

## 고정표본채취시점을 갖는 가변표본채취간격 다변량 $T^2$ 관리도

장영순·배도선

한국과학기술원 산업공학과

### Multivariate $T^2$ Variable Interval Control Chart with Sampling at Fixed Times

Young Soon Chang · Do Sun Bai

Dept. of Industrial Engineering, Korea Advanced Institute of Science and Technology

#### Abstract

This paper proposes a multivariate  $T^2$  variable interval control chart with sampling at fixed times, where samples are taken at specified equally spaced fixed time points, and additional samples are allowed between these fixed times when indicated by the preceding  $T^2$  statistics. At fixed sampling times, the  $T^2$  statistics are composed of all quality characteristics, and a part of quality characteristics are selected to obtain  $T^2$  statistics at additional sampling times. A Markov chain approach is used to evaluate the performance of the proposed chart.

#### 1. 서론

최근 품질관리분야에서는 공정을 검사하여 불량품이 생산되지 않도록 예방활동을 하는 통계적 공정관리(statistical process control: SPC)의 체계적인 활용이 강조되고 있다. SPC의 중요한 목적중의 하나는 공정에 이상이 발생했을 때 이를 빨리 탐지하여 적절한 수정조치를 취하는 것이다. 이러한 목적으로 사용되는 가장 대표적인 기법이 관리도이다. 관리도란 관리한계선이 있는 그래프에 시간의 흐름에 따른 품질의 변동상황을 나타낸 것으로, 표본으로부터 계산한 타점통계량의 값을 순차적으로 타점하여 공정의 이상유무를 판단한다. 만일 모든 점들이 관리한계선 안에 놓이고 점들의 행태로부터 별

다른 이상 징후가 발견되지 않으면 공정이 관리상태에 있다고 보고, 만약 한 점이라도 관리한계선을 벗어나거나 점들이 비정상적인 행태를 보이면 공정이 이상상태에 있다고 판단하여 공정을 면밀히 조사하게 된다.

관리도는 관리하고자 하는 품질특성치의 개수에 따라 단변량관리도와 다변량관리도로 나누어 볼 수 있다. 단변량관리도는 품질특성치의 개수가 하나인 경우에 사용되는 것으로 공정평균의 변화를 탐지하기 위해서는  $\bar{X}$  관리도가 널리 사용된다. 다변량관리도는 상관관계가 존재하는 다수의 품질특성치들을 관리하기 위해 사용하는 것으로, Hotelling의  $T^2$  통계량을 이용한  $T^2$  관리도가 널리 사용된다.

단변량 혹은 다변량 관리도는 일반적으로 동일한 간격의 고정된 표본채취시점들에서 표본을 얻어서 운영되는데, 이러한 관리도를 고정표본채취간격(fixed sampling interval: FSI) 관리도라 한다. 최근 들어 FSI 관리도의 효율을 높이기 위하여 타점통계량의 값에 따라 표본채취간격을 변화시키는 가변표본채취간격(variable sampling interval: VSI) 관리도에 관한 연구가 이루어지고 있다. VSI 관리도는 표본으로부터 계산된 타점통계량의 값이 관리한계선 근처에 타점되면 공정이 이상상태일 가능성이 높다고 보고 다음 표본채취시점까지의 간격을 짧고 하고, 그렇지 않으면 표본채취간격을 길게 한다. Reynolds 등 (1988), Amin 과 Hemasinha (1993), Prybutok 등 (1997), Lee 와 Bai (2000) 는 VSI

$\bar{X}$  관리도를, Chengalur 등 (1993) 과 Aparisi 와 Haro (2001)는 VSI  $T^2$  관리도를 제안하였으며, 이들은 모두 VSI 관리도를 사용하는 것이 FSI 관리도를 사용하는 것에 비해 공정의 이상을 더 빨리 탐지할 수 있음을 보였다.

FSI 관리도를 사용할 경우에는 제품의 가공시간이나 작업자의 교대시간 등으로부터 정해지는 시점에 맞추어 표본채취간격을 고정하여 사용할 수 있으나 VSI 관리도를 사용할 경우에는 타점통계량의 값에 따라 표본채취시점이 랜덤하게 결정되므로 공정상황에 따라 표본채취가 필요한 시점에 표본을 얻을 수 없는 경우가 자주 발생한다. Reynolds (1996a) 와 Reynolds (1996b) 는 이러한 점을 개선하여 고정표본채취시점을 갖는 단변량 가변표본채취간격 (variable sampling interval with fixed times: VISFT) 관리도를 제안하였다. VSIFT 관리도는 공정상황을 고려하여 미리 정해놓은 고정된 표본채취시점에서 표본을 채취하다가 타점통계량의 값이 관리한계선 근처에 타점되면 다음 고정표본채취시점 이전에 표본을 추가로 얻는다. 이러한 방식은 FSI 관리도와 같이 공정상황에 적합하도록 표본채취시점을 미리 정해놓으므로 실제 공정에서 사용하기가 용이하다는 장점이 있다.

이 논문에서는 기존의 단변량 VSIFT 관리도를 확장하여 다변량 VSIFT  $T^2$  관리도를 제안한다. 다변량 관리도를 운영하기 위해서는 한 표본채취시점에서 검사해야 하는 품질특성치가 많으므로 이를 줄이기 위한 연구가 필요하다. 따라서 이 논문에서는 고정된 표본채취시점에서는 모든 품질특성치들을 검사하다가 타점통계량이 관리한계선 근처에 타점되면, 추가표본채취시점에서는 이상상태일 가능성이 높은 일부 품질특성치들만을 검사하여 공정을 관리하는 VSIFT  $T^2$  관리도를 제안한다. 이러한 방법은 공정이 이상상태에 있는 경우 이상 발생의 가능성이 높은 품질특성치들을 선별하여 관리할 수 있으므로 다변량 관리도의 성능을 향상시킬 수 있을 뿐만 아니라 관측치의 개수를 줄일 수 있다는 장점이 있다.

## 2. VISFT $T^2$ 관리도

### 2.1. $T^2$ 관리도

$\bar{X}_i$  를 평균벡터가  $\mu = (\mu_1, \dots, \mu_p)^T$  이고 분산-공분산행렬이

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\ & \sigma_2^2 & \cdots & \rho_{2p}\sigma_2\sigma_p \\ & & \ddots & \vdots \\ & & & \sigma_p^2 \end{bmatrix}$$

인  $p$  변량 정규분포  $N_p(\mu, \Sigma)$  를 따르는 공정으로부터 크기  $n$  인 표본을 관찰하여 얻은  $i$  번째 표본평균이라 할 때,  $T^2$  관리도의 타점통계량은

$$T_i^2 = n(\bar{X}_i - \mu)^T \Sigma^{-1} (\bar{X}_i - \mu) \quad (2.1)$$

와 같이 정의된다. 여기서 'T'는 벡터 또는 행렬의 전치(transpose)를 나타낸다. 식 (2.1)의  $T^2$  통계량은 자유도가  $p$  인  $\chi^2$  분포를 따른다는 것이 알려져 있으므로 관리도의 거짓경보비율을  $\alpha$  라 할 때, 관리상한선은 자유도  $p$  를 갖는  $\chi^2$  분포의  $100(1-\alpha)$  백분위수인  $\chi_\alpha^2(p)$  가 된다.

만일 각 표본채취시점에서 얻어지는 품질특성치의 개수가 다른 경우에는 이를 하나의 관리도에 나타내기 위해 다음과 같이  $T^2$  를 표준화한  $Q_i$  를 타점통계량으로 사용할 수 있다.

$$Q_i = \Phi^{-1}(G_{p_i}(T_i^2)), \quad (2.2)$$

여기서  $p_i$  는  $i$  번째 표본채취시점에서 얻어지는 품질특성치의 개수이고,  $T_i^2$  는 해당되는 품질특성치의 평균벡터와 분산-공분산행렬로만 구해진  $T^2$  값이다. 또한,  $G_{p_i}(\cdot)$  는 자유도가  $p_i$  인  $\chi^2$  분포의 분포함수이고,  $\Phi^{-1}(\cdot)$  는 표준정규분포함수의 역함수이다. 따라서 식 (2.2)의 통계량  $Q_i$  에 대한 관리한계선은  $z_\alpha$  가 된다.

### 2.2. VSIFT $T^2$ 관리도

공정운영상황을 고려하여 정해진 고정표본채취간격을  $d_f$  라 하고 이를  $\eta$  개의 구간으로 나누는  $d_v = d_f/\eta$  를 추가표본채취간격이라 하자.

VSIFT  $T^2$  관리도는 매 고정표본채취시점에서는  $p$  개의 품질특성 모두를 검사하여 타점통계량  $Q$ 를 구하고 그 값이 경고한계선(warning limit)  $g$ 보다 작으면 공정이 관리상태라 판단하여 다음 고정표본채취시점으로 넘어가고, 관리한계선  $h$ 보다 크면 공정에 이상이 발생했다고 판단하여 이상신호를 발생한다. 또한,  $g < Q_i \leq h$  이면 공정이 이상상태일 가능성이 높다고 판단하고 추가표본채취시점으로 넘어가게 된다. 이때 추가표본채취시점에서는 모든 품질특성치들을 검사하는 것이 아니라 이상상태일 가능성이 높은  $p_v (\leq p)$  개의 품질특성치만을 선별하여 검사하게 된다.  $p_v$  개의 품질특성을 선별하는 방법은 여러가지가 있을 수 있으나 이 논문에서는 관리도 운영의 편의를 고려하여 각 품질특성치의 표본평균을 표준화한 값

$$|\bar{z}_j| = \left| \frac{\bar{x}_j - \mu_j}{\sigma_j / \sqrt{n}} \right| \quad (2.3)$$

가 큰 순서로  $p_v$  개를 선택하는 방법을 사용한다. 추가표본채취시점에서는 선별된 품질특성치들과 관리한계선  $h$ , 경고한계선  $g$ 를 이용하여 공정의 이상유무를 판단하게 된다. VSIFT  $T^2$  관리도의 운영절차는 다음과 같다.

- (i)  $i$  번째 표본채취시점이 고정표본채취시점인 경우  $p$  개의 모든 품질특성치들로부터  $Q$ 를 구한 후,
  - a)  $Q_i > h$ 이면 이상신호를 발생한다.
  - b)  $Q_i \leq g$  이면 공정이 관리상태라고 판단하여 다음 고정표본채취시점으로 넘어간다.
  - c)  $g < Q_i \leq h$  이면 이상상태일 가능성이 높은  $p_v$  개의 품질특성치들을 선별한 후 추가표본채취시점으로 넘어간다.
- (ii)  $i$  번째 표본채취시점이 추가표본채취시점인 경우 미리 선별된  $p_v$  개의 품질특성치들로부터  $Q$ 를 구한 후,
  - a)  $Q_i > h$  이면 이상신호를 발생한다.
  - b)  $Q_i \leq g$  이면 공정이 관리상태라고 판단하여 고정표본채취시점으로 넘어간다.
  - c)  $g < Q_i \leq h$  이면 마지막 추가표본채취시점인 경우에는 고정표본채취시점으로 넘어가

고 그렇지 않은 경우에는 다음 추가표본채취시점으로 넘어간다.

이때, 관리한계선  $h$ 는 원하는 수준의 거짓경보비율이  $\alpha$  라면  $h = z_\alpha$ 로 설정하고, 경고한계선  $g$ 는 공정이 관리상태일 때 타점통계량이 관리한계선과 경고한계선 사이에 타점될 확률이  $q$ 가 되도록  $g = z_{\alpha+q}$ 로 설정한다.

<그림 2.1>은  $\eta = 3$  인 VSIFT  $T^2$  관리도의 예이다. 그림에서 시간축위의 굵은 세로선은 고정표본채취시점을, 얇은 세로선은 추가표본채취가 가능한 시점을 나타낸다. 관리도상에 검은색 점은 고정표본채취시점에서 타점된 통계량을, 흰색 점은 추가표본채취시점에 타점된 통계량을 나타낸다. 첫번째 고정표본채취시점  $t_1$ 에서는 타점통계량이 경고한계선  $g$  아래에 위치하므로 공정이 관리상태에 있다고 판단하고 다음 고정표본채취시점으로 넘어간다.  $t_2$ 에서 얻은 타점통계량은 관리한계선  $h$ 와 경고한계선  $g$  사이에 타점되었으므로 이상상태일 가능성이 높은 품질특성치들을 선별한 후 추가표본채취시점으로 넘어간다. 추가표본채취시점  $t_{2,1}$ 에서는 공정이 안정상태인 것으로 판단되어 다음 고정표본채취시점  $t_3$ 로 넘어간다.  $t_{3,2}$ 에서는 타점통계량의 값이 관리한계선  $h$ 를 벗어났으므로 이상신호를 발생한다.

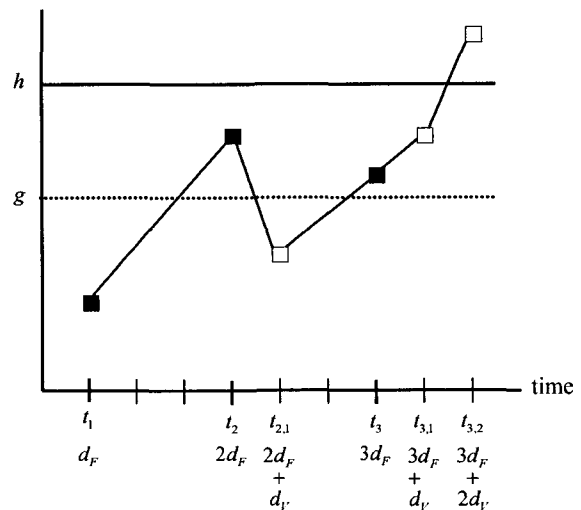


그림 2.1  $\eta = 3$  인 VSIFT  $T^2$  관리도

### 3. 수행도 평가 측도

관리도의 성능을 평가하는 여러 가지 수행도 평가 측도들 중에서 널리 사용되는 것으로 관리도에 의하여 이상신호가 발생될 때까지의 평균 시간인 *ATS* (average time to signal), 이상신호가 발생될 때까지 얻게 되는 총 관측치 수의 평균인 *ANOS* (average number of observations to signal) 등이 있다. 실제로 관리도를 운영하는 경우에는 공정이 안정상태에서 시작하여 임의의 시점에서 이상상태로 변하게 된다. 이러한 점을 고려하여 사용할 수 있는 측도로는 *AATS* (adjusted *ATS*)와 *AANOS* (adjusted *ANOS*)가 있다.

관리도의 수행도를 평가하기 위한 식들은 마야코프 연쇄(Markov chain) 혹은 적분방정식(integral equation)을 이용하여 구할 수 있다. 이 논문에서는 비교적 적용범위가 넓은 마야코프 연쇄를 이용하여 이들을 유도한다.

마야코프 연쇄를 이용하여 제안된 VSIFT  $T^2$  관리도의 성능평가측도들을 유도하기 위해 타점통계량  $Q_l$ 의 값과 표본채취시점 및 검사하는 품질특성치에 따라 다음과 같이 마야코프 연쇄의 상태(state)들을 정의한다.

- (i)  $S_A$ :  $Q_l > h$ 인 경우
  - (ii)  $S_{0,0}$ : 고정표본채취시점이고  $Q_l \leq g$ 인 경우
  - (iii)  $S_{0,m}$ : 고정표본채취시점이고  $g < Q_l \leq h$ 가 되어 다음 표본채취시점이 추가표본채취시점이며, 추가표본채취시점에서 검사할 품질특성치들의 집합이  $E_m$ 인 경우,  $m=1, \dots, \phi$
  - (iv)  $S_{l,0}$ :  $l$ 번째 추가표본채취시점이고  $Q_l \leq g$  또는  $l=\eta-1$ 이 되어 다음 표본채취시점이 고정표본채취시점인 경우,  $l=1, \dots, \eta-1$
  - (v)  $S_{l,m}$ :  $l$ 번째 추가표본채취시점이고 검사할 품질특성치들의 집합이  $E_m$ 이며  $g < Q_l \leq h$ 가 되어 다음 표본채취시점 역시 추가표본채취시점인 경우,  $l=1, \dots, \eta-1$ ,  $m=1, \dots, \phi$
- 추가표본채취시점에서 검사하게 되는 품질특성치들의 조합은 총  $\phi = \binom{p}{p_v}$  개가 된다. 이 때, <그림 3.1>과 같이 품질특성치들을 1 열로 나

$X_1$	$X_2$	$X_3$	$X_4$	집합
0	0	1	1	$\rightarrow E_1$
0	1	0	1	$\rightarrow E_2$
0	1	1	0	$\rightarrow E_3$
1	0	0	1	$\rightarrow E_4$
1	0	1	0	$\rightarrow E_5$
1	1	0	0	$\rightarrow E_6$

그림 3.1  $p=4, p_v=2$ 인 경우의 선별된 품질특성치들의 집합

열하고 각각의 품질특성치가 선택되었으면 1, 그렇지 않으면 0을 배정하여 2진수를 만든다. 이렇게 만들어진 2진수의 값이 작은 것부터 큰 순서대로 나열하여 선택된 품질특성치들로만 이루어진 집합을  $E_1$ 에서  $E_\phi$ 라 정의한다. <그림 3.1>은  $p=4$ 이고  $p_v=2$ 인 경우에 추가표본채취시점에서 검사하게 되는 품질특성치들의 조합에 대한 집합  $E_1, \dots, E_\phi$ 를 얻는 방법에

대해 설명한 것으로  $\phi = \binom{4}{2} = 6$ 이 된다.

이렇게 나누어진 상태들 간의 전이확률행렬  $R$ 은

$$R = \begin{bmatrix} Q & (I-Q)\mathbf{1} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3.1)$$

으로 표현할 수 있다. 여기서  $Q$ 는 관리도의 운영을 계속하게 되는 일시상태(transient stat)  $S_{l,m}$ ,  $l=0, \dots, \eta-1$ ,  $m=0, \dots, \phi$ 들 간의 전이확률을 나타내는  $\eta(\phi+1) \times \eta(\phi+1)$ 차 정방행렬이며,  $(I-Q)\mathbf{1}$ 는 일시상태에서 관리도에 의해 이상신호가 발생되는 흡수상태(absorbing state)  $S_A$ 로의 전이확률을 나타내는  $\eta(\phi+1)$ 차 열벡터이고,  $\mathbf{0}$ 는 모든 원소들이 0으로 이루어진  $\eta(\phi+1)$ 차 열벡터이다.

식 (3.1)의 전이확률행렬을 이용하면 다음과 같은 VSIFT  $T^2$  관리도의 *ATS*와 *ANOS*를 얻을 수 있다.

$$ATS = \mathbf{b}^T (I-Q)^{-1} \mathbf{d}, \quad (3.2)$$

$$ANOS = \mathbf{b}^T (I-Q)^{-1} \mathbf{v}_s, \quad (3.3)$$

여기서  $\mathbf{b}$ 는 관리도 시작시의 상태를 나타내는

초기확률벡터로 일반적으로 관리도는 고정표본 채취시점에서 시작하므로  $S_{0,0}$ 에 해당하는 원소는 1 이고 나머지는 모두 0 을 사용한다.  $\mathbf{d}$  와  $\mathbf{v}_s$  는 각각 다음 표본채취시점까지의 시간 간격과 한 표본채취시점에서의 관측치의 수를 나타내는 열벡터로  $S_{l,m}$ 에 해당하는 원소는 각각

$$d(l,m) = \begin{cases} (\eta-l)d_v, & m=0 \\ d_v, & \text{기타} \end{cases}, \quad (3.4)$$

$$v_s(l,m) = \begin{cases} np, & m=0 \\ np_v, & \text{기타} \end{cases} \quad (3.5)$$

이 된다.

공정이 관리상태에서 오랜 시간 운영된 후 임의의 시점에서 타점 통계량이 마야코프 연쇄의 각 상태에 있을 안정상태(steady state) 확률을 나타내는 벡터를  $\boldsymbol{\pi}$ 라 하면,  $S_{l,m}$ 에 해당하는 원소는

$$\pi(l,m) = \frac{\left[ \mathbf{b}^T (\mathbf{I}-\mathbf{Q})^{-1} \right]_{l,m} \times d(l,m)}{ATS} \quad (3.6)$$

이 된다. 특정표본채취간격 내에서 평균변화가 일어나는 시점은 균일분포를 따른다고 가정하면 (Reynolds 등, 1990) AATS 와 AANOS 는 각각

$$AATS = \boldsymbol{\pi}^T (\mathbf{I}-\mathbf{Q})^{-1} \mathbf{d} - \frac{1}{2} \boldsymbol{\pi}^T \mathbf{d}, \quad (3.7)$$

$$AANOS = \boldsymbol{\pi}^T (\mathbf{I}-\mathbf{Q})^{-1} \mathbf{v}_s, \quad (3.8)$$

이 된다.

## 참고문헌

- Amin, R. W. and Hemasinha, R. (1993), The Switching Behavior of  $\bar{X}$  Charts with Variable Sampling Intervals, *Communications in Statistics - Theory and Methods* **22**, 2081-2102.
- Aparisi, F. and Haro, C. L. (2001), Hotelling's  $T^2$  Control Chart with Variable Sampling Intervals, *International Journal of Production Research* **39**, 3127-3140.
- Chengalur, I. N., Arnold, J. C., and Reynolds, M. R. (1993), Multiple Means Shewhart Charts Using Variable Sampling Intervals, *Journal of*

- Statistical Computation and Simulation* **45**, 1-21.
- Lee, K. T. and Bai, D. S. (2000), Variable Sampling Interval  $\bar{X}$  Control Charts with Run Rules, *International Journal of Industrial Engineering* **7**, 147-158.
- Prybutok, V. R., Clayton, H. R., and Harvey, M. M. (1997), Comparison of Fixed versus Variable Sampling Interval Shewhart  $\bar{X}$  Control Charts in the Presence of Positively Autocorrelated Data, *Communications in Statistics - Simulation and Computation* **26**, 83-106.
- Reynolds, M. R. (1996a), Variable-Sampling-Interval Control Charts with Sampling at Fixed Times, *IIE Transactions* **32**, 371-384.
- Reynolds, M. R. (1996b), Shewhart and EWMA Variable Sampling Interval Control Charts with Sampling at Fixed Times, *Journal of Quality Technology* **28**, 199-212.
- Reynolds, M. R., Amin, R. W., and Arnold, J. C. (1990), CUSUM Charts with Variable Sampling Intervals, *Technometrics* **32**, 371-384.
- Reynolds, M. R., Amin, R. W., Arnold, J. C., and Nachlas, J. A. (1988),  $\bar{X}$  Charts with Variable Sampling Intervals, *Technometrics* **30**, 181-192.