

변량표본추출간격을 이용한 범위관리도와 분산관리도의 ATS비교

- A Comparison of Average Time Rate with range and variance chart
when using variable sampling interval -

李喜春*
池善水**

ABSTRACT

The basic of the VSI charts is that if the sample statistic computed after a sample is taken shows some indication of a process change then the sampling interval before the next sample should be short, otherwise long. This paper was shown the VSS chart can be considerably more efficient than the FSS chart and the effectiveness of VSI R chart with S chart used for monitoring a process standard deviation.

1. 서론

산업생산공정에서 생산라인을 조사하여 합리적인 조치를 취하는 것은 양질의 물품을 생산하는데 대단히 중요하며, 공정의 이상유무에 대한 판단을 내리며 그 원인을 찾아내는 일반적인 기법으로 관리도가 있다.

관리도 설계에서 관리한계선, 표본추출간격, 표본크기를 정하는 것이 중요하므로 이에 대한 연구가 활발히 진행되고 있으며 연구의 결과가 실제상황에 쉽게 적용될 수 있도록 간단하면서도 효율적인 관리도 개발에 관한 연구가 진행되고 있다.

관리도 형태는 크게 3가지로 나누어 볼 수 있는데, Shewhart(1931)에 의해 제안된 Shewhart 관리도, Page(1954)에 의해 연구 제안된 누적합(cumulative sum:CUSUM) 관리도, 그리고 Roberts(1959)에 의해 개발된 지수가중이동평균(exponentially weighted moving average:EWMA) 관리도 등이 있다.

지금까지 관리도에 대한 연구는 설계모수가 항상 일정한 값을 가지는 것으로 설계되어 왔으나 최근에는 설계모수를 변량화하여 이상원인을 가능한 빨리 지적할 수 있는 관리도 설계가 연구, 대두되었다. 설계모수중에서 표본크기를 변량화하는 연구는 박창순과 Reynolds(1993)의 변량표본크기(variable sample size)를 이용한 \bar{X} -관리도를 들 수 있으며, 박창순 지선수(1993)는 변량표본크기를 이용한 \bar{X} -R 관리도를 연구하였다.

표본추출간격을 변량화하는 연구는 Reynolds등(1988)에 의해 제안된 변량표본추출간격(variable sampling interval)을 이용한 \bar{X} (VSI \bar{X}) 관리도를 들 수 있으며, 변량추출간격은 표본통계량 값에 따라 다음 표본추출까지의 시간간격을 다르게하는 관리형태를 나타낸다. 즉 표본통계량이 관리한계선 가까이 나타나면 추출간격을 짧게, 그렇지 않으면 추출간격을 길게하여 관리도의 효율을 증대하려는 관리계획이다. 이와같이 표본크기와 추출간격을 변량화하여 관리도의 효율을 증대시키려는 관리형태를 변량추출비율(variable sampling rate)이라 한다.

* 상지대학교 통계학과

** 원주전문대학 사무자동화과

접수 : 1994년 4월 27일

확정 : 1994년 5월 10일

관리도에 대한 연구모델 대부분이 공정평균에 관한 것이었으나 최근에는 공정의 산포에 관한 연구가 진행되고 있으며 이는 정보의 손실측면에서 대단히 중요한 일이라 하겠다. 이 논문에서는 변량표본추출간격을 이용한 R (VSI R) 관리도와 S (VSI S) 관리도를 비교 연구하였다.

2. 변량표본추출간격을 이용한 관리도 설계

관찰된 확률표본값은 평균이 μ 이고 분산이 σ^2 인 정규분포를 따르는 공정에서 택해졌다고 가정한다. t 번째 시점에서 동시에 n개가 택해진 표본을 $X_t = (X_{t1}, \dots, X_{tn})$ 이라 하자. 관리도 적용의 복잡성을 피하기 위해 가능한 단순한 형태를 고려하는데 이 논문에서는 두개의 표본추출간격을 갖고 관리도 과정을 진행한다. 즉 d_1 또는 d_2 두 종류만을 고려한다. 표본으로부터 표본평균과 분산을 $\bar{X}_t = \frac{1}{n} \sum_{j=1}^n X_{tj}$, $S_t^2 = \frac{1}{n-1} \sum_{j=1}^n (X_{tj} - \bar{X}_t)^2$ 와 같이 계산할 수 있으며, 또한 $R_t(n) = \max(X_{t1}, \dots, X_{tn}) - \min(X_{t1}, \dots, X_{tn})$ 이라 하면 다음과 같은 표준화된 통계량을 생각할 수 있다.

$$V_{Rt} = R_t(n) / \sigma_0$$

$$= W_t \cdot \delta \tag{1}$$

$$V_{S_t} = \sqrt{\frac{(n-1) \cdot S_t^2}{\sigma_0^2}}$$

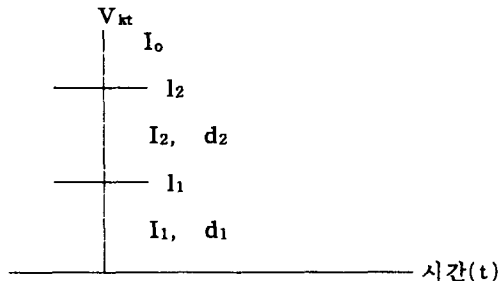
$$= \sqrt{\chi^2(n-1)} \cdot \delta \tag{2}$$

여기서 $\delta = \sigma / \sigma_0$ 이며 σ_0 는 σ 의 관리값(control value)이다. 즉 공정이 정상상태일 때 택해지는 값이다.

위 통계량과 <그림 1>을 참고로 t번째 시점의 표본통계량을 사용하여 다음과 같은 변량표본추출간격을 이용하는 공정관리방법을 제안한다. $V_{Rt} \in I_0$ 이면 공정을 중지한다. $V_{Rt} \in I_1$ 이면 공정을 계속하고 d_1 단위시간 이후에 다음 표본을 추출한다. $V_{Rt} \in I_2$ 이면 현공정을 유지하면서 d_2 단위시간 이후에 다음 표본을 추출한다. 여기에서 $d_1 > 1.0 > d_2$ 이다. 따라서 X_t 와 X_{t+1} 의 표본추출간격을 다음과 같이 표본추출함수로 나타낼 수 있다.

$$d(V_{Rt}) = d_j, \quad V_{Rt} \in I_j, \quad j=1,2, \quad k=1,2$$

그러므로 $d(V_{Rt})$ 는 두 표본사이의 추출간격이 t번째 표본통계량값에 따라 결정됨을 알 수 있다.



<그림 1>. VSI를 이용한 관리도에서 한계영역과 표본추출간격

여기에서 물론 $d_1 = d_2$ 일 경우 일정한 시간간격으로 표본을 추출하는 고정표본추출간격(fixed sampling interval)을 이용한 FSI 관리도의 관리절차와 동일하다.

공정에서 이상원인이 발생한 상태에서 표본통계량 V_{1t} 와 V_{2t} 가 관리한계를 벗어날 확률은 각각 다음과 같이 구할 수 있다.

$$q_R = 1.0 / (1.0 - \psi_n(l_2/\delta))$$

$$q_S = 1.0 / (1.0 - \phi_n(l_2/\delta)) \quad (3)$$

공정이 정상상태와 이상상태일 때 표본통계량 V_{1t} 와 V_{2t} 가 I_1 영역에 떨어질 확률은 아래와 같이 각각 계산할 수 있다.

$$p_{R01} = \psi_n(l_1)$$

$$p_{R11} = \psi_n(l_1/\delta) \quad (4)$$

$$p_{S01} = \phi_n(l_1)$$

$$p_{S11} = \phi_n(l_1/\delta) \quad (5)$$

공정이 정상상태와 이상상태일 때 표본통계량 V_{1t} 와 V_{2t} 가 I_2 영역에 떨어질 확률은 다음과 같이 각각 계산할 수 있다.

$$p_{R02} = \psi_n(l_2) - \psi_n(l_1)$$

$$p_{R12} = \psi_n(l_2/\delta) - \psi_n(l_1/\delta) \quad (6)$$

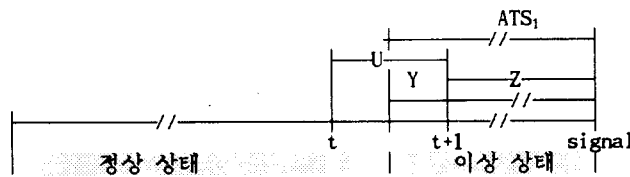
$$p_{S02} = \phi_n(l_2) - \phi_n(l_1)$$

$$p_{S12} = \phi_n(l_2/\delta) - \phi_n(l_1/\delta) \quad (7)$$

$\psi_n(\cdot)$ 는 $W_K = R_K(n)/\sigma$ 의 분포함수이며, $R_K(n)$ 는 분산이 σ^2 인 정규분포로부터 표본크기가 n 개가 택해질 때의 표본의 범위이다. W_K 에 관한 분포는 Pearson과 Hartley (1942)가 제안한 식을 그대로 사용하였다. $\phi_n(\cdot)$ 는 자유도가 $n-1$ 인 카이스퀘어 분포함수의 제곱근을 취한 것이다

3. VSI 관리도를 이용한 ATS 계산

관리도의 효율을 평가할 때 대부분 평균런의길이(average run length :ARL)를 많이 이용하여 왔으나 이 논문에서는 공정에 이상원인이 발생한 경우 신호를 줄 때까지의 평균시간(average time to signal:ATS)을 사용하여 관리도를 평가하는 방법을 사용한다. <그림 2>에서 이상원인이 t 와 $t+1$ 시점에서 발생할 경우 ATS_1 가 가능한 짧게 설계된 것이 최적인 관리도라 할 수 있다.



<그림 2> 공정변화가 t 와 $t+1$ 사이에서 발생할 경우 신호를 줄 때까지의 공정주기

이상상태일 때 신호까지 평균표본수($ANSS_1$)는

$$ANSS_1 = 1/q_m, m=R,s \quad (8)$$

이상원인 발생시점부터 신호까지의 시간을 W 로 한다면 다음과 같이 나타낼 수 있다.

$$W = Y + Z$$

$$= Y + \sum_{i=1}^{M-1} R_i \quad (9)$$

신호까지 표본수(M)가 상수이고 $R_i, i=1, \dots, M-1$,이 독립이고 동일분포를 따르므로 Wald등식을 이용하여

ATS는 다음과 같이 쓸 수 있다.

$$ATS = E(Y) + (EM-1) \cdot E(R_i) \quad (10)$$

위식의 $E(Y)$, EM 와 $E(R_i)$ 를 계산하여 공정에서 이상원인이 발생된 시점부터 신호까지 평균시간 (ATS_{mi})은 다음과 같이 간단히

$$ATS_{mi} = \frac{\sum_{j=1}^2 d_j^2 \cdot p_{m0j}}{2 \cdot \sum_{j=1}^2 d_j \cdot p_{m0j}} + \frac{\sum_{j=1}^2 d_j \cdot p_{m1j}}{q_m} \quad (11)$$

로 쓸 수 있다.

4. 공정모수 결정방법

Pearson(1942)은 범위를 고려한 상태에서의 일반적인 관리한계는 $1.25\sigma \leq R_i(n) \leq 4.49\sigma$ 에서 가장 많이 사용되는 것을 보였다. 이 논문에서는 계산의 간편함을 위해 $I_2=4.16$ 을 이용한다. 통계량 V_R 를 이용한 R 관리도와 V_S 를 이용한 S 관리도를 사용하여 신호까지 평균시간을 비교하여 본다.

공정변화(δ)를 1.0, 1.25, 1.5, 2.0, 2.25, 2.50, 3.0로 변함에 따라 표본크기(n)가 3, 5, 7, 9일 때를 고려한다. d_1 와 d_2 의 범위를 $1.0 \leq d_1 \leq 5.0$ 와 $0.1 \leq d_2 \leq 0.9$ 로 주고 각각의 경우에 ATS_{mi} 가 최소가 되는 모수값(관리한계선(I_1), 표본추출간격(d_1, d_2))을 찾는다. 여기에서 신호까지 평균시간이 최소가 되는 모수를 찾는 방법으로 GRG-II 프로그램을 이용한다. 고정표본추출간격을 이용할 때, 즉 $d_1=d_2=1$ 인 경우의 신호까지 평균시간과 비교하여 효율이 얼마나 개선되는지를 알아본다. 결과를 <표 1>- <표 4>에 수록하며 표에서 %는 $\frac{\text{FSI를 이용한 } ATS_1 - \text{VSI를 이용한 } ATS_1}{\text{FSI를 이용한 } ATS_1}$ 의 값을 나타낸다.

5. 결 론

고정표본추출간격을 이용하는 것보다 변량표본추출간격을 이용하는 것이 공정변화에 따라 R-관리도의 경우 23에서 47 % 빠르게 이상원인을 지적할 수 있으며, S-관리도의 경우 13에서 49 % 정도 빨리 공정변화를 지적함을 확인할 수 있다. 표본크기가 3-5일 경우 공정변화에 대해서 d_2 의 최적값이 항상 0.1에 가깝게 나타나며 d_1 은 변화의 크기에 따라 조금씩 변한다는 것을 알 수 있다. 또한 공정변화가 클수록 VSI R-관리도 보다 VSI S- 관리도가 효율면에서 약간 우수하다는 것을 볼 수 있다. 표본크기가 클 수록 R-관리도(9 이상)에서 공정에서 하나의 표본을 추출하기까지 걸리는 최소의 시간(shortest sampling interval) d_2 와 최대의 시간간격(longest sampling interval) d_1 이 거의 같게 됨을 알 수 있다.

공정의 변화가 I_1 영역에서 발생되면 신호를 주기까지 긴시간동안 공정이 이상상태를 유지하게 되므로 표본추출비(d_1/d_2)가 크게 간격을 조절하면, 즉 d_1 을 크게 하면 오히려 관리도의 효율이 떨어질 수 있다. 따라서 d_1 의 선택이 매우 중요하게 작용되므로 신중히 선택되어야 한다. R-관리도와 S-관리도 모두 표본이 클수록 관리도의 효율을 크게 개선하지 못했으며 $n=7$ 일 경우 오히려 관리도의 효율이 $n=3-5$ 일 경우보다 조금씩 떨어짐을 알 수 있다.

결론적으로 FSI 관리도보다 VSI 관리도가 우수하며 표본크기가 작을 경우 VSI R 관리도를 표본크기가 클수록 VSI S 관리도를 이용하는 것이 유리하다는 것을 알 수 있다. 가능하면 표본크기를 3-5로 정하는 것이 효율적 경제적인면에서 유리하며, 표본통계량이 관리한계선에 가까이 떨어진 영역에 위치하면 표본추출비를 작게, 아니면 크게하여 추출간격을 정하는 것이 바람직하다.

<표 1> n=3, R 관리도(S 관리도)의 모수값.

변화의 크기	l_1	d_2	d_1	%	비 고
1.00	2.980 (1.519)	0.10 (0.10)	1.08 (1.28)		
1.25	1.221 (0.524)	0.10 (0.10)	2.74 (1.86)	0.231 (0.130)	
1.50	1.523 (0.670)	0.10 (0.10)	1.99 (1.70)	0.329 (0.217)	
2.00	1.980 (0.951)	0.10 (0.10)	1.45 (1.49)	0.394 (0.326)	
2.25	2.151 (1.085)	0.10 (0.10)	1.34 (1.43)	0.405 (0.362)	
2.50	2.296 (1.215)	0.10 (0.10)	1.27 (1.37)	0.413 (0.390)	
3.00	2.530 (1.461)	0.10 (0.10)	1.18 (1.30)	0.423 (0.432)	

<표 2> n=5, R 관리도(S 관리도)의 모수값.

변화의 크기	l_1	d_2	d_1	%	비 고
1.00	3.007 (1.679)	0.10 (0.10)	1.22 (1.73)		
1.25	2.037 (1.990)	0.10 (0.10)	2.29 (1.57)	0.288 (0.170)	
1.50	2.412 (2.160)	0.10 (0.31)	1.64 (1.43)	0.372 (0.276)	
2.00	2.939 (2.269)	0.10 (0.67)	1.24 (1.28)	0.414 (0.399)	
2.25	3.132 (2.281)	0.10 (0.77)	1.17 (1.25)	0.423 (0.435)	
2.50	3.293 (2.284)	0.10 (0.84)	1.12 (1.22)	0.431 (0.461)	
3.00	3.543 (2.344)	0.14 (0.93)	1.07 (1.18)	0.446 (0.492)	

<표 3> n=7, R 관리도(S 관리도)의 모수값.

변화의 크기	l_1	d_2	d_1	%	비 고
1.00	3.098 (1.942)	0.10 (0.10)	1.34 (2.06)		
1.25	2.251 (2.316)	0.10 (0.89)	1.96 (1.50)	0.297 (0.192)	
1.50	2.956 (2.512)	0.10 (1.00)	1.45 (1.40)	0.370 (0.302)	
2.00	3.496 (3.468)	0.10 (1.00)	1.15 (1.30)	0.414 (0.398)	
2.25	3.588 (3.931)	0.32 (1.00)	1.10 (1.27)	0.431 (0.416)	
2.50	3.612 (4.160)	0.56 (1.00)	1.07 (1.25)	0.446 (0.426)	
3.00	3.645 (4.160)	0.81 (1.00)	1.04 (1.22)	0.467 (0.432)	

<표 4> n=9, R 관리도(S 관리도)의 모수값.

변화의 크기	l_1	d_2	d_1	%	비 고
1.00	3.214 (1.929)	0.10 (0.10)	1.43 (2.38)		
1.25	2.928 (3.667)	0.10 (1.00)	1.76 (1.49)	0.292 (0.188)	
1.50	3.24 (4.160)	0.10 (1.00)	1.34 (1.43)	0.361 (0.271)	
2.00	3.628 (4.160)	0.55 (1.00)	1.11 (1.43)	0.424 (0.324)	
2.25	3.669 (4.160)	0.76 (1.00)	1.07 (1.43)	0.446 (0.329)	
2.50	3.693 (4.160)	0.88 (1.00)	1.06 (1.43)	0.461 (0.328)	
3.00	3.694 (4.160)	0.99 (1.00)	1.04 (1.43)	0.473 (0.323)	

참고문헌

1. Abramowitz M. and Stegun I. A., "Handbook of mathematical function ", Dover Publication, New York, 1964.
2. Box G. E. P. and Jenkins G. M., "Time series analysis: forecasting and control", San francisco, Holden-day, U.S.A., 1976.
3. Crowder S. V., "Computation of ARL for combined indeividual measurement and moving range chart", J.Q.T, vol 19. No. 2, 1987.
4. Ji S. S., "A study on the R chart with variable samling interval", unpublished paper, 1994.
5. Lasdon L. S. and Allan D. W., "GRG2 User's guide", 1983.
6. Montgomery D. C. "Introduction to Statistical Quality Control Chart", New Yprk, John Wiley&Sons, 1985.
7. Park C. S. and Ji S. S., "An economic design of \bar{X} -R Chart with variable sampling size scheme", Ph. D. thesis, Chung-Ang University, korea, 1993.
8. Pearson E. S. and Hartley H. O., "The probability integral of the range in samples of n observation from a noraml population", Biometrika, vol 32, pp 301-310, 1942.
9. Reynolds Jr. M. R., Amin R. W., Arnolds J. C. and Nachlas J. A. , " \bar{X} charts with variable sampling interval", Technometrics, vol 30, no 2, pp 181-192, 1988.
10. Reynolds Jr. M. R., " Optimal VSI control charts", Sequential Analysis, pp 361-379, 1989.