 5.3는 $n = 4$와 $d = 4$인 경우이다. 각 표에서 SSATS는 비교를 위하여 FSR 관리도에 대한 값들이 주어져 있다. 특히, 표 5.1-표 5.2에서 SZ관리도는 [1]로 이름붙인 첫 번째 행에 있고 행 [2]와 [3]은 VSI SZ관리도에 대한 것이다. 표 5.3은 $n = 4$인 경우이다.

$n > 1$일 때 SZ관리도는 σ에서 증가를 탐지하는데 비효율적이므로 표 5.3에서 SZ와 SZ^{2}의 조합관리도를 포함하였다.

표 5.1 FSR과 VSI 다변량 슈하르트 관리도의 평균 SSATS ($p = 4,\ n = 1,\ \rho = 0$)

	δ		0	0.4	0.8	1.2	1.6	2.0	5.0	0	0	0	0	0	μ	Σ
	ψ		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.6	2.0	3.0	10.0		
FSR	$d_1 = 1.0$, $d_2 = 1.0$	SZ [1]	800.0	590.5	290.8	123.9	52.5	23.4	0.7	224.3	36.5	12.0	3.0	0.6	h=17.9715, g=——	h=——, g=——
VSI	$d_1 = 0.1$, $d_2 = 1.25$	SZ [2]	800.0	578.3	266.1	101.4	36.6	13.3	0.6	204.2	27.4	7.8	1.7	0.6	h=17.9715, g=5.7530	h=——, g=——
VSI	$d_1 = 0.1$, $d_2 = 1.90$	SZ [3]	800.0	569.3	250.1	88.8	29.6	10.2	0.9	192.2	23.9	6.8	1.8	1.0	h=17.9715, g=3.3527	h=——, g=——

표 5.2 FSR과 VSI 다변량 슈하르트 관리도의 평균 SSATS ($p = 4,\ n = 1,\ \rho = 0.9$)

	δ		0	0.4	0.8	1.2	1.6	2.0	5.0	0	0	0	0	0	μ	Σ
	ψ		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.6	2.0	3.0	10.0		
FSR	$d_1 = 1.0$, $d_2 = 1.0$	SZ [1]	800.0	590.5	290.8	123.9	52.5	23.4	0.7	165.5	18.1	5.9	1.7	0.6	h=17.9715, g=——	h=——, g=——
VSI	$d_1 = 0.1$, $d_2 = 1.25$	SZ [2]	800.0	578.3	266.1	101.4	36.6	13.3	0.6	149.3	13.2	3.8	1.2	0.6	h=17.9715, g=5.7530	h= ——, g=——
VSI	$d_1 = 0.1$, $d_2 = 1.90$	SZ [3]	800.0	569.3	250.1	88.8	29.6	10.2	0.9	140.4	11.7	3.5	1.3	0.9	h=17.9715, g=3.3527	h=——, g=——

6. 표에 대한 결론

6.1. FSR 관리도

표 5.1-표 5.3에서 μ의 이동인 경우, 행 [1]에 있는 SZ와 SZ^{2}의 조합관리도는 SZ관리도에 비하여 큰

표 5.3 FSR과 VSI 다변량 슈하르트 관리도의 평균 SSATS ($p = 4,\ n = 4,\ \rho = 0$)

			δ	0	0.4	0.8	1.2	1.6	2.0	5.0	0	0	0	0	0	μ	Σ
			ψ	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.6	2.0	3.0	10.0		
FSR	$d_1 = 1.0$, $d_2 = 1.0$	SZ & SZ^2	[1]	799.9	411.7	105.9	27.9	9.4	4.2	2.0	195.7	26.5	8.7	2.8	2.0	h=16.2578 g=——	h=60.7308 g=——
VSI	$d_1 = 0.1$, $d_2 = 1.25$	SZ & SZ^2	[2]	800.1	386.3	80.8	15.8	4.8	2.9	2.5	166.6	16.2	5.1	2.7	2.5	h=16.2578 g=6.9951	h=60.7308 g=22.7388
VSI	$d_1 = 0.1$, $d_2 = 1.90$	SZ & SZ^2	[3]	800.0	367.3	68.2	12.9	5.1	4.0	3.7	147.9	13.7	5.4	3.8	3.7	h=16.2578 g=4.4939	h=60.7308 g=15.3007

이동에 대하여는 효과적이나 작은 이동에 대하여는 비효율적이다. σ증가를 탐지하는데 SZ와 SZ^2의 조합관리도는 SZ관리도에 비하여 작은 증가에 대하여는 효율적이나 큰 증가에 대하여는 비효율적이다.

6.2. 가변추출간격을 갖는 슈하르트 관리도

표로부터 VSI SZ관리도가 μ의 중간 정도의 이동을 탐지하는 능력을 크게 향상시킴을 알 수 있다. 예로서 행 [1]에 있는 FSR SZ관리도는 크기 $\delta = 2.0$의 μ의 이동을 탐지하는데 평균 23.4시간이 소요되나 행 [3]의 $(d_1, d_2) = (0.1, 0.9)$를 갖는 VSI 관리도는 평균 10.2시간이 소요된다. σ의 증가를 탐지하는데 VSI 관리도가 더 효율적이다. 예로서 FSR SZ관리도는 크기 $\psi = 2.0$의 σ의 증가를 탐지하는데 평균 12.0시간이 소요되나 행 [3]의 $(d_1, d_2) = (0.1, 1.9)$를 갖는 VSI 관리도는 평균 6.8시간이 소요된다. 그러나 VSI 관리도는 큰 이동에 대하여는 비효율적이다.

$n = 4$일 때 VSI SZ와 SZ^2의 조합관리도는 작은 이동에 대하여 매우 효율적이다. 예로 표 5.3에서 크기 $\delta = 1.2$의 μ의 이동을 탐지하는데 행 [1]에 있는 FSR SZ와 SZ^2의 조합관리도는 평균 27.9시간이 소요되나 행 [3]의 VSI SZ와 SZ^2의 조합관리도는 평균 12.9시간이 소요된다.

공정에서 변화가 있을 때 VSI 관리도의 사용은 매우 효율적이며 만약 d_1이 작다면 추가적인 정보를 신속히 얻을 수 있다.

6.3. 추출간격 선택

표 5.1-표 5.3에서 VSI 관리도에 대한 SSATS는 두 개의 긴 추출간격 $d_2 = 1.25d$와 $d_2 = 1.9d$에 대한 값이다. 기대하는 바와 같이 $d_2 = 1.25d$는 큰 이동에 $d_2 = 1.9d$는 작은이동에 효율적이다. 전반적으로 좋은 효율을 얻기 위하여 d_2값은 $d_2 = 1.25d$와 $d_2 = 1.9d$사이의 값이 합리적이 값이다. VSI 관리도에서 짧은 간격 d_1은 가능한 작을수록 최상이다. 표에서 $d_1 = 0.1d$이다.

6.4. 표본크기 n의 선택

작은 표본을 자주 뽑는 것이 좋은지 혹은 큰 표본을 덜 자주 뽑는 것이 좋은지를 생각해 보자. 여기서 비 n/d는 상수로 가정 하자. 예로써 만약 FSR 관리도가 사용되면 매 $d = 1$시간 마다 $n = 1$ 표본,

매 $d = 4$시간 마다 $n = 4$ 표본, 매 $d = 8$시간 마다 $n = 8$ 표본을 취하는 것이 최상이냐? Hawkins와 Olwell (1998), Reynolds와 Stoumbos (2004a, 2004b, 2005)는 이 이슈를 일변량 셋팅에서 연구하였고, Reynolds와 Cho (2006)는 다변량 셋팅에서 연구하였다. 얻어진 결론은 최상의 표본추출 패턴은 사용할 관리도의 형태에 의존한다.

슈하르트 관리도는 n과 d의 선택에 매우 민감하다. 특히 n의 작은 값은 큰 이동을 탐지하는데 좋고, n의 큰 값은 작은 이동을 탐지하는데 좋다. 작은 이동과 큰 이동 둘 다를 탐지하기 위하여 $n = 4$와 같은 n의 중간 값 정도가 합리적인 타협 값이 될 것이다.

VSI 슈하르트 관리도에서 n 선택의 효과를 조사하기 위하여 표 5.1의 행 [2]와 [3]에 있는 $n = 1$과 $d = 1$을 갖는 VSI SZ관리도와 표 5.3의 행 [2]와 [3]에 있는 $n = 4$과 $d = 4$을 갖는 SZ와 SZ^2의 조합관리도를 비교한다. $n = 1$을 사용하는 것이 μ혹은 σ에서 큰 이동을 탐지하는 것에 좋고, $n = 4$을 사용하는 것이 작은 이동을 탐지하는 것에 좋다. 수치 결과가 표에는 없지만 $n = 8$와 같은 n의 큰 값을 사용하는 것이 작은 이동을 탐지하는 능력을 향상 시키지만 큰 이동을 탐지하는 능력을 해친다. 따라서 FSR 슈하르트 관리도와 같이 작은 이동과 큰 이동 둘 다를 탐지하기 위하여 $n = 4$와 같은 n의 중간 값 정도가 합리적인 타협 값이 될 것이다.

7. 결론

본 연구는 평균벡터 μ와 표준편차벡터 σ를 동시에 모니터링하기 위한 다변량 슈하르트 관리도에서 VSI 사용의 성능을 조사하는 것이다. 모니터링의 목적은 평균벡터에서 작은 이동과 큰 이동, 표준편차벡터 σ에서 작은 혹은 큰 증가를 탐지하는데 있다고 가정하자. 여기서 주요 관심사가 특별한 이동 방향에 있는 것은 아니다.

VSI 사용은 μ 혹은 σ에서 대부분의 이동에 대하여 슈하르트 관리도의 성능을 크게 향상 시킨다. VSI 사용은 짧은 추출간격 d_1이 매우 작을 때 가장 큰 성능 향상을, d_1이 아주 작지 않을 때도 상당한 성능 향상을 가져 온다.

슈하르트 관리도에 대하여 $n > 1$을 선택하는 것이 성능을 가장 크게 향상 시킨다. 따라서, VSI 관리도에서 최상의 n 선택은 표준 FSR 관리도의 경우에서 결론과 거의 유사하다.

최상의 성능을 갖는 구조에 대한 본 연구의 논의는 모든 구조들이 관리 상태에서 똑 같은 평균 표본추출비를 갖는 다는 가정 하에서 가장 작은 SSATS를 갖는 구조로 짜여져 있다. 어떤 경우에서는, VSR 관리도를 사용하는 동기가 공정 모수의 변화를 탐지하는 적당한 능력을 유지하면서 표본 추출 비용을 절감하는 것이다 (Baxley, 1995; Reynolds, 1996). 만약 현재 사용 중인 구조가 매 $d = 4$시간당 $n = 4$표본을 추출하는데 근거한 FSR 구조라면, 샘플 추출비율은 매 시간당 1.0표본이다. 1시간당 0.5표본을 취하는 표본 추출비율을 갖는 VSR 구조로 바꾸면 표본 추출 비용을 50% 절감하면서 대략적으로 모수 이동을 탐지하는 능력은 근사적으로 같다.

참고문헌

Alt, F. A. (1984). *Multivariate quality control*, S. Kotz, N. L. Johnson, and C. R. Reid (Eds), The Encyclopedia of statistical Sciences, 110-122, Wiley, Newyork.

Aparisi, F. (1996). Hotelling's T^2control chart with adaptive sample sizes. *International Journal of Production Research*, **34**, 2853-2862.

Aparisi, F. and Haro, C. L. (2001). Hotelling's T^2control chart with variable sampling intervals. *International Journal of Production Research*, **39**, 3127-3140.

Aparisi, F. and Haro, C. L. (2003). A comparison of T^2control charts with variable sampling schemes as opposed to MEWMA chart. *International Journal of Production Research*, **41**, 2169-2182.

Baxley, R. V. Jr. (1995). An application of variable sampling interval control charts. *Journal of quality Technology*, **27**, 275-282.

Domangue, R. and Patch, S. C. (1991). Some omnibus exponentially weighted moving average statistical process monitoring schemes. *Technometrics*, **33**, 299-313.

Hawkins, D. M. and Olwell, D. H. (1998). *Cumulative sum control charts and charting for quality improvement*, Springer-Verlag, New York.

Hotelling, H. (1947). *Multivariate quality control-illustrated by the air testing of sample bombsights, in techniques of statistical analysis*, C. Eisenhart, M. W. Hastay, and W. A. Wallis, eds., McGraw-Hill, New York.

Lowry, C. A. and Montgomery, D. C. (1995). A review of multivariate control charts. *IIE Transactions*, **27**, 800-810.

MacGregor, J. F. and Harris, T. J. (1993). The exponentially weighted moving variance. *Journal of Quality Technology*, **25**, 106-118.

Mason, R. L. and Young J. C. (2002). *Multivariate statistical process control with industrial applications*, ASA-SIAM, Philadelphia.

Montgomery, D. C. (2005). *Introduction to statistical quality control*, 5th Edition, John Wiley & Sons, New York.

Reynolds, M. R., Jr. (1995). Evaluating properties of variable sampling interval control charts. *Sequential Analysis*, **14**, 59-97.

Reynolds, M. R., Jr. (1996). Shewhart and EWMA variable sampling interval control charts with sampling at fixed times. *Journal of Quality Technology*, **28**, 199-212.

Reynolds, M. R., Jr., Amin, R. W., Arnold, J. C. and Nachlas, J. A. (1988). $\bar{X}$charts with variable sampling intervals. *Technometrics*, **30**, 181-192.

Reynolds, M. R., Jr. and Cho, G. Y. (2006). Multivariate control charts for monitoring the mean vector and covariance matrix. *Journal of Quality Technology*, **38**, 230-253.

Reynolds, M. R., Jr. and Ghosh, B. K. (1981). Designing control charts for means and variances. *35th Annual Quality Congress Transactions*, 400-407.

Reynolds, M. R., Jr. and Stoumbos, Z. G. (2001). Monitoring the process mean and variance using individual observations and variable sampling intervals. *Journal of Quality Technology*, **33**, 181-205.

Reynolds, M. R., Jr. and Stoumbos, Z. G. (2004a). Control charts and the efficient allocation of sampling resources. *Technometrics*, **46**, 200-214.

Reynolds, M. R., Jr. and Stoumbos, Z. G. (2004b). Should observations be grouped for effective process monitoring?. *Journal of Quality Technology*, **36**, 343-366.

Reynolds, M. R., Jr. and Stoumbos, Z. G. (2005). Should exponentially weighted moving average and cumulative sum charts be used with Shewhart limits?. *Technometrics*, **47**, 409-424.

Shamma, S. E. and Amin, R. W. (1993). An EWMA quality control procedure for jointly monitoring the mean and variance. *International Journal of Quality and Reliability Management*, **10**, 58-67.

Stoumbos, Z. G., Mittenthal, J. and Runger, G. C. (2001). Steady-state-optimal adaptive control charts based on variable sampling Intervals. *Stochastic Analysis and Applications*, **19**, 1025-1057.

Stoumbos, Z. G. and Reynolds, M. R., Jr. (2005). Economic statistical design of adaptive control schemes for monitoring the mean and variance: an application to analyzers. *Nonlinear Analysis - Real-World Applications*, **6**, 817-844.

Tagaras, G. (1998). A survey of recent developments in the design of adaptive control charts. *Journal of Quality Technology*, **30**, 212-231.

Wierda, S. J. (1994). Multivariate statistical process control - recent results and directions for future research. *Statistica Neerlandica*, **48**, 147-168.

Zhang, S. and Wu, Z. (2006). Monitoring the process mean and variance using a weighted loss function CUSUM scheme with variable sampling intervals. *IIE Transactions*, **38**, 377-387.

Journal of the Korean Data & Information Science Society
2010, **21**(6), 999–1008

한국데이터정보과학회지

Multivariate Shewhart control charts with variable sampling intervals†

Gyo Young Cho[1]

[1]Department of Statistics, Kyungpook National University

Received 10 May 2010, revised 30 September 2010, accepted 5 October 2010

Abstract

The objective of this paper is to develop variable sampling interval multivariate control charts that can offer significant performance improvements compared to standard fixed sampling rate multivariate control charts. Most research on multivariate control charts has concentrated on the problem of monitoring the process mean, but here we consider the problem of simultaneously monitoring both the mean and variability of the process.

Keywords: Average time to signal, multivariate control chart, statistical process control, steady state ATS, variable sampling interval.

† This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0076310).

1 Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
E-mail: gycho@knu.ac.kr