

Design of Robust \bar{x} Control Chart Using a Location Parameter

Young-Bae Chung[†] · Yon-Soo Kim

Dept. of Industrial and Management Engineering, Incheon National University

위치모수를 이용한 로버스트 \bar{x} 관리도의 설계

정영배[†] · 김연수

인천대학교 산업경영공학과

Control charts are generally used for process control, but the role of traditional control charts have been limited in case of a contaminated process. Traditional \bar{x} control charts have not been activated well for such a problem because of trying to control processes as center line and control limits changed by the contaminated value. This paper is to propose robust \bar{x} control charts which is considering a location parameter in order to respond to contaminated process. In this paper, we consider \bar{x}_α , that is trimmed rate; typically ten percent rate is used. By comparing with p, ARL value, the responding results are decided. The comparison resultant results of proposed two control charts are shown and are well contrasted.

Keywords : Robust Control Chart, Contaminated Process, Location Parameter

1. 서론

품질은 산포를 하기 때문에 제품의 품질을 향상시키는데 있어서는 항상 공정관리의 문제가 발생하게 된다.

이러한 품질의 산포를 관리하기 위해, 산포의 원인을 우연원인과 이상원인으로 구분하여 통제 가능한 이상원인을 찾아 현장조치를 통해 해결함으로써 공정을 관리상태로 유지하는 방법이 사용되고 있으며, 대표적으로 사용되는 도구로써 Shewhart에 의해 제창된 관리도(Control Chart)가 있다. Shewhart가 소개한 관리도에는 공정의 평균치를 관리하는 관리도와 산포를 관리하는 관리도가 있으며, 평균치 관리도는 군내, 군간 산포를, 산포 관리도는 군내 산포를 관리한다.

현재 가장 널리 사용되는 관리도는 $\bar{x}-R$ 관리도가 있다. 그러나 이러한 전통적인 관리도는 공정이 오염되어 있을 경우에는 이러한 오염된 데이터의 영향을 받아 그

한계가 드러날 수가 있다. 즉, 평균치는 이상치의 영향을 곧바로 받기 때문에 이러한 오염된 데이터가 있는 경우의 평균치는 중심을 나타내는 척도로서 문제점을 나타내게 된다. 이러한 이유로 관측치가 정상상태가 아닌 상황, 즉 공정에 오염된 관측치가 침투하였을 경우의 관리선은 오염된 관측치를 여과 없이 받아들이기 때문에 관리도가 해야 하는 본연의 기능인 이상군의 탐지 능력에 문제가 생긴다. 따라서 기존의 $\bar{x}-R$ 관리도에서의 평균추정치보다 오염된 공정변화에 덜 영향을 받을 수 있는 로버스트 평균치에 대한 추정치의 연구가 필요하게 되었고, 이러한 로버스트 평균치를 이용한 관리도를 공정관리에 도입함으로써 오염된 데이터가 있는 경우에도 효과적으로 공정관리를 할 수 있다는 것이다.

현재까지 기존 \bar{x}, R 과 같은 평균, 산포의 추정치보다 오염된 공정에서의 로버스트 추정치에 대한 연구가 다양하게 진행되어 왔다. 로버스트 평균 추정치의 연구로서 Rocke는 절사평균을, Tukey와 Wang은 표본의 중앙값과 다른 두 사분위수의 가중평균에 근거한 trimean이라는 평균 추정치를 개발하였다[4, 8, 9]. Wu는 중앙값으로부

터 절대편차의 중앙값, 중앙값으로부터 절대편차의 평균 등의 평균 추정치를 개발하였다[10].

Schoonhoven은 여러 학자들에 의해 주장된 평균, 표준편차의 추정치에 관해 각각의 효율성을 비교하였다[6].

본 논문에서는 Schoonhoven에 의해 증명되어진 여러 가지 평균치와 산포의 추정치 중 평균 제곱 오차의 변화가 적은 추정치 중 하나로 절사하는 방법을 채택하였으며, 절사하는 방법으로는 각 부분군의 평균과 산포를 모두 구한 후에 절사를 시행한 후 평균을 구하는 방법과 각각의 부분군에서 절사를 시행한 후 평균과 산포의 추정치를 구한 후 나중에 평균을 구하는 방법 두 가지를 사용하였다. 그리고 절사하는 이론을 채택하여 설계한 새로운 관리도가 기존의 관리도와 비교하여 어떠한 차이를 나타내는지 시뮬레이션을 통해 확인하였다.

본 논문은 위 두 가지 방법에 의하여 구해진 위치모수를 이용하여 평균치와 산포의 강건한 추정치를 추정하고, 새로운 관리선을 설정함으로써 공정에 오염된 데이터가 있을지라도 이에 영향을 받지 않고 공정을 관리할 수 있는 로버스트 \bar{x} 관리도를 설계하고자 한다.

2. 기존 연구 고찰

평균치 관리도는 공정의 중심이 변화되고 있는가를 감시하는 관리도로 평균이, 산포 관리도는 공정의 산포가 변화되고 있는가를 감시하는 관리도로 산포가 통계적 관리 상태를 벗어나고 있을 때 이를 검출할 수 있는 기준을 제시해준다.

2.1 \bar{x} -R 관리도

\bar{x} 관리도는 부분군에서 계산된 \bar{x} 를 타점하여 공정의 중심이 변화되고 있는가를 감시하는 관리도이고, 중심선인 CL 과 관리한계선인 UCL 과 LCL 은 다음 식 (1)과 같다.

$$\begin{aligned}
 CL &= E(\bar{x}) = \mu & (1) \\
 UCL &= E(\bar{x}) + 3D(\bar{x}) \\
 &= \mu + 3 \frac{\sigma}{\sqrt{n}} \\
 LCL &= E(\bar{x}) - 3D(\bar{x}) \\
 &= \mu - 3 \frac{\sigma}{\sqrt{n}}
 \end{aligned}$$

표준값이 정해져 있지 않은 경우에 모수 추정치를 사용하여 다음 식 (2)와 같이 관리선(중심선, 관리한계선)을 구하게 된다.

$$CL = \hat{\mu} = \bar{\bar{x}} \tag{2}$$

$$\begin{aligned}
 UCL &= \hat{\mu} + 3 \frac{\hat{\sigma}}{\sqrt{n}} \\
 &= \bar{\bar{x}} + \frac{3}{\sqrt{n}} \bar{R}
 \end{aligned}$$

$$\begin{aligned}
 LCL &= \hat{\mu} - 3 \frac{\hat{\sigma}}{\sqrt{n}} \\
 &= \bar{\bar{x}} - \frac{3}{\sqrt{n}} \bar{R}
 \end{aligned}$$

여기서 $\bar{\bar{x}} = \sum \bar{x}_i/k$, $\bar{R} = \sum R_i/k$ 로 계산된다.

R 관리도는 부분군에서 계산된 R 을 타점한 후, 공정의 산포가 변화되고 있는가를 감시하는 관리도이고, 중심선인 CL 과 관리한계선인 UCL 과 LCL 은 다음 식 (3)과 같다.

$$CL = E(R) = d_2\sigma \tag{3}$$

$$\begin{aligned}
 UCL &= E(R) + 3D(R) \\
 &= d_2\sigma + 3d_3\sigma
 \end{aligned}$$

$$\begin{aligned}
 LCL &= E(R) - 3D(R) \\
 &= d_2\sigma - 3d_3\sigma
 \end{aligned}$$

표준값이 정해져 있지 않은 경우는 모수 추정치를 사용하여 다음 식 (4)와 같이 관리선(중심선, 관리한계선)을 구하게 된다.

$$CL = d_2\hat{\sigma} = d_2 \frac{\bar{R}}{d_2} = \bar{R} \tag{4}$$

$$\begin{aligned}
 UCL &= (d_2 + 3d_3)\hat{\sigma} \\
 &= (d_2 + 3d_3) \frac{\bar{R}}{d_2} \\
 &= (1 + 3 \frac{d_3}{d_2}) \bar{R} \\
 &= D_4 \bar{R}
 \end{aligned}$$

$$\begin{aligned}
 LCL &= (d_2 - 3d_3)\hat{\sigma} \\
 &= (d_2 - 3d_3) \frac{\bar{R}}{d_2} \\
 &= (1 - 3 \frac{d_3}{d_2}) \bar{R} \\
 &= D_3 \bar{R}
 \end{aligned}$$

여기서 $\bar{R} = \sum R_i/k$,이며 $D_4 = 1 + 3 \frac{d_3}{d_2}$, $D_3 = 1 - 3 \frac{d_3}{d_2}$ 는 역시 관리계수로서 n 에 의하여 정의되는 계수값이다.

2.2 로버스트 추정치

Rocke는 자신의 논문에서 식 (5)~식 (7)과 같이 세 가지 평균치의 로버스트 추정치를 제시하였다[2].

식 (5)는 측정치들을 크기순으로 나열한 뒤, 측정치들의 중앙값을 구하는 방법이다.

$$M(\bar{X}) = \text{median}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k) \quad (5)$$

식 (6)은 식 (5)의 추정치의 평균을 내는 방법이다.

$$\bar{M} = \frac{1}{k} \sum_{i=1}^k M_i \quad (6)$$

식 (7)은 측정치를 $\alpha\%$ 의 비율로 큰 값과 작은 값 양쪽에서 절사하여 평균 \bar{X} 의 기대값을 구하는 방법이다.

$$\bar{X}_\alpha = \frac{1}{k - 2 \lceil k\alpha \rceil} \times \left[\sum_{v=\lceil k\alpha \rceil+1}^{k-\lceil k\alpha \rceil} \bar{X}_{(v)} \right] \quad (7)$$

이 추정치는 여기서 α 는 잘라낼 부분군의 백분율(절사율)이며, $\lceil x \rceil$ 은 천정함수를 나타내는 것으로 x 보다 작지 않은 정수들 중 가장 작은 것을 의미한다. k 는 부분군의 총수, $\bar{X}_{(v)}$ 는 v 번째 군에서의 \bar{X} 를 의미한다.

Tukey는 중앙값(median)과 hinge들의 결합을 통해 중심 위치를 측정하는 방법을 제시하였다[6]. hinge는 데이터를 크기순으로 배열 하였을 때 중앙값을 중심으로 양 옆으로 배열되는 데이터들을 말하며, 이때 trimean은 중앙값에 가중치를 주고 다른 두 hinge들과 합의 평균을 구하는 방식으로 산출식은 다음 식 (8)과 같다.

$$\text{trimean} = \frac{\text{lower hinge} + 2(\text{median}) + \text{upper hinge}}{4} \quad (8)$$

여기서 lower hinge, upper hinge는 크기순으로 배열된 데이터들 중에서 각각 1/4지점(Q_1), 3/4지점(Q_3)의 데이터를 의미한다.

3. 로버스트 \bar{x} 관리도의 설계

로버스트 \bar{x} 관리도는 각각의 부분군에서 절사율 α 를 적용하여 절사를 시행한 후 구한 \bar{x}_α 값의 평균을 $\hat{\mu}_x$ 의 추정치로 사용하는 관리도로서 \bar{x}_α 관리도라 부른다. 여기서 절사율 α 는 $0 < \alpha < 0.5$ 의 값을 갖는다.

중심선(Center Line)인 CL 과 관리한계선(Control Limit)인 UCL 과 LCL 은 다음 식 (9)와 같다.

$$CL = \hat{\mu}_x = \bar{x}_\alpha \quad (9)$$

$$UCL = \hat{\mu}_x + 3 \frac{\hat{\sigma}_x}{\sqrt{n}} \\ = \bar{x}_\alpha + \frac{3}{\sqrt{n}} \times \frac{\bar{R}_\alpha}{d_2}$$

$$LCL = \hat{\mu}_x - 3 \frac{\hat{\sigma}_x}{\sqrt{n}} \\ = \bar{x}_\alpha - \frac{3}{\sqrt{n}} \times \frac{\bar{R}_\alpha}{d_2}$$

단, $\bar{x}_\alpha = \frac{1}{n - 2 \lceil n\alpha \rceil} \times \left[\sum_{v=\lceil n\alpha \rceil+1}^{n-\lceil n\alpha \rceil} x_{i(v)} \right]$ 이고, α 는 잘라낼 부분군의 절사율, $\lceil x \rceil$ 는 천정함수를 나타내는 것으로 x 보다 작지 않은 정수들 중 가장 작은 것을 의미한다. n 은 각 군의 측정치 수, $x_{i(v)}$ 는 i 군에서의 v 번째의 값을 의미한다.

R_α 관리도는 부분군에서 계산된 R 들을 타점하고, 공정 산포가 변화되고 있는가를 감시하는 관리도로 각각의 부분군에서 절사율 α 를 적용하여 절사를 시행한 후 구한 R_α 의 평균을 $\hat{\mu}_R$ 의 추정치로 사용하는 관리도이다. 중심선(Center Line)인 CL 과 관리한계선(Control Limit)인 UCL 과 LCL 은 다음 식 (10)과 같다.

$$CL = d_2 \hat{\sigma} \\ = d_2 \frac{\bar{R}_\alpha}{d_2} \\ = \bar{R}_\alpha \quad (10)$$

$$UCL = (d_2 + 3d_3) \hat{\sigma} \\ = (d_2 + 3d_3) \frac{\bar{R}_\alpha}{d_2} \\ = (1 + 3 \frac{d_3}{d_2}) \bar{R}_\alpha \\ = D_4 \bar{R}_\alpha$$

$$LCL = (d_2 - 3d_3) \hat{\sigma} \\ = (d_2 - 3d_3) \frac{\bar{R}_\alpha}{d_2} \\ = (1 - 3 \frac{d_3}{d_2}) \bar{R}_\alpha \\ = D_3 \bar{R}_\alpha$$

단, $\bar{R}_\alpha = \frac{1}{n - 2 \lceil n\alpha \rceil} \times \left[\sum_{v=\lceil n\alpha \rceil+1}^{n-\lceil n\alpha \rceil} R_{i(v)} \right]$ 이고, $D_4 = 1 + 3 \frac{d_3}{d_2}$, $D_3 = 1 - 3 \frac{d_3}{d_2}$ 은 관리계수로서 n 에 의하여 정의되는 계수값이다. 또한 $n \leq 5$ 인 경우는 D_3 의 값은 음의 값이 나타난다. 따라서 관리도를 그릴 때 \bar{R}_α 관리도의 관리하한선은 의미가 없으므로 표기하지 않는다. α 는 잘

라벨 부분군의 절사율이며, $[x]$ 은 천정함수를 나타내는 것으로 x 보다 작지 않은 정수들 중 가장 작은 것을 의미한다. n 은 각 군의 측정치 수, $R_{i(v)}$ 는 i 군에서의 v 번째 값을 의미한다.

관리도의 성능은 관리한계선을 벗어날 확률인 p 와 타점된 관측치가 처음으로 관리한계를 벗어날 때까지의 평균 런의 길이(ARL : Average Run Length)에 의해 평가한다. E_i 는 관리한계선 밖에 있는 사건, $P(E_i)$ 는 그 확률이라고 정의하면, 첫 번째로 관리한계선을 벗어나는 평균 \bar{x}_i , 범위 R_i 가 나오기 전까지의 표본의 수를 런의 길이(RL : Run Length)이라고 표현한다.

σ 가 알려져 있는 경우 일반적으로 RL은 $p = P(E_i) = \alpha$ 로 기하분포를 따른다. 평균 런의 길이(ARL)는 $1/p$ 로 런 길이의 표준편차(SDRL)은 $\sqrt{1-p}/p$ 의 확률을 따른다. σ 가 알려져 있지 않은 경우는 $p = E(P(E_i|\hat{\mu}, \hat{\sigma}))$ 와 $ARL = E(1/P(E_i|\hat{\mu}, \hat{\sigma}))$ 로 정의한다. 이러한 기대값들은 충분한 반복을 통한 데이터 세트를 발생시킴으로 구해질 수 있다.

4. 실험 및 분석

\bar{x}_α 관리도의 성능은 관리한계선을 벗어날 확률인 p 와 타점된 관측치가 처음으로 관리한계를 벗어날 때까지의 평균 런의 길이(ARL : Average Run Length)에 의해 평가한다. $\hat{\mu} = \bar{x}_\alpha$, $\hat{\sigma} = \bar{R}_\alpha/d_2$ 를 사용하고, 절사율 α 는 가장 대표적으로 사용되는 10%로 설정하여 시뮬레이션을 하였다. 평가의 비교대상으로 기존의 \bar{x} 관리도를 삼았으며, 공정의 변화를 첫째, 공정의 평균이 변화하는 경우, 둘째, 공정의 산포가 변화하는 경우, 셋째, 공정의 평균과 산포가 변화하는 경우로 나누어 각각 비교하였다.

4.1 공정변화 설계

공정의 변화 설계는 정상적인 경우에 공정에서 발생할 수 있는 평균과 산포가 변화한 오염된 비정상적인 공정이 혼입된 다음 세 가지 경우로 시뮬레이션을 실시한다.

- (1) 정상적인 공정 $N(0,1^2)$ 90%에 평균에 변화를 준 오염된 공정 $N(1.5,1^2)$ 10%의 비율로 혼입된 경우
- (2) 정상적인 공정 $N(0,1^2)$ 90%에 산포에 변화를 준 오염된 공정 $N(0,2.5^2)$ 10%의 비율로 혼입된 경우
- (3) 정상적인 공정 $N(0,1^2)$ 90%에 평균과 산포를 동시에

변화를 준 오염된 공정 $N(1.5,2.5^2)$ 10%의 비율로 혼입된 경우

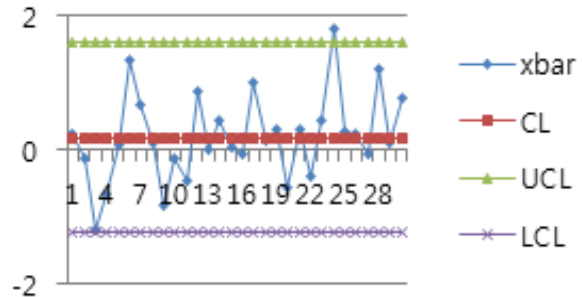
$n = 5, k = 30$ 을 기준으로 시뮬레이션을 실시하였고, 시뮬레이션 횟수 N 은 각 case 별 50회를 실시하였다.

4.2 수치 예

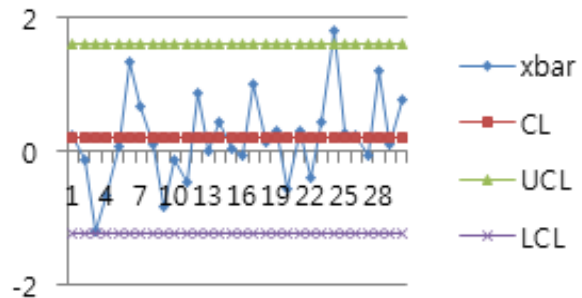
위에서 설계한 정상적인 공정에 오염된 공정 혼입의 세 가지 경우에 대해서 기존의 \bar{x} 관리도와 제안한 로버스트 $\bar{x}_{0.1}$ 관리도로서의 $\bar{x}_{0.1}$ 관리도와의 성능 비교는 <Table 1>, <Table 2>, <Table 3>에, 관리도는 <Figure 1>, <Figure 2>, <Figure 3>에 나타냈다.

<Table 1> Comparison p and ARL in $N(0,1^2)$ and $N(1.5,1^2)$

	p	ARL
\bar{x}	0.0811	12.326
$\bar{x}_{0.1}$	0.0814	12.378



(a) \bar{x} control chart

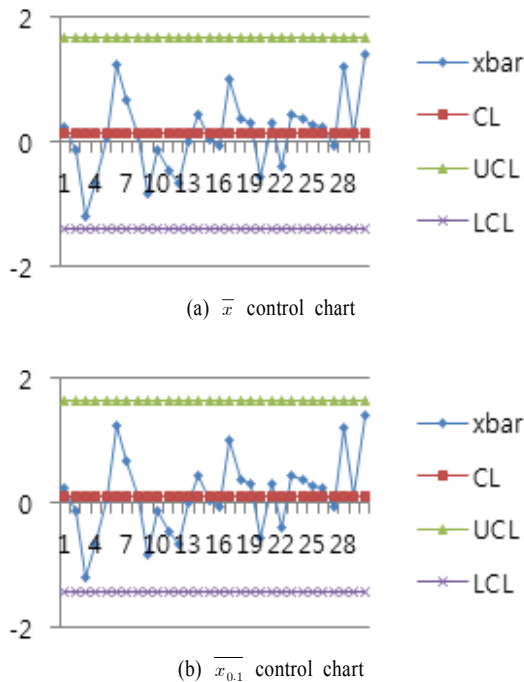


(b) $\bar{x}_{0.1}$ control chart

<Figure 1> \bar{x} and $\bar{x}_{0.1}$ Control Chart in $N(0,1^2)$ and $N(1.5,1^2)$

<Table 2> Comparison p and ARL in $N(0,1^2)$ and $N(0,2.5^2)$

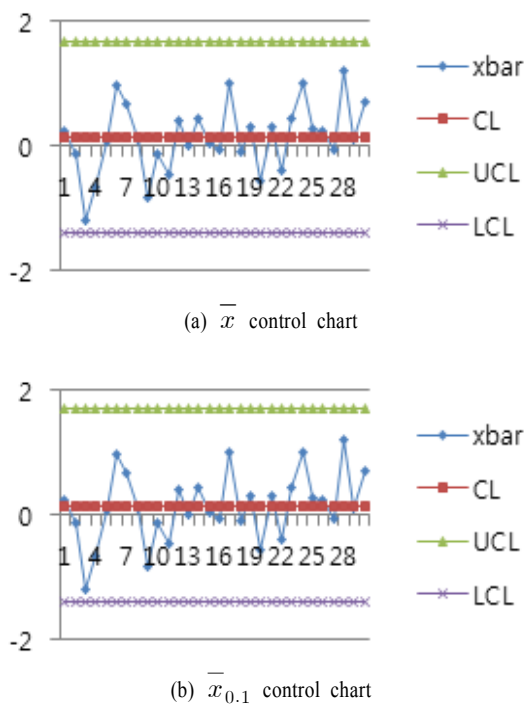
	p	ARL
\bar{x}	0.0545	18.362
$\bar{x}_{0.1}$	0.0547	18.269



<Figure 2> \bar{x} and $\bar{x}_{0.1}$ Control Chart in $N(0,1^2)$ and $N(0,2.5^2)$

<Table 3> Comparison p and ARL in $N(0,1^2)$ and $N(1.5,2.5^2)$

	p	ARL
\bar{x}	0.0423	23.661
$\bar{x}_{0.1}$	0.0424	23.595



<Figure 3> \bar{x} and $\bar{x}_{0.1}$ Control Chart in $N(0,1^2)$ and $N(1.5, 2.5^2)$

본 연구에서 제안한 로버스트 관리도와 기존의 공정 관리를 위해 널리 이용되고 있는 전통적인 관리도의 성능 비교를 위하여 p와 ARL값을 구했다. 공정이 오염된 세 가지 경우에서 관리도의 객관적인 비교를 위해 본 논문에서는 관측치를 동일하게 조절된 상태에서 p값과 ARL값을 비교하였다. 일반적으로 데이터가 오염되어 있는 경우 p값이 높고, ARL값이 작은 관리도 방법이 공정 관리에 더욱 우수한 방법이라고 할 수 있다. 앞서 제안한 공정변화 설계의 3가지 경우 모두 로버스트 관리도가 우수함을 알 수 있다.

5. 결론

일반적으로 기업에서 공정관리를 함에 있어서 작업자의 숙련도 차이, 부적합한 자재의 유입 등 제어하지 못하는 환경인자나 초기공정 경우 등으로 인해 공정의 오염을 겪게 될 수 있다. 이러한 상황에서 오염된 관측치를 그대로 사용하여 기존의 전통적인 관리도로서 공정을 관리하게 되면 관리한계선이 오염된 관측치 값의 영향을 받게 되고, 이와 같은 오염된 관측치의 영향을 받은 관리한계선으로 공정관리를 하게 됨으로써 오염된 군을 판별하지 못하는 경우가 발생할 수 있다. 이러한 이유로 인하여 실제 공정에서 공정을 제대로 관리하지 못하여 발생하는 물질적, 시간적인 불필요한 손실을 야기함으로써 기업에 손해를 끼칠 수 있다.

본 논문에서는 이러한 상황에서 관측치에 이상값이 침투한 오염된 공정을 고려한 강건한 관리도로서 기존의 관리도에 비해 민감하게 대응 할 수 있는 절사 모수를 적용하는 관리도법을 제시하였고, 이러한 절사 모수를 적용한 로버스트 관리도가 오염된 공정에 반응하지 않고 기존 관리도보다 성능이 좋은 관리도임을 보여 주었다.

본 논문에서 제안한 절사모수를 적용한 관리도 기법은 강건한 공정관리를 수행하는데 도움을 주어 기업의 불필요한 손실을 줄여줄 수 있을 것이라 기대된다.

Acknowledgement

This work was supported by Incheon National University research grant in 2012.

References

[1] Jensen, W.A., Jones-Farmer, L.A., Champ, C.W., and Woodall, W.H., Effects of Parameter Estimation on Control Chart Properties : A Literature Review. *Journal*

- of *Quality Technology*, 2006, Vol. 38, No. 4, p 349-364.
- [2] Chung, Y.B. and Kim, Y.S., Comparison and Application of Process Capability Indices. *Journal of Society of Korean Industrial and System Engineering*, 2007, Vol. 30, No. 4, p 182-189.
- [3] Kim, D.H. and Chung, Y.B., Design of Expected Loss Control Chart considering Economic Loss. *Journal of Society of Korean Industrial and System Engineering*, 2013, Vol. 36, No. 2, p 56-62.
- [4] Rocke, D.M., Robust Control Charts. *Technometrics*, 1989, Vol.31, No. 2, p 173-184.
- [5] Schoonhoven, M., Riaz, M., and Does, R.J.M.M., Design and Analysis of Control Charts for Standard Deviation with Estimated Parameters. *Journal of Quality Technology*, 2011, Vol. 43, No. 4, p 307-333.
- [6] Schoonhoven, M. and Does, R.j.M.M., A Robust Standard Deviation Control Chart. *Technometrics*, 2012, Vol. 54, No. 1, p 73-82.
- [7] Tatum, L.G., Robust Estimation of the Process Standard Deviation for Control Charts. *Technometrics*, 1997, Vol. 39, No. 2, p 127-141.
- [8] Tukey, J.W. Exploratory Data Analysis. Reading, MA : Addison-Wesley; 1977.
- [9] Wang, T., Li, Y., and Cui, H., On Weighted Randomly Trimmed Means. *Journal of Systems Science and Complexity*, 2007, Vol. 31, No. 4, p 376-386.
- [10] Wu, C., Zhao, Y., and Wang, Z., The Median Absolute Deviation and Their Applications to Shewhart \bar{X} Control Charts. *Communications in Statistics-Simulation and Computation*, 2002, Vol. 31, No. 3, p 425-442.