

응용논문

개별 관측치에서 지수변환을 이용한 EWMA 관리도 적용기법
 - EWMA chart Application using the Transformation of the
 Exponential with Individual Observations -

지 선 수¹⁾
 Ji, Seon Su

Abstract

The long-tailed, positively skewed exponential distribution can be made into an almost symmetric distribution by taking the exponent of the data. In these situations, to use the traditional shewhart control limits on an individuals chart would be impractical and inconvenient. The transformed data, approximately bell-shaped, can be plotted conveniently on the individuals chart and exponentially weighted moving average chart. In this paper, using modifying statistics with transformed exponential of the data, we give a method for constructing control charts. Selecting method of exponent for individual chart is evaluated. And consider that smaller weight being assigned to the older data as time process and properties and taking method of exponent(θ), weighting factor(α) are suggested. Our recommendation, on the basis result of simulation, is practical method for EWMA chart.

1. 서론

산업 생산공정라인에서 공정상태를 조사하여 적절한 조치를 취하는 것은 좋은 품질을 생산하는데 대단히 중요하며, 공정의 이상유무에 대한 판단을 내리며 그 원인을 찾아내는 대표적인 기법으로 관리도가 있다. 최적 관리도 설계를 위해 많은 학자들에 의해 여러 가지 기법이 제안되어 관리도의 효율성을 증가시켜왔다. 관리도의 일반적인 형태는 W.A. Shewhart(1931)에 의해 Shewhart 관리도가 제안되었으며, E.S. Page(1954)는 누적합(cumulative sum:CUSUM) 관리도를 연구하였다. 또한 S.M. Roberts(1959)에 의해 개발된 지수 가중 이동평균(exponentially weighted moving average:EWMA) 관리도 등이 있으며 이들 관리도들은 여러 학자들에 의해 연구, 발전되어 왔다.

Shewhart 관리형태는 제조와 서비스공정 모두에 품질개선도구로서 이용되어 왔으며 이상원인을 지적하는 기본적인 법칙은 3σ 관리한계선 밖에 혹은 안에 통계량이 위치하는 것을 조사하는 것이다. 즉, 관리한계선 안쪽으로 통계량이 위치해 있으면 정상상태로, 바깥쪽에 위치해 있으면 공정이 이상상태임을 지적한다. 전통적인 관리도 기법에서 몇 가지 개선이 이루어졌는데 가장 개선이 많이 이루어진 것은 CUSUM 관리도의 도입이었다. CUSUM 관리도는 공정의 작은 변화를 지적하는데 있어 민감도를 향상시켰으나 이용면에서는 Shewhart 관리도보다 쉽게 적용하지 못한다는 단점이 있다. 즉, 계산의 복잡성과 적용의 어려움 때문에 실제 사용에 있어 많은 제약이 따른다. EWMA 관리도는 전통적인 Shewhart 관리도의 구조를 개선하여 개발하

1) 원주대학/사무자동화과

였으며 공정의 작은 변화와 큰 변화 모두에 관리능력이 향상되었으며 적용면에서도 용이하다. C 관리도는 생산공정에서 제품 단위당 결점수를 관리하기 위해 고안된 관리도이며 포아송 분포를 따르는 사건들을 분석하는데 매우 유용하게 이용된다. 각각의 검사단위에는 여러 종류의 결점이 있을 수 있는데, 결점의 종류를 구분하지 않고 관리하고자 할 때 사용한다. 일반적으로 결점수를 나타내는 확률변수가 평균치 \bar{C} (결점수)인 포아송 분포를 따르므로 t 번째 얻은 시점에서 얻은 결점수는 3σ 관리한계를 고려한다면 $\bar{C} \pm 3\sqrt{\bar{C}}$ 으로 관리한계선을 설정한다. 여기에서 \bar{C} 는 「검사에서 발견된 총 결점수/총 검사 단위수」로 결정한다.

생산공정에서 다음과 같은 과정을 고려한다. 0 시점부터 t-1 시점까지는 신호가 없었으나 t-1 시점과 t 시점 사이에서 이상신호가 있었다. 그러나 C 관리도에 표시해 본 결과는 정상으로 판정되었다. t 시점과 t+1 시점 사이에 신호가 있었으나 마찬가지로 정상으로 판정되었다. 여기에서 공정라인에서는 실질적으로 어떤 일이 발생했으며 전통적인 C 관리도에서는 신호를 감지하지 못했다면 공정의 개선을 위해 C 관리도를 보완할 수 있는 방법을 제시할 필요가 있다. 이것에 관련된 연구는 Draper와 Cox(1969), Nelson(1994), White, Keats, Stanley(1997) 등에 의해 진행되어 왔다.

이 논문에서는 C 관리도를 이용한 공정개선을 위해 이상신호가 발생하는 시간 간격을 기준으로 X와 EWMA 관리도를 적용하기 위한 자료변환의 타당성을 조사해 본다. EWMA 관리도 적용시 적절한 가중치 적용방법과 자료변환을 위한 지수와의 관계에 대해 평가한다. 지수분포 과정의 확률수를 구하여 X와 EWMA 관리도에서 기존에 사용하던 관리한계를 적용할 때와 지수 변환자료를 이용한 경우 각각의 관리도 효율성을 평가한다.

2. 지수분포와 변환

주어진 단위시간 안에 발견된 결점의 수는 근사적 포아송 과정을 기초로 공정관리를 진행할 수 있으며 결점이 발견되는 기간을 고려할 경우 다음과 같은 지수분포를 이용할 수 있다. 확률 변수 x는 공정변화가 일어날 때까지의 대기기간이다. 따라서 λx 는 포아송 과정에서 x 단위구간에서 발생하는 사건의 평균횟수를 나타낸다.

$$f(x) = \lambda e^{-\lambda x}, \quad 0 \leq x \leq \infty \quad (1)$$

여기에서 λ 는 단위시간당 사건의 발생 비를 나타내며 $E(x) = 1/\lambda$, $V(x) = 1/\lambda^2$ 이다. 일반적인 관리도에서 공정변화에 있어서 변화를 보고자 하는 품질특성이 정규분포를 가정하면 그 측정치는 공정에 이상요인이 작용하지 않는 한 평균치를 중심으로 $\pm 3\sigma$ 안에 99.73%가 포함된다. 이는 이론적 기초를 바탕으로 출발한다. 따라서 관리하한선(LCL)과 관리상한선(UCL)을 이용하여 특성치가 벗어날 확률은 다음과 같이 쓸 수 있다.

$$P(x \leq LCL) = 0.00135 \quad (2)$$

$$P(x \geq UCL) = 1 - P(x \leq UCL) = 0.00135 \quad (3)$$

(2)식과 (3)식을 이용하여 관리한계선을 구할 수 있다. 관리하한선(LCL)은 다음과 같이 쓸 수 있다.

$$LCL = -\frac{1}{\lambda} \ln(0.99865) \quad (4)$$

관리상한선(UCL)은 다음과 같이 계산할 수 있다.

$$UCL = -\frac{1}{\lambda} \ln(0.00135) \tag{5}$$

예를 들어 어떤 생산공정에서 공정변화가 일어날 평균 대기시간이 12인 지수분포 과정을 따른다고 한다면 $E(x) = \frac{1}{\lambda} = 12$ 이 되므로 $\lambda = \frac{1}{12}$ 이 된다. 여기에서 관리한계선을 구하면 $LCL = -12 \ln(0.99865) = 0.01621$, $UCL = -12 \ln(0.00135) = 79.2918$ 이 된다. 그러나 관리한계의 폭이 넓어 효율적이지 못하다.

각 표본점에서 하나 이상의 측정치를 얻는 것이 적당하지 않는 경우 개별 관측치를 고려한 공정관리 연구가 많은 학자들에 의해 진행되었다. 개별 관측치를 이용하여 일반적인 Shewhart 관리도를 이용할 경우 이러한 관리한계의 적용은 현실적으로 유용하지 않다고 증명되었다. 많은 학자들에 의해 관측치의 정보를 잃지 않는 조건하에 근사적인 종형분포를 하여 Shewhart 관리도를 쉽게 적용할 수 있도록 하는 적절한 자료변환 방법에 대한 연구가 진행되어 왔다. 또한 공정관리를 진행하는 현장에서 관리도 적용을 위한 단순한 기법이 필요하게 되었다.

Nelson(1994)은 0.2777 지수변환에 의한 자료를 이용하여 적용하였으며, Draper, Cox(1969)는 0.268, Taylor(1985)는 0.2654 지수변환에 의해 얻어진 자료를 이용하는 것에 대한 타당성을 제시하였다. Rudolf(1999)는 계산의 간편함과 공정관리과정에서 쉽게 적용하기 위해 0.25 지수변환 방법을 제안하였으며, 적용결과 평균과 중앙값과의 차이는 0.7% 정도로 Nelson이 제안한 0.2% 보다는 못하지만 실제 적용면에서 유리함을 장점으로 제시하였다.

<표1> 지수분포의 지수 변환에 따른 평가값

구분	exponent(θ)				
	0.25 (Rudolf,1999)	0.2654 (Taylor,1985)	0.2777 (Nelson,1994)	0.2907 (Rudolf,1999)	0.3333 (Rudolf,1999)
평균	0.9064	0.9034	0.9011	0.8989	0.8930
중앙값	0.9124	0.9085	0.9032	0.8989	0.8850
표준편차	0.2543	0.2698	0.2779	0.2890	0.3246
skewness	-0.087	-0.038	0.0	0.041	0.168
kurtosis	2.748	2.732	2.717	2.711	2.729

3. 공정관리 과정

공정라인으로부터 얻은 관측치를 조사한 결과 이상신호가 발생된 시점의 자료만을 고려한다면 시점 t에서의 자료는 지수분포 과정을 따르며 $x_t, t=1,2,3, \dots$ 으로 나타낼 수 있다. 확률변수 x는 실제적용에서의 단순성을 위해 다음과 같은 변환자료를 사용한다.

$$y_t = x_t^\theta$$

Nelson등에 의해 θ 값이 0.25 내외인 경우 변환자료 y_t 는 근사적인 종형분포를 따른다는 것이 증명되었으며, 지수분포의 변환은 일반적으로 skewness가 0에 가깝고, kurtosis가 3 정도 된다. 그러므로 변환된 자료를 Shewhart 관리도에 적용시키는 것은 큰 무리가 없다. 그러나 kurtosis

가 3보다 작을 경우 관리한계를 계산할 때 표준편차를 사용하면 관리한계의 폭이 지나치게 넓게 되므로 사용을 하지 않는 것이 좋으므로 다음의 방법에 의해 관리한계를 설정하는 것이 일반적이다.

3.1 X 관리도

X 관리도의 관리한계는 (4)식을 이용하여 변환자료의 적당한 관리하한선을 다음과 같이 계산할 수 있다.

$$LCL = \left[-\frac{1}{\lambda} \ln(0.99865) \right]^{\theta} \tag{6}$$

중심선은 $CL = \left[-\frac{1}{\lambda} \ln(0.5) \right]^{\theta}$ 으로 정할 수 있으며 X 관리도의 관리한계는 (5)식을 이용하여 변환자료의 적당한 관리상한선을 다음과 같이 얻을 수 있다.

$$UCL = \left[-\frac{1}{\lambda} \ln(0.00135) \right]^{\theta} \tag{7}$$

지수분포의 기대값은 $\frac{1}{\lambda}$ 이고 $\frac{1}{\lambda}$ 의 최우추정(MLE) 값은 \bar{X} 이므로 (6)과 (7)식은 다음과 같이 쓸 수 있다.(Rudolf, 1999)

$$LCL_x = \left[-\bar{X} \ln(0.99865) \right]^{\theta} \tag{8}$$

$$CL_x = \left[-\bar{X} \ln(0.5) \right]^{\theta} \tag{9}$$

$$UCL_x = \left[-\bar{X} \ln(0.00135) \right]^{\theta} \tag{10}$$

t 시점에서의 변환된 자료에서 $Y_t < LCL_x$ 혹은 $Y_t > UCL_x$ 이면 공정의 주의를 준다.

3.2 EWMA 관리도

EWMA 관리도를 적용하기 위한 기본통계량으로 다음을 계산할 수 있다.

$$Z_t = \alpha y_t + (1 - \alpha)Z_{t-1}, \quad 0 < \alpha < 1 \tag{11}$$

여기에서 α 는 가중치(weighting factor) 이다. Z_t 의 평균과 분산은 다음과 같이 계산할 수 있다.

$$E(Z_t) = \mu, \quad V(Z_t) = \frac{\alpha}{(2 - \alpha)} \cdot \sigma^2$$

EWMA 관리도에서는 다음과 같은 관리한계선을 이용한다.

$$LCL_{EWMA} = E(Z_t) - 3\sqrt{\frac{\alpha}{(2 - \alpha)}} \sigma \tag{12}$$

$$CL_{EWMA} = E(Z_t) \tag{13}$$

$$UCL_{EWMA} = E(Z_t) + 3\sqrt{\frac{\alpha}{2-\alpha}} \sigma \quad (14)$$

여기에서 표준편차는 $\sigma = (UCL_X - LCL_X)/6$ 을 이용하여 계산한다. t 시점에서 변환된 자료의 지수 가중 이동 통계량을 참고로 $Z_t < LCL_{EWMA}$ 혹은 $Z_t > UCL_{EWMA}$ 이면 공정의 주의를 준다. 가중치(α)를 정하는 논리적인 이론은 없으며, 전통적인 EWMA 관리도에서는 경험적으로 공정의 변화를 감지하기 위해 $0.20 \leq \alpha \leq 0.35$ 에서 정하여 준다.

4. 적용

지수분포 과정을 따르는 확률수를 통해 관측치를 구하여 지수변환을 한 후 기본 통계량을 구한다. 이때 자료변환을 위한 지수(θ)는 0.24, 0.25, 0.2654, 0.2777, 0.2907, 0.3333으로 설정한다. X 관리도를 적용하기 위한 관리한계선은 (8)식과 (10)식을 참고하며, EWMA 관리도를 적용하기 위한 관리한계선은 (12)식과 (14)식을 이용한다. 이때 가중치 α 는 0.1:0.45(0.025)로 변화시켜 본다. 두 관리도의 효율성을 평가하기 위해 관리한계선을 벗어나는 관측치 수(ONEC)를 이용한다. 이때 IMSL의 GGEXN을 이용하여 확률수를 구하며, 신뢰성을 고려하여 10,000 번의 반복실험을 한다.

<표2>는 X 관리도를 적용할 때 지수의 크기에 따라 ONEC 변화를 구한 것이다. 여기에서 Taylor가 제안한 0.2654일 때 공정의 주의를 주는 것에 매우 민감함을 알 수 있으며, Rudolf가 제안한 0.25 지수가 다른 학자들이 주장한 것에 비해 타당하다는 것을 알 수 있다.

<표2> X 관리도에서 지수변화에 따라 ONEC 변화

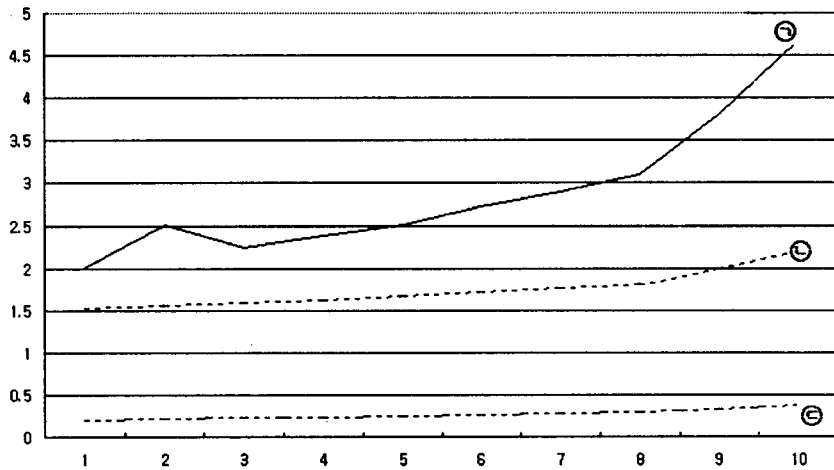
구분	exponent (θ)					
	0.2400	0.2500	0.2654	0.2777	0.2907	0.3333
ONEC	25	33	36*	21	27	26

<표3>은 EWMA 관리도를 적용할 때 지수(θ)의 크기와 가중치(α)의 변화에 따라 ONEC의 변화를 구한 것이다.

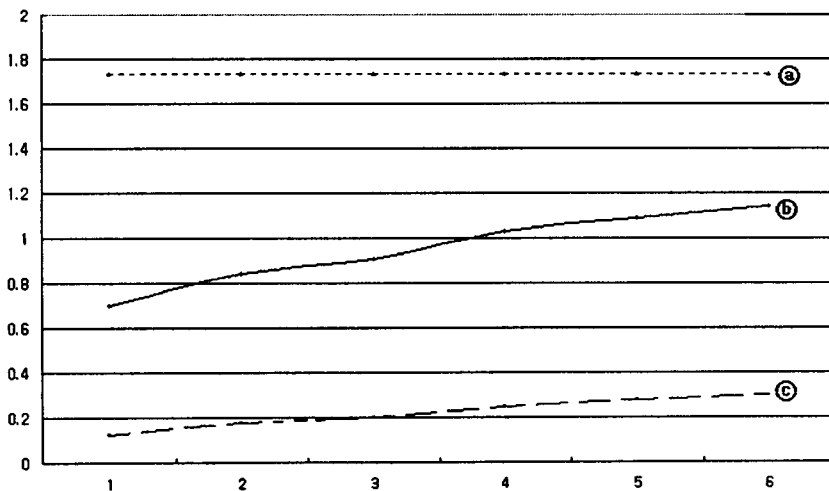
<표3> EWMA 관리도에서 지수변화와 가중치에 따라 ONEC 변화

구분	α	exponent (θ)					
		0.2400	0.2500	0.2654	0.2777	0.2907	0.3333
ONEC	0.125	49	146*	58	41	50	57
		1	2	3	1	3	2
	0.15	65	56	59	73	52	82*
		0	0	1	0	0	1
	0.20	75	42	47	47	102*	72
		0	0	0	0	0	0
	0.25	38	50	45	70*	62	48
	0.30	42	50	55	60	61*	59
0.35	62	46	37	56	63*	49	
0.40	41	34	38	47	61	64*	

<표3>에서 진하게 표시된 부분은 기존의 방법으로 관리한계선을 이용하여 ONEC를 계산한 것으로서 표준편차의 영향으로 관리한계의 폭이 넓어져서 민감도가 떨어짐을 확인할 수 있다. 또한 가중치가 0.20 이상이 되면 관리도에서 신호를 감지할 수 없음을 확인할 수 있다. <표2>와 <표3>을 비교하면 EWMA 관리도를 이용하는 것이 X 관리도를 이용하는 것보다 민감하다는 것을 알 수 있다. Rudolf가 제안한 0.25 지수를 사용할 때는 가중치를 작게 사용하는 것이 바람직하지만 지수를 크게 할 경우 가중치도 크게 하는 것이 관리도의 민감도를 높일 수 있다. 기존의 EWMA 관리도에서는 가중치를 $0.20 \leq \alpha \leq 0.35$ 내외에서 주지만 지수변환이 된 자료를 가지고 적용할 때에는 $0.125 \leq \alpha \leq 0.25$ 내외로 주는 것이 효과적임을 확인할 수 있다. <그림1>은 지수(θ :㉔)를 0.21:0.3777으로 변화시킬 때 변환자료의 평균 $E(Y_t)$ (㉓), 관리한계 폭(㉒)의 변화를 나타낸 그림이다. 여기에서 지수를 크게 할수록 변환된 자료의 변화에 비해 관리한계의 폭이 급격히 증가함을 알 수 있다. 따라서 지수변환에 의해 X 관리도를 적용할 때 지수를 크게 하는 것보다 작게 하는 것이 효율적임을 알 수 있다. 즉, Rudolf가 제안한 0.25 내외에서 적용하는 것이 바람직함을 알 수 있다.



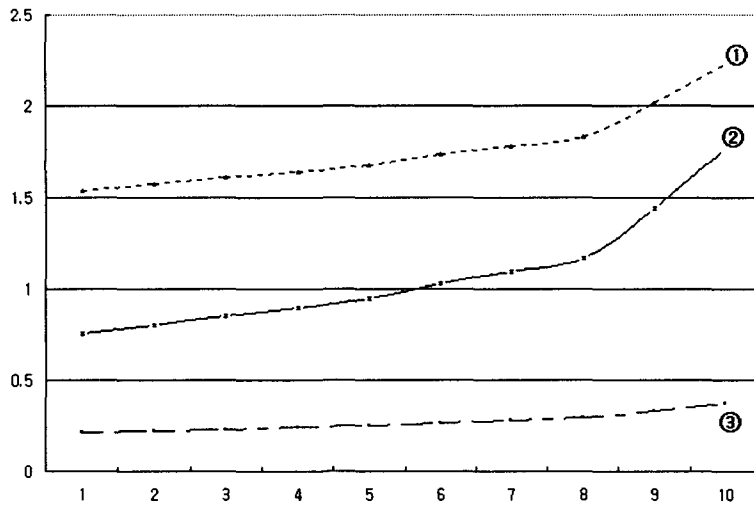
<그림1> 지수(θ :㉔), $E(Y_t)$ (㉓), 관리한계 폭(㉒)의 변화



<그림2> 지수(θ)=0.2654일 때 가중치(α :㉓), 관리한계 폭(㉒), $E(Z_t)$ (㉔)의 변화

<그림2>는 지수(θ)=0.2654일 경우 가중치(α)를 0.125:0.3으로 변화시킴에 따라 관리한계 폭(⑥)과 EWMA 통계량의 평균값의 $E(Z_t)$ (②)의 변화를 나타낸 그림이다. 여기에서 가중치를 크게 할수록 관리한계의 폭이 증가함을 알 수 있으며 EWMA 통계량의 변화에는 영향이 거의 없음을 알 수 있다. 따라서 가중치를 크게 하는 것보다 작게 하는 것이 효율적임을 알 수 있다.

<그림3>은 가중치(α)를 0.25로 할 경우 지수(③)를 0.21:0.3777까지 변화시킴에 따라 관리한계 폭(②)과 EWMA 통계량의 평균값 $E(Z_t)$ (①)의 변화를 나타낸 그림이다. 여기에서 지수를 크게 할수록 관리한계의 폭과 EWMA 통계량이 증가함을 알 수 있다.



<그림3> 가중치(α)를 0.25로 할 때 지수(θ :③), 관리한계 폭(②), $E(Z_t)$ (①)의 변화

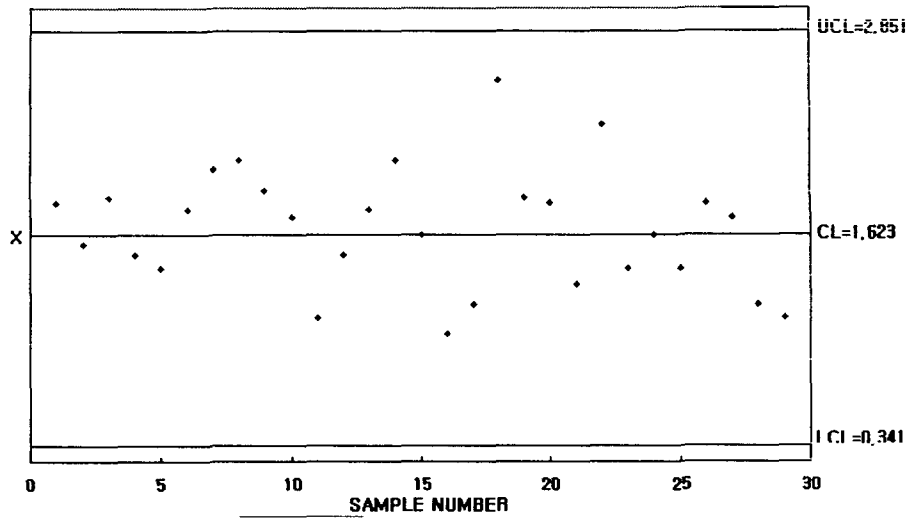
[적용 예]

생산공정에서 부정기적으로 조사된 자료를 이용하여 공정관리를 하기 위해 C 관리도를 이용한 다. 그러나 개별 관측치를 이용할 경우 C 관리도를 이용하여 이상신호를 감지하기란 쉬운 일 이 아니다. 여기에서 조사된 기간내에 29개의 이상신호 점의 자료가 있으며, 이상신호가 발생 한 간격이 다음과 같다. 이들 자료는 1 단위 시간부터 45 단위시간까지의 범위에서 전형적인 지수분포를 한다.

11	6	12	5	4	10	17	19	13	9
1.5	5	10	19	7	1	2	45	12	11
3	29	4	7	4	11	9	2	1.5	

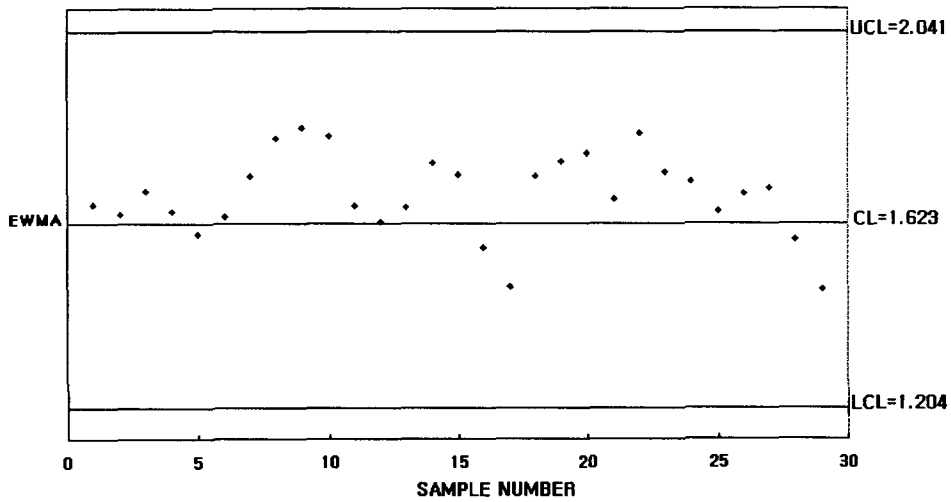
자료의 빈도수를 가지고 그래프를 그리면 근사적인 지수분포 형태를 나타냄을 알 수 있다. 자료의 평균은 10.3035, 표준편차는 9.3068, 중앙값은 9이다. $\theta=0.25$ 일 경우 변환된 자료를 기준으로 평균은 1.6764, 표준편차는 0.3739, 중앙값은 1.7321이다.

<그림4>는 지수 변환된 자료를 이용하여 X 관리도를 적용한 결과이다. 여기에서 관리한계선을 구하기 위해 (8), (9), (10) 식을 이용한다. 중심선(CL)을 기준으로 관리상한선의 폭이 관리하한선의 폭보다 좁음을 확인할 수 있다.



<그림4> $\theta=0.25$ 일 때 변환된 자료를 기준으로 적용한 개별관리도

<그림5>는 $\theta=0.25, \alpha=0.20$ 일 때 변환된 자료를 기준으로 적용한 EWMA 관리도이다. 개별관리도의 관리한계 폭에 비해 EWMA 관리도 관리한계의 폭이 29% 정도 작다는 것을 확인할 수 있다. 관리한계를 구하기 위해 (12), (13), (14) 식을 이용한다.



<그림5> $\theta=0.25, \alpha=0.20$ 일 때 변환된 자료를 기준으로 적용한 EWMA 관리도

<그림4>와 <그림5>에서 어떠한 신호를 감지하지는 못하였지만 지수분포 과정을 따르는 자료들을 변환하여 공정의 개선을 위한 이상신호의 감지 및 주의 조치를 취할 수 있는 방법을 단순화하였다.

5. 결론 및 향후전망

C 관리도의 보안을 위해 지수를 적용한 변환자료를 이용할 수 있다. 변환자료의 이상신호를 관리하기 위해 X 관리도보다 EWMA 관리도가 공정의 주위를 감지하는 면에서 효율적임을 확

인할 수 있다. 지수변환에 의해 X 관리도를 적용할 때 지수를 크게 하면 관리한계의 폭이 넓어져 민감도가 떨어지므로 지수를 작게 하는 것이 효율적임을 알 수 있다. 공정에 이상신호가 발생한 간격을 참고로 지수분포과정을 따르는 자료들을 변환한 후 EWMA 관리기법을 이용하는 것이 이상신호를 감지하는 민감도를 높일 수 있다. EWMA 관리도를 적용할 때 지수와 가중치를 크게 할수록 관리한계의 폭이 지나치게 넓어지므로 지수변환을 위한 지수는 0.25 내외에서 적용하는 것이 바람직하다. 또한 가중치는 $0.125 \leq \alpha \leq 0.2$ 에서 정하는 것이 효율적임을 제시한다.

공정분산을 기준으로 지수변환을 이용한 EWMS 관리기법의 타당성에 대해서는 더 많은 연구가 진행되어야 할 것이다.

참고문헌

- [1] Adke S.R. and Hong X., "A Supplementary Test Based on the Control Chart for Individuals", *Journal of Quality Technology*, Vol 29, No 1, pp 16-20, 1997.
- [2] Borror C. M. and Montgomery, D. C. and Runger G. C., "Robustness of the EWMA control chart to non-normality", *Journal of Quality Technology*, Vol 31, pp 309-316, 1999.
- [3] Draper N. R. and Cox D. R., "On distribution and their transformation to normality", *J.R.S.S.*, B 31, pp 472-476, 1969.
- [4] Felisa J. Vazquez, "Generation of random variables-exponential distribution", <http://www.ee.mu.oz.au>, 1999.
- [5] IMSL inc, "The I.M.S.L. library", 1980.
- [6] Nelson L. S., "Control charts for individual measurements", *Journal of Quality Technology*, Vol 14, No 3, pp 172-173, 1982.
- [7] Nelson L. S., "The shewhart control chart-test for special causes", *Journal of Quality Technology*, Vol 16, pp 237-239, 1984.
- [8] Nelson L. S., "A control charts for parts per million nonconforming items", *Journal of Quality Technology*, Vol 26, pp 239-240, 1994.
- [9] Rudolf G. K., Jr, "Transforming the Exponential for SPC Applications", *Journal of Quality Technology*, Vol 31, No 3, pp 301-308, 1999.
- [10] Susan L, Albin and Lan Kang and Gerald Shea, "An X and EWMA chart for individual observations", *Journal of Quality Technology*, Vol 29, No 1, pp 41-48, 1997.
- [11] 지선수의외, "통계적 품질관리", 경문사, 1997.