변량추출비 $\overline{X}$ 관리도의 통계적 효율 비교

이재헌[1], 박창순[2], 전성호[3]

요약

변량추출비 관리도는 현재의 관측값에 기초하여 다음 시점의 표본크기와 표본추출간격을 변화시키면서 공정의 변화를 탐지하는 관리도 절차이다. 만일 공정에서 추출한 현재의 관측값을 살펴볼 때 공정변화의 징후가 있는 경우에는 다음 시점의 표본추출비를 증가시켜, 즉 표본크기를 크게 하고 표본추출간격을 작게 하여 예상되는 공정변화를 더 빠르고 정확하게 탐지함으로 보다 효율적인 공정관리를 수행하는 것이다. 본 연구는 변량추출비 $\overline{X}$ 관리도에서 사용하는 표본크기와 표본추출간격의 수를 달리하며 각각의 경우에 대한 통계적 효율을 계산하고 이를 비교하고자 한다.

주요용어 : 변량표본크기, 변량표본추출간격, 변량추출비, $\overline{X}$ 관리도, Markov chain

1. 서론

관리도(control chart)는 나쁜 품질의 공정 결과물을 줄 수 있는 공정모수의 변화를 탐지하기 위하여 널리 사용되고 있다. 관리도에서 표본을 추출하는 전통적인 방법은 고정된 표본추출간격(fixed sampling interval: FSI)에서 고정된 표본크기(fixed sample size: FSR)를 추출하는 고정표본추출비(fixed sampling rate: FSR)를 사용하는 것이다.

그러나 최근 들어 공정에서 얻어지는 현재의 관측값을 기초하여 표본크기와 표본추출간격을 변화시키는 것에 대한 연구가 활발하게 진행되고 있다. 기본적인 아이디어는 공정에서 추출된 현재의 관측값을 살펴볼 때 공정변화의 징후가 있는 경우와 그렇지 않은 경우로 분리하여, 변화의 징후가 있는 경우에는 다음 시점에서의 표본추출비를 증가시켜, 즉 표본추출시 더 많은 관측치를 추출하고 표본추출간격을 작게 함으로써 예상되는 공정변화를 빠르고 정확하게 탐지하자는 것이다. 만일 예상되지 않는 경우에는 표본추출시 상대적으로 작은 수의 관측치와 큰 추출간격으로 공정을 관리하는 것이다.

표본추출비중 표본크기만을 변화시키는 관리도는 변량표본크기(variable sample size: VSS) 관리도, 표본추출간격만을 변화시키는 관리도는 변량표본추출간격(variable sampling interval: VSI) 관리도, 그리고 VSS와 VSI 방법을 동시에 수행하는 관리도를 변량추출비(variable sampling rate: VSR) 관리도라 한다. 이에 대한 대표적인 연구로는 Reynolds, Amin, Arnold와 Nachlas(1988), Phabhu, Runger와 Keats(1993), Phabhu, Montgomery와 Runger(1994), Park과 Reynolds(1994, 1999), Zimmer, Montgomery와 Runger(1998), 그리고 Reynolds와 Arnold(2001) 등을 들 수 있다.

전통적으로 사용하는 Shewhart의 FSR $\overline{X}$ 관리도는 현재의 관측값만으로 공정을 판단하기 때문에 사용하기에 간편하고 모수의 큰 변화에는 효율적이지만, 모수의 꾸준하면서 작은 변화

1) (503-703) 광주광역시 남구 진월동 592-1 광주대학교 산업정보공학과 부교수
2) (156-756) 서울특별시 동작구 흑석동 221 중앙대학교 수학통계학부 교수
3) (503-703) 광주광역시 남구 진월동 592-1 광주대학교 산업정보공학과 석사과정

는 빨리 탐지하지 못한다는 단점을 가지고 있다. 이와 같은 단점을 해결할 수 있는 방법 중에 하나가 VSR 방법을 적용하는 것이다.

지금까지의 VSR $\overline{X}$ 관리도에 대한 연구를 살펴보면 2가지 표본크기를 사용하는 VSS 방법과 2가지 표본추출간격을 사용하는 VSI 방법을 결합시키는 것이 대부분이었다. VSI 관리도에서는 2가지 표본추출간격을 사용하는 것이 통계적인 효율 면에서 최적이라는 사실이 알려져 있지만, VSI 관리도에서 표본크기의 개수에 대한 최적 성질에 대해서는 연구된 바가 없다. 단지 Zimmer, Montgomery와 Runger(1998)는 3가지 표본크기를 사용하는 VSS $\overline{X}$ 관리도가 2가지 표본크기를 사용하는 것에 비해 공정평균의 변화가 작은 경우 훨씬 효율적이라는 연구 결과를 발표했다. 본 논문에서는 VSI 방법에서 표본추출간격은 2가지로 고정하고, VSS 방법에서 표본크기는 2가지, 3가지, 그리고 4가지를 사용하는 VSR $\overline{X}$ 관리도의 통계적 효율에 대하여 비교하고 이를 분석하고자 한다.

2. VSR $\overline{X}$ 관리도의 절차

먼저 공정에서 추출하는 품질특성치의 분포는 평균 μ, 분산 σ^2을 갖는 정규분포에 따름을 가정하고, 공정평균 μ가 목표치 μ_0에서 변화하는지를 탐지하는 공정관리를 고려해 보자. N_t는 t번째 표본추출시점에서의 표본크기, H_t는 표본 $t-1$과 t간의 표본추출간격을 나타낸다. 시점 t에서의 공정평균을 $\overline{X}_t$, 이를 표준화시킨 표본평균을 $Z_t = \sqrt{N_t}(\overline{X}_t - \mu_0)/\sigma$라 할 때, 공정평균의 변화를 탐지하는 절차는 $|Z_t|$가 미리 설정된 관리한계 $\pm c$를 벗어날 경우 이상상태라는 신호를 주게된다.

VSI 방법에서 표본추출간격 H_t는 2가지를 사용하며, VSS 방법에서 표본크기 N_t는 2가지, 3가지, 그리고 4가지를 사용한다. 그러나 표현의 간편성을 위하여 이후의 N_t 표현은 3가지를 사용하는 것에 한정하기로 한다.(N_t를 2가지 사용하는 것은 기존의 많은 연구에 나타나 있으며, 4가지를 사용하는 것은 3가지를 사용하는 경우를 그대로 확장시킬 수 있다.) N_t와 H_t는 $t \geq 2$일 때 관리한계내의 영역에 대하여 다음과 같이 설정할 수 있다.

$$N_t = \begin{cases} n_1 & \text{만일 } |Z_{t-1}| < c_{S1} \\ n_2 & \text{만일 } c_{S1} \leq |Z_{t-1}| < c_{S2} \\ n_3 & \text{만일 } c_{S2} \leq |Z_{t-1}| < c \end{cases}, \quad H_t = \begin{cases} h_1, & \text{만일 } |Z_{t-1}| < c_I \\ h_2, & \text{만일 } c_I \leq |Z_{t-1}| < c \end{cases}.$$

여기서 $n_1 \leq n_2 \leq n_3$, $h_2 \leq h_1$, 그리고 c_{S1}, c_{S2}와 c_I는 각각 3가지 표본크기와 2가지 표본추출간격을 결정짓는 영역의 분계선(threshold limit)을 나타낸다. c_{S1}, c_{S2}, 그리고 c_I의 크기에 따라 다음과 같이 3가지 경우를 고려해 보자.

$$\text{경우 1: } 0 < c_I < c_{S1}, \text{ 경우 2: } c_{S1} \leq c_I < c_{S2}, \text{ 경우 3: } c_{S2} \leq c_I < c.$$

또한 관리한계내의 영역을 다음과 같이 4개 영역으로 구분해보자.

$$I_1 = (-c_1, c_1),\ I_2 = (-c_2, -c_1] \cup [c_1, c_2),$$

$$I_3 = (-c_3, -c_2] \cup [c_2, c_3),\ I_4 = (-c_4, -c_3] \cup [c_3, c_4),$$

여기서 $$(c_1, c_2, c_3, c_4) = \begin{cases} (c_I, c_{S1}, c_{S2}, c), & \text{경우 1에 대하여} \\ (c_{S1}, c_I, c_{S2}, c), & \text{경우 2에 대하여} \\ (c_{S1}, c_{S2}, c_I, c), & \text{경우 3에 대하여} \end{cases}.$$

VSR $\overline{X}$ 관리도에서 사용하는 (N_t, H_t)는 3가지 경우에 대하여 각각 Z_{t-1}에 따라

Z_{t-1}	경우 1	경우 2	경우 3
$Z_{t-1} \in I_1$	(n_1, h_1)	(n_1, h_1)	(n_1, h_1)
$Z_{t-1} \in I_2$	(n_1, h_2)	(n_2, h_1)	(n_2, h_1)
$Z_{t-1} \in I_3$	(n_2, h_2)	(n_2, h_2)	(n_3, h_1)
$Z_{t-1} \in I_4$	(n_3, h_2)	(n_3, h_2)	(n_3, h_2)

와 같이 표현될 수 있다.

3. VSR $\overline{X}$ 관리도의 통계적 효율에 대한 측도

일반적으로 VSR 관리도의 통계적 효율은 이상원인이라는 신호까지의 시간으로 측정될 수 있다. 신호까지의 평균시간(average time to signal: ATS)은 공정의 시작으로부터 관리도가 신호를 줄 때까지의 평균시간을 의미한다. 공정이 관리상태에서 계산된 ATS는 오경보에 관련된 측도이며, 이상상태에서 계산된 ATS는 공정이 $\mu=\mu_1$에서 시작되었을 때 이를 탐지하는 관리도 효율에 대한 측도가 된다. 그러나 많은 경우에 공정은 관리상태, 즉 $\mu=\mu_0$에서 출발하여 진행되다가 어느 시점에서 공정평균이 $\mu=\mu_1$으로 이동하는 것이 대부분이다. 이와 같은 경우 관리도의 효율은 이동이 발생한 시점부터 관리도의 이상신호까지의 평균시간으로 측정되어야 한다. 이에 대한 측도로서 SSATS(steady-state ATS)를 사용할 수 있다. 이 SSATS에 대한 자세한 내용은 Reynolds와 Arnold(2001)를 비롯한 기존의 많은 연구를 참조할 수 있다.

SSATS_δ를 공정평균이 μ_0에서 $\mu_1=\mu_0+\delta\sigma$로 이동한 경우 신호까지의 평균시간이라 하자. SSATS_δ는 Markov chain을 이용하여 계산할 수 있다. Markov chain은 2장에서 정의한 4개의 영역에 해당되는 4가지 일시적 상태(transient state)를 갖는다. 여기서 일시적 상태 전이 확률행렬(transition probability matrix)은 $\mathbf{Q}_\delta=[q_{ij}^\delta]_{4\times4}$로 표현된다. 여기서

$$
\begin{aligned}
q_{ij}^\delta &= \Pr(Z_t \in I_j \mid Z_{t-1} \in I_i) \\
&= \begin{cases} \Phi(c_1-\sqrt{n(i)}\,\delta)-\Phi(-c_1-\sqrt{n(i)}\,\delta), & \text{만일 } j=1 \\ \Phi(c_j-\sqrt{n(i)}\,\delta)-\Phi(c_{j-1}-\sqrt{n(i)}\,\delta) \\ \quad +\Phi(-c_{j-1}-\sqrt{n(i)}\,\delta)-\Phi(-c_j-\sqrt{n(i)}\,\delta), & \text{만일 } j=2,3,4 \end{cases}
\end{aligned}
$$

$$
(n(1), n(2), n(3), n(4)) = \begin{cases} (n_1, n_1, n_2, n_3), & \text{경우 1에 대하여} \\ (n_1, n_2, n_2, n_3), & \text{경우 2에 대하여} \\ (n_1, n_2, n_3, n_3), & \text{경우 3에 대하여} \end{cases}.
$$

Markov chain의 성질과 Reynolds와 Arnold(2001)의 결과를 이용하면 SSATS_δ는

$$\mathrm{SSATS}_\delta = \mathbf{s}'\{[\mathbf{I}-\mathbf{Q}_\delta]^{-1}-\mathbf{I}/2\}\,\mathbf{h}$$

로 나타낼 수 있다. 여기서 $\mathbf{s}'=(s_1, s_2, s_3, s_4)$는 초기확률의 벡터, $\mathbf{I}$는 4×4의 단위행렬, 그리고 $\mathbf{h}$는 원소가 다음과 같은 벡터이다.

$$
\mathbf{h}' = \begin{cases} (h_1, h_2, h_2, h_2), & \text{경우 1에 대하여} \\ (h_1, h_1, h_2, h_2), & \text{경우 2에 대하여} \\ (h_1, h_1, h_1, h_2), & \text{경우 3에 대하여} \end{cases}.
$$

안정상태(steady state)의 초기확률값은

$$s_i = \frac{q_{ii}^0}{q_{i1}^0 + q_{i2}^0 + q_{i3}^0 + q_{i4}^0}, \quad i = 1,2,3,4$$

로 계산할 수 있다(Phabhu, Montgomery와 Runger(1994)와 Reynolds와 Arnold(2001)를 참조).

4. VSR $\overline{X}$ 관리도의 통계적 효율 비교

전통적으로 사용하고 있는 FSR 관리도와 본 논문에서 사용하는 VSR 관리도의 통계적 효율을 비교하기 위해서는 관리상태에서의 수행 능력을 동일하게 하여야 한다. 이것은 공정이 관리상태일 경우 VSR 관리도의 평균표본크기와 평균표본추출간격을 FSR의 고정된 표본크기와 표본추출간격과 동일한 값으로 설정하며, 오경보율을 동일하게 함으로 달성할 수 있다. 따라서 VSR 관리도를 설계할 때 다음과 같은 3가지 제약조건을 고려할 수 있다.

(1) $E[N_t \mid -c < Z_{t-1} < c, \delta = 0] = n_0$,

(2) $E[H_t \mid -c < Z_{t-1} < c, \delta = 0] = h_0$,

(3) $\mathrm{ATS}_0 = A_0$.

여기서 n_0와 h_0는 각각 FSR 관리도에서 고정된 표본크기와 표본추출간격을 나타내며, A_0는 관리상태에서 달성하고자 하는 ATS 값이다.

결국 VSR 관리도(표본크기가 3가지인 VSS 방법을 사용)의 통계적 설계는 위의 3가지 제약조건하에서 SSATS_δ를 최소로 하는 관리모수 ($n_1, n_2, n_3, h_1, h_2, c_{S1}, c_{S2}, c_I, c$)를 찾는 것이다. 제약조건 (1)과 (2)의 기대값을 계산하면, 각각

$$2\Phi(c_{S1})(n_1 - n_2) + 2\Phi(c_{S2})(n_2 - n_3) + 2\Phi(c)(n_3 - n_0) = n_1 - n_0,$$

$$2\Phi(c_I)(h_1 - h_2) - 2\Phi(c)(h_0 - h_2) = h_1 - h_0$$

와 같으며, 이 식들로부터

$$c_{S2} = \Phi^{-1}\left[\frac{(n_1 - n_0) - 2\Phi(c_{S1})(n_1 - n_2) - 2\Phi(c)(n_3 - n_0)}{2(n_2 - n_3)}\right],$$

$$c_I = \Phi^{-1}\left[\frac{(h_1 - h_0) - 2\Phi(c)(h_2 - h_0)}{2(h_1 - h_2)}\right]$$

을 얻을 수 있다. 여기서 $\Phi(\cdot)$는 표준정규분포의 누적분포함수를 나타낸다. 또한 제약조건 (3)으로부터 관리한계 c는

$$c = \Phi^{-1}\left[1 - \frac{h_0}{2A_0}\right]$$

로 계산할 수 있다.

본 논문에서는 VSI 방법에서 표본추출간격은 2가지로 고정하고, VSS 방법에서 표본크기는 2가지, 3가지, 그리고 4가지를 사용하는 VSR $\overline{X}$ 관리도의 통계적 효율을 비교하였다. 이를 위하여 $n_0 = \{3, 5\}$, $h_0 = 1.0$, $A_0 = 370.4$, $\delta = \{0.5, 1.0, 1.5, 2.0, 3.0\}$, 표본크기의 최소값과 최대값은 각각 1과 50, 그리고 표본추출간격의 최소값과 최대값은 각각 0.1과 5.0을 사용하였다.

<표 1>에는 FSR과 VSR $\overline{X}$ 관리도에서 SSATS_δ를 최소로 하는 관리모수들과 그 때의 SSATS_δ 값이 비교되어 있다. 각 경우에서 첫 번째 행은 FSR 관리도, 두 번째 행은 VSS에서 사용하는 표본크기의 수가 2가지, 세 번째 행은 3가지, 그리고 네 번째 행은 4가지인 VSR

관리도의 경우를 나타낸다. 이 표를 살펴본 때 δ가 커질수록 표본크기의 수를 많이 사용하는 것의 효과는 거의 없는 것으로 나타났다. 물론 δ가 큰 경우에는 VSR $\overline{X}$ 관리도 대신 전통적인 FSR $\overline{X}$ 관리도를 사용하는 것만으로 큰 효과를 얻을 수 있다고 알려져 있으므로 위의 결과는 자명한 것이라 판단된다. 그러나 δ가 아주 작은 경우에는 표본크기의 수를 3가지 사용하는 것이 기존에 많이 사용하는 2가지 표본크기에 비하여 통계적 효율이 어느 정도 증가하는 것으로 나타났다. 표본크기의 수를 4가지 사용하는 것은 3가지를 사용하는 것에 비하여 효율이 크게 증가하지 않기 때문에, 아주 작은 δ를 탐지하고자 하며 불량에 대한 비용의 부담이 큰 경우 표본크기의 수가 3가지인 VSS와 표본추출간격이 2가지인 VSI를 병행한 VSR $\overline{X}$ 관리도의 사용은 아주 유용한 일이라고 생각된다.

5. 결론

일반적으로 VSR 관리도는 전통적인 FSR 관리도에 비하여 좀 더 빠른 시간 내에 공정의 이상원인을 탐지할 수 있는 것으로 알려져 있다. 본 논문에서는 VSI 방법에서 표본추출간격을 2가지로 고정하고, VSS 방법에서 표본크기를 2가지, 3가지, 그리고 4가지로 사용하는 VSR $\overline{X}$ 관리도에 대한 통계적 효율을 비교하였다. 그 결과 표본크기를 4가지 사용하는 경우 그 절차의 복잡성에 비하여 효율이 별로 증가하지 않았으며, 공정평균이 변화가 아주 작은 경우 표본크기를 3가지 사용하는 VSR $\overline{X}$ 관리도는 기존의 2가지 사용하는 것에 비하여 효율이 많이 증가함을 알 수 있었다.

본 논문에서는 관리도의 통계적 효율 측면만 생각했지만, 향후 여러 가지 비용을 고려한 경제적 설계를 통하여 적절한 표본크기의 수에 대한 연구가 진행되어야 할 것으로 판단된다.

참고문헌

[1] Park, C. and Reynolds, M. R., Jr. (1994). Economic Design of a Variable Sample Size $\overline{X}$-Charts, *Communications in Statistics: Simulation and Computation*, Vol. 23, 467-483.

[2] Park, C. and Reynolds, M. R., Jr. (1999). Economic Design of a Variable Sampling Rate $\overline{X}$ Chart, *Journal of Quality Technology*, Vol. 31, 427-443.

[3] Prabhu, S. S., Montgomery, D. C., and Runger, G. C. (1994). A Combined Adaptive Sample Size and Sampling Interval $\overline{X}$ Control Scheme, *Journal of Quality Technology*, Vol. 26, 164-176.

[4] Prabhu, S. S., Runger, G. C., and Keats, J. B. (1993). An Adaptive Sample Size $\overline{X}$ Chart, *International Journal of Production Research*, Vol. 31, 2895-2909.

[5] Reynolds, M. R., Jr., Amin, R. W., Arnold, J. C., and Nachlas, J. A. (1988). $\overline{X}$ Charts with Variable Sampling Interval, *Technometrics*, Nol. 30, 181-192.

[6] Reynolds, M. R., Jr. and Arnold, J. C. (2001). EWMA Control Charts with Variable Sample Sizes and Variable Sampling Intervals, *IIE Transactions*, Vol. 33, 511-530.

[7] Zimmer, L. S., Montgomery, D. C., and Runger, G. C. (1998). Evaluation of a Three-State Adaptive Sample Size $\overline{X}$ Control Chart, *International Journal of Production Research*, Vol. 36, pp. 733-743.

<표 1> FSR과 VSR $\overline{X}$ 관리도의 통계적 효율 비교

δ	n_0 / n_1	n_2	n_3	n_4	h_0 / h_1	h_2	c_S / c_{S1}	c_{S2}	c_{S3}	c_I	c	$SSATS_\delta$
0.5	3				1.0						3.00	60.69
	1	29			5.0	0.1	1.79			0.23	3.00	11.83
	1	15	37		5.0	0.1	1.60	2.24		0.23	3.00	10.95
	1	13	25	42	5.0	0.1	1.60	2.00	2.45	0.23	3.00	10.71
1.0	3				1.0						3.00	9.76
	2	9			5.0	0.1	1.46			0.23	3.00	1.47
	2	5	11		5.0	0.1	1.20	1.92		0.23	3.00	1.36
	2	4	7	13	5.0	0.1	1.10	1.60	2.23	0.23	3.00	1.33
1.5	3				1.0						3.00	2.91
	2	6			5.0	0.1	1.15			0.23	3.00	0.69
	2	3	7		5.0	0.1	0.30	1.87		0.23	3.00	0.66
	2	3	5	8	5.0	0.1	0.30	1.70	2.26	0.23	3.00	0.66
2.0	3				1.0						3.00	1.47
	2	4			5.0	0.1	0.67			0.23	3.00	0.55
	2	3	6		5.0	0.1	0.10	2.18		0.23	3.00	0.54
	2	3	4	6	5.0	0.1	0.10	1.90	2.43	0.23	3.00	0.54
3.0	3				1.0						3.00	1.01
	2	4			5.0	0.1	0.67			0.23	3.00	0.51
	2	3	4		5.0	0.1	0.10	1.74		0.23	3.00	0.50
	2	3	4	5	5.0	0.1	0.10	1.80	2.48	0.23	3.00	0.50
0.5	5				1.0						3.00	33.40
	2	29			5.0	0.1	1.58			0.23	3.00	5.55
	1	12	36		5.0	0.1	1.10	2.00		0.23	3.00	4.95
	1	8	20	41	5.0	0.1	1.00	1.60	2.22	0.23	3.00	4.74
1.0	5				1.0						3.00	4.50
	4	12			5.0	0.1	1.52			0.23	3.00	0.82
	4	7	15		5.0	0.1	1.20	2.03		0.23	3.00	0.79
	4	6	10	16	5.0	0.1	1.10	1.70	2.29	0.23	3.00	0.79
1.5	5				1.0						3.00	1.57
	4	8			5.0	0.1	1.15			0.23	3.00	0.55
	4	5	9		5.0	0.1	0.20	2.03		0.23	3.00	0.55
	4	5	7	11	5.0	0.1	0.30	1.70	2.36	0.23	3.00	0.55
2.0	5				1.0						3.00	1.08
	4	6			5.0	0.1	0.67			0.23	3.00	0.51
	4	5	7		5.0	0.1	0.10	2.03		0.23	3.00	0.51
	4	5	6	7	5.0	0.1	0.10	1.80	2.48	0.23	3.00	0.51
3.0	5				1.0						3.00	1.00
	4	6			5.0	0.1	0.67			0.23	3.00	0.50
	4	5	6		5.0	0.1	0.10	1.74		0.23	3.00	0.50
	4	5	6	7	5.0	0.1	0.10	1.80	2.48	0.23	3.00	0.50