Zone, 다변량 T^2, ARIMA를 이용한 통합관리도의 적용방안
- Implementation of Integrated Control Chart Using Zone, Multivariate T^2 and ARIMA -

최 성 운*
Sung-Woon Choi*

Abstract

The research discusses the implementation of control charts tools of MINITAB which are classified according to the type of data and the existence of subgrouping, weight and multivariate covariance. The paper presents the three integrated models by the use of zone, multivariate T^2-GV(Generalized Variance) and ARIMA(Autoregressive Integrated Moving Average).

Keywords : Implementation, MINITAB, Integrated Models, Zone, Multivariate T^2-GV, ARIMA

1. 서 론

식스시그마 운동에서 통계적 품질관리(SQC, Statistical Quality Control)와 통계적 공정관리(SPC : Statistical Process Control)는 품질 개선 및 혁신을 위한 중요한 도구이다. 샘플링 검사(Acceptance Sampling)는 제품스펙(Specification)의 검사항목에 대한 사후관리적인 SQC 방법으로 로트(Lot) 불합격시 폐기, 재작업, 특채, 용도변경 등의 조치를 취한다. 반면에 관리도(Control Chart)는 제품스펙의 관리항목의 이상원인(Assignable Cause)을 설비, 금형, 치공구의 점검항목으로 즉시적으로 개선하는 사전관리적인 SPC 도구이다.

* 경원대학교 산업공학과

우리나라에서 매년 개최되고 있는 대통령상 국가품질경영대회의 식스시그마와 품질분임조 수상기업에서 주로 사용하는 기법으로는 검정, 추정, 실험계획법과 더불어 SPC의 관리도, 히스토그램, 공정능력지수 등이 있다. 이 기법들 중에서 관리도는 주제선정(Define), 현상파악(Measure), 결과분석 및 사후관리(Control) 단계에서 사용되는 대표적인 기법이나 이에 대한 전문가, 심사원, 기업실무자의 이해부족으로 관리그래프로 대치하거나 관리용 관리도의 오적용이 모든 발표원고에서 공통적으로 나타나고 있다.

현상파악단계에서 제조기업인 경우 통상 1달간의 일별 관리도를 작성하여 부서목표치에 대한 평균데이터와 비교를 하기 전에 산포의 원인을 정상적으로 인정하는 우연원인(Random Cause)과 피할 수 있는 이상원인(Assignable Cause)으로 구분하고 이상원인인 평균데이터는 원인분석과 Quick-Fix 또는 장기개선의 과제로 넘겨야 한다..

이 경우 관리도의 CL(Central Line), UCL(Upper Control Limit), LCL(Lower Control Limit)이 부서목표치를 긋는 선과 혼동되고 관리도 사용이 실무자에게 부담스럽다고 꺾은선 그래프에 부서목표치를 기입하는 관리그래프 사용이 일부 전문가에 의해 잘못 지도교육되어 적용되고 있는 실정이다. 이는 산포가 큰 평균은 의미가 없다는 기본적인 통계이론의 무지에서 오는 오류로 반드시 시정해야 할 부분이다. 또한 개선을 실시한 후 표준화를 유지하는 사후관리(Control) 단계의 관리용 관리도는 1달을 기준으로 월초에 미리 관리한계를 긋고 모니터링하는 방법으로, 1달을 기준으로 월말에 관리한계를 나중에 작성하는 현상파악(Measure) 단계의 해석용 관리도와 똑같이 작성하는 오류를 범하고 있다.

따라서 본 연구에서는 국가품질경영대회에 출전한 기업에서 가장 많이 사용하고 있는 MINITAB 통계 패키지에서 지원하고 있는 관리도의 적용방안을 논의한다. 또한 Zone관리도[8]와 관리그래프의 통합 모형, 다변량 T^2관리도[2,5,6,9]와 PCA(Principal Component Analysis)의 통합모형, ARIMA(Autoregressive Integrated Moving Average)[1,3,4,7]와 I-MR관리도의 통합모형을 제시한다.

2. MINITAB 지원 관리도의 적용방안

MINITAB에서 지원하는 관리도의 종류로는 부분군(Subgroup) 계량연속형 관리도, 개별값 계량연속형 관리도, 계수이산형 관리도, 시간가중 관리도, 다변량 관리도 등이 있다.

부분군 계량연속형 관리도 종류로는 $\overline{x}-R$관리도, $\overline{x}-S$관리도, I-MR-R/S(군간/군내)관리도, Zone관리도가 있다. 부분군 관리도에서 군내표준편차(σ_w)는 작게 하고 군간 표준편차(σ_b)는 크게 하는 합리적인 군구분(Rational Subgrouping) 원칙이 선행되어야 한다. 이는 부분군(Subgroup)을 5M 1J PE(Man, Machine, Material, Method, Measurement, Jig, Part, Fixture)에 따라 층별된 로트나 배취로 구성되어야 한다는 원리에 있다. 군내 표준편차는 피할 수 없는 랜덤한 우연원인으로만 구성되고 군간 표준편차는 피할 수 있는 개선을 필요로 하는 이상원인으로 탐지될 수 있도록 부분군의 샘플링을 실시하여야 한다. 이상원인이 있는 5개의 로트 또는 3개의 작업반(Shift)을

부분군으로 형성할 경우 합리적인 군구분 원칙의 층별 오류로 관리도의 기능을 상실하게 된다. 종합 표준편차 $\sigma_o = (\sigma_b^2 + \sigma_w^2)^{1/2}$로 계산되며 관리도가 정상(In-Control)으로 판정된 경우 군내 표준편차(σ_w), 이상(Out-of-Control)으로 판정된 경우 종합 표준편차(σ_o)를 각각 사용하는 공정능력지수(PCI : Process Capability Index)와 공정성능 지수(PPI : Process Performance Index)가 있다. 여기서 PPI는 초기공정 셋팅 후 안정화하기까지의 이상원인이 되는 생산기술조건을 모니터링하고 개선하기 위해 사용되는 방법으로 일부 전문가와 기업 실무자가 정확도의 1.5σ 치우침의 정적(Static), 동적(Dynamic)개념과 혼동하여 단기, 장기의 용어로 사용하고 있다. R관리도는 계산기가 없던 시절 간편하게 사용해 오던 방법이나 n이 커질 경우 관리도의 성능이 떨어진다. S관리도는 계산기가 없으면 구하기가 어려우나 n이 커질 수록 관리도의 성능이 좋아진다. MINITAB의 사용자는 군구분 크기(Subgroup Size) n이 클 경우 이상값에 영향을 덜 받으면서 계산의 부담이 없는 S관리도의 이용을 권장한다. 1로트 또는 배취에서 몇 개씩 부품 또는 제품을 생산 측정하는 경우 군내 표준편차가 작아지거나 커지게 되어 $\bar{x}-R/S$(군내)관리도를 작성할 경우 관리도를 하나도 벗어나지 않거나 대부분 벗어나는 기현상이 일어난다. 이 경우 부분군간 평균은 I관리도, 부분군간 이동범위는 MR관리도, 부분군내 정밀도는 R/S관리도로 작성하는 도구가 $I-MR-R/S$관리도이다. 이는 $\bar{x}-R/S$관리도가 군내 정밀도에 의해 작성되는 문제점을 보완하는 방법으로 군간 정밀도를 MR관리도에 의해 파악한다. Zone 관리도는 3.1절에서 언급하기로 한다.

개별값 계량연속형 관리도에는 $I-MR$관리도, $Z-MR$관리도가 있다. $I-MR$관리도는 부분군을 형성할 수 없거나 의미가 없는 경우 즉 n=1로(그러나 관리도 계수표를 찾을 경우는 이동범위를 고려해주는 개수의 의미로 사용되므로 n=2로 사용) $x-Rs$로도 사용되고 있다. 이 관리도는 장치산업에서 수율(Yield)의 관리항목에 대해 작성되며 전문가나 기업의 실무자가 수율에 대한 개념을 이해하지 못하여 단위가 %라하여 계수 이산형 P관리도를 적용하는 실수를 자주 저지르고 있다. 수율은 자재의 생산성(Productivity)으로 Output인 제품의 중량(ton, Kg)을 Input인 자재의 중량(ton, Kg)으로 나누어 산출하며 작은 소숫점을 큰 정수로 만들어 주기 위해 통상 100을 곱한 퍼센트(%)의 단위를 사용한다. 여기서 수율의 단위는 결과적으로 %이지만 원래의 산출식에서 ton, Kg을 사용했기 때문에 $I-MR$계량연속형 관리도를 적용해야 한다. 구매

또는 에너지 사용부서에서는 수율의 역수인 원단위(Standard Material Usage)를 사용하는데 이 역시 $I-MR$계량연속형 관리도를 작성해야 하며 TPM분임조에서 시간가동률향상을 위한 손실유휴시간 역시 $I-MR$관리도를 적용해야 한다. 다품종 소량생산(Short-Run Production)인 경우 부분군 로트의 단위를 없애고 정규화(Normalizing)하는 표준화 Z = (데이터 - 평균)/표준편차 또는 Z = (데이터-기준 공칭 명목치수)/허용차로 변환된 $Z-MR$관리도를 적용한다.

계수이산형 관리도에는 부적합품의 P, NP 관리도와 부적합의 C, U 관리도가 있다.

계측기를 사용하여 물리적, 화학적 측정 분석 단위를 갖는 계량연속형 관리도와 다르게 계수이산형 관리도의 기초단위는 개(EA : Each)이다. 부적합이 2개이상일 경우의 Unit에 대해 부적합품이라 한다. 부적합품에서 부분군의 크기(n)가 일정한 소품종 대량 계획생산에서는 NP관리도를, 일정하지 않은 다품종 소량 주문생산에서는 P관리도를 사용한다. 여기서 부적합품률 P = 부적합품수(x)/부분군의 크기(n)로 소숫점의 숫자를 크게 하기 위해 10^2,10^6,10^9을 각각 곱하여 %, PPM(Parts Per Million), PPB(Parts Per Billion)의 P관리도를 사용한다.

로트산업에서 수율 = (Output 제품 개수/Input 자재 개수)*100(%)의 단위는 %이지만 원래의 산출식에서 단위가 개수이므로 이러한 산업에서의 수율은 P관리도를 적용해야 한다. 앞에서 언급한 바와 같이 철강, 화학, 식품, 제약의 장치배취산업에서 $I-MR$계량연속형 관리도를 적용해야 하는 데도 불구하고 수율 단위가 %라 하여 P계수이산형 관리도를 잘못 적용하는 경우가 빈번하게 발생한다. 부적합인 경우 부적합품 관리도와 같이 계획생산인 경우 C관리도를, 주문생산인 경우 U관리도를 작성한다.

시간가중 관리도에는 MA(Moving Average)관리도, $EWMA$(Exponentially Weighted Moving Average)관리도, $CUSUM$(Cumulative Sum)관리도 등이 있다. MA관리도는 전체 부분군의 수(k)를 사용하지 않고 이동평균의 길이만큼만 데이터를 고려하는 방법이나 기간 설정에 객관성이 결여될 수 있다. $EWMA$관리도는 현재에 비중을 많이 주고 과거로 올라갈수록 비중을 지수적으로 적게 주는 기하이동평균(Geometric Moving Average)의 가중관리도이나 가중치의 주관성 부여에 따라 관리도의 성능이 좌우될 수 있다.

다변량 관리도는 T^2-GV(Generalized Variance)관리도와 $MEWMA$(Multivariate EWMA)관리도가 있다. $MEWMA$관리도 역시 $EWMA$관리도와 마찬가지로 가중치 설정의 객관성 문제와 평균 런 길이 ARL(Average Run Length)에 대한 실무에서의 개념의 이해 및 적용의 어려움이 있다. T^2-GV관리도는 3.2절에서 언급하기로 한다.

3. 통합관리도의 적용방안

3.1 Zone관리도와 관리그래프의 통합모형

관리그래프는 데이터의 꺾은선 그래프에 회사의 목표치를 그어 평균값과 비교하는 원시적인 방법이다. 최근 우리나라의 국가품질경영대회에서 일부 전문가의 잘못된 지도와 사용의 용이성으로 인한 기업실무자의 무조건적인 도입으로 품질 개선시 그릇된 현상파악(Measure)과 사후관리(Control) 활동을 펼치고 있다. 정밀도가 나쁜 정확도는 무의미하다는 통계의 기본원리를 무시한 관리그래프에서 산포를 고려하지 않고 데이터의 평균값과 회사의 목표치를 단순 비교하여 공정의 해석을 내려서는 절대 안된다. 만약 중소기업과 같이 인적, 물적여건이 어려워 Shewhart 3σ관리도의 교육이 현실적으로 불가능한 경우 기업실무자가 이해하기 쉬운 Zone관리도를 통합적용해야 한다.

MINITAB에서 지원하는 Zone관리도는 $\overline{x}$(I)관리도와 CUSUM관리도를 혼합하여 구역마다 점수(Score)를 부여해 판정하는 간단한 도구로 정규분포에 대한 가정과 데이터의 이상원인에 대한 전문적인 패턴구분이 필요없다. 1σ이내는 Zone 1, 1σ와 2σ사이는 Zone 2, 2σ사이와 3σ 사이는 Zone 3, 3σ밖은 Zone 4로 각 Zone에 한계 점수 0, 2, 4, 8를 부여하고 데이터의 누적점수가 한계점수보다 클 경우 이상원인으로 탐지한다. 관리그래프의 부서 목표치 비교기능과 Zone관리도의 이상원인 탐지 기능의 통합과 더불어 2가지 도구 모두 사용이 간편하다는 실무적 장점이 있다.

3.2 다변량 T^2관리도와 PCA의 통합모형

MINITAB에서 지원하는 다변량 T^2-GV(Generalized Variance)관리도는 단변량 $\overline{x}-S$관리도의 확장형태로 관리항목 x간 공분산(Covariance) 또는 상관(Correlation)이 존재할 경우 적용된다. 예를 들어 인장강도와 신율은 음의 상관관계가 존재하며 두 스펙의 관리항목을 동시에 관리하는 경우 다변량 T^2-GV가 적용된다. 그러나 이 관리도에서 정상이 아닐 경우 어떤 스펙의 관리항목이 이상인지를 파악하기가 곤란하다.

이에 대한 연구가 상당히 많이 진행되고 있으나 이론적인 통계분석 및 설계, 평가방법에 치우쳐 기업 실무에서 적극적으로 활용되지 못하고 있는 실정이다. 따라서 MINITAB 사용자는 다변량 T^2-GV에서 관리이탈(Out of-Control)로 판정된 경우 주성분 분석(PCA : Principal Component Analysis)을 사용하여 이상원인을 구체적으로 파악해야 한다. PCA는 공분산을 고유 벡터(Eigen Vector)와 고유치(Eigen Value)의 스펙트랄 분해(Spectral Decomposition)로 표현하는 방법이다. Scree 도표에 의해 고유치의 기여율로 원래의 변수 x 또는 표준화 변수 Z와 고유벡터 e로 변환하는 방법이다. 따라서 다변량 T^2-GV관리도는 이상의 탐지(Detection)기능으로, PCA는 이상원인의 해석(Analysis)으로 통합사용하는 것이 바람직하다. 예를 들어 철강제품의 NTB(Nominal-The-Best)인 치수 스펙, STB(Smaller-The-Better)인 불순물의 함유량, 신율 스펙, LTB(Larger-The-Better)인 인장강도, 경도, 순도 스펙에 대해 1단계에서 다변량 T^2-GV관리도를 적용하여 이상원인을 탐지하고 2단계에서 PCA에 의한 주성분별 이상원인을 해석한다. 2단계의 PCA에서 $y=ex=ez$에 의해 $y_1=e_1z_1$, $y_2=e_2z_2+e_3z_3-e_4z_4$, $y_3=e_5z_5-e_6z_6$으로 변환된 경우 y_1은 치수, y_2는 물리적 스펙인 인장강도, 경도, 신율을, y_3는 화학적 스펙인 순도와 불순물의 함유량을 나타내며 이를 근거로 이상원인의 해석이 가능하다. 1단계의 다변량 T^2-GV관리도는 해석용 관리도로 사후관리에서 사용하는 다변량 χ^2-GV관리용 관리도가 있으며 관리한계는 <표1>과 같다.

<표1> 다변량 관리도의 관리한계

	다변량 T^2 해석용 관리도	다변량 χ^2 관리용 관리도
플롯되는 점	$T^2 = n(\overline{x} - \overline{\overline{x}})S^{-1}(\overline{x} - \overline{\overline{x}})$	$\chi^2 = n(\overline{x} - \mu)\Sigma^{-1}(\overline{x} - \mu)$
관리 한계	UCL= $(p(k-1)(n-1))/(k(n-1)-(p-1))\cdot F_{1-\alpha}(k(n-1)-(p-1))$	UCL = $\chi^2_{1-\alpha}(p)$

GV관리용 관리도는 $W_k = -pn + pn\ln(n) - n\ln(|A_i|/|\Sigma|) + tr(\Sigma^{-1}A_i)$ (여기서 $A_i = (n-1)S_k$), $UCL = \chi^2_{1-\alpha}(p(p+1)/2)$이며 해석용 관리도의 $CL = (|S|/b_1)b_1$, $UCL, LCL = (|S|/b_1)(b_1 \pm 3b_2^{1/2})$이며 $b_1 = (1/(n-1)^p)\cdot\prod_{j=1}^{p}(n-i), b_2 = (1/(n-1)^{2p})\prod_{j=1}^{p}(n-i)[\prod_{j=1}^{p}(n-j+2) - \prod_{j=1}^{p}(n-j)]$이며 플롯되는 점은 $|S_k|$이다.[8] 관리용 관리도에서 Σ는 목표 공분산이고 해석용 관리도에서 S는 예비 데이터에 의해 구해진다.

3.3 ARIMA와 $I-MR$관리도의 통합모형

스펙의 관리항목간의 공분산 또는 상관관계에 의한 다변량 T^2-GV관리도와 다르게 측정 데이터간 시계열로 자동 상관된(Serially Autocorrelated) 경우 Box-Jenkins가 제안한 ARIMA(Autoregressive Integrated Moving Average) 모형을 사용한다. 비정상(Nonstationary) ARIMA(p, d, q)는 자동상관 p차원, 차분(Differencing) d차원, 이동범위 q차원으로, 정상(Stationary) 시계열 ARIMA(p,q)는 ARIMA의 차분으로 구해질 수 있다. ARIMA(p,q) 모형은 식별(Identification), 추정(Estimation), 진단(Diagnosis)의 3단계로 분석, 평가된 후 구해진 iid(Independently and Identically Distributed) 잔차(Residual)에 I-MR관리도를 적용하는 통합모형을 사용할 경우 기업 실무자가 손쉽게 해석이 가능하다.

4. 결 론

본 연구에서는 MINITAB에서 지원되는 부분군 계량연속형 관리도, 개별값 계량연속형 관리도, 계수이산형 관리도, 시간가중 관리도, 다변량 관리도의 활용 및 적용방안을 제시하였다. 사용의 용이성을 위한 효율성 관점과 관리도의 이상원인의 탐지 및 해석 기능의 효과성 관점에서 Zone 관리도와 관리 그래프, 다변량 T^2-GV관리도와 PCA, ARIMA와 $I-MR$관리도의 3가지 통합모형의 적용방안을 제안하였다.

5. 참 고 문 헌

[1] 김은정, 최병선, 미니탭을 사용한 시계열 분석 입문, 자유아카데미, 1990.
[2] 최용석, 정광모, 실무자를 위한 MINITAB 다변량 분석, 이레테크, 2001.
[3] Box G.E.P. Jenkins G.M., Reinsel G.C., Time Series Analysis : Foecasting and Control, 4th Edition, Wiley, 2008.
[4] Choi S., Lee S., " Serially Correlated Process Monitoring Using Foreward and Backward Prediction Errors From Linear Prediction Lattice Filter ", 품질경영학회지, 61(4) (1998): 143-150.
[5] Hawkins D.M., Choi S., Lee S., " A General Multivatiate Exponentially Weighted Moving Average Control Chart ", Journal of Quality Technology, 39(2)(2007): 118-125.
[6] Jackson J.E., A User's Guide to Principal Components, Wiley, 2003.
[7] Lee S., Choi S., " Adaptive Process Monitoring Using Scale CUSUM for Serially Correlated Process ", Computer and Industrial Engineering, 33(3) (1997): 737-740.
[8] Montgomery D.M., Introduction to Statistical Quatity Control, 6th Edition, Wiley, 2008.
[9] Wilkinson J.H., The Algebraic Eigenvalue Problem, Oxford University Press, 1996.