

비정규성 데이터에 대한 단일 관리도들의 비교

강명구¹, 이장택²

^{1,2}단국대학교 응용통계학과

접수 2015년 3월 23일, 수정 2015년 5월 11일, 게재확정 2015년 5월 24일

요약

품질특성치의 중심과 산포를 하나의 통계량으로 관리하는 단일 관리도는 품질특성치가 정규분포를 따른다고 가정하지만 실제 데이터들은 왜도가 양수이거나 첨도가 양수인 경우가 많다. 본 논문에서는 품질특성치가 정규분포를 따르지 않은 경우에 가짜 알람률 (false alarm rate; FAR)을 이용하여 단일 관리도 성능을 비교하였다. 고려된 단일 관리도는 반원관리도, 최대 관리도 및 평균제곱오차 관리도이며 모의실험 결과, 공정이 안정 상태인 경우는 최대관리도의 성능이 좋았으며, 공정이 불안정 상태인 경우에는 왜도가 양수일 때 최대관리도, 첨도가 큰 경우에는 평균제곱오차 관리도의 성능이 우수하였다.

주요용어: 가짜 알람률, 단일 관리도, 반원 관리도, 최대 관리도, 평균제곱오차 관리도.

1. 서론

제조현장에서 공정의 이상 유무를 탐지하는 데 사용되는 관리도 (control chart)는 Bell 전화연구소에 근무하던 Shewhart (1926)에 의하여 처음으로 소개되었으며, 그의 저서는 오늘날의 관리도에 대한 기초적인 이론을 제시하여 주었다. 제조현장에서 제품의 강도, 무게 등 연속형 수치를 관리하는 관리도는 1930년대 Shewhart가 관리도를 소개한 이후로 지속적으로 사용되고 있지만 컴퓨터의 발달에 힘입어 산업 분야가 다양해지며 첨단화되는 연유로 고객의 요구사항은 점점 더 복잡하고 까다로워지기 때문에 보다 쉽고 편리한 관리도의 필요성이 한층 더 강조되고 있다. 이러한 관점에서 1990년대 후반부터 Shewhart 관리도보다 계산이 편하고 직관적으로 이해할 수 있는 단일 관리도 (single chart)에 대한 연구가 본격화 되었다. Chao와 Cheng (1996)은 실제 공정에서의 품질특성치 (quality characteristic)에 대한 중심과 산포를 동시에 알 수 있는 반원 관리도 (semicircle control chart)를 발표하였고, Chen과 Cheng (1998)은 공정의 순서를 잃어버리는 반원 관리도의 단점을 개선한 최대 관리도 (max chart)를 발표하였다. 또한 Spiring과 Cheng (1998)은 한 개의 단일 관리도에서 2개의 변수를 타점하여 품질특성치의 중심과 산포를 관리하는 평균제곱오차 관리도 (MSE chart)를 개발하였다.

한편 관리도에 대한 최근의 국내연구들을 살펴보면 누적이동평균모형에서 공정 변화시점의 추정문제를 다룬 Lee와 Lee (2009), 가변 추출간격을 갖는 다변량 Shewhart 관리도를 연구한 Cho (2010), 통합 공정관리에서 출력변수와 입력변수를 탐지하는 절차를 비교한 Lee (2011), 상관관계가 존재하는 다수의 품질 특성치들을 관리하기 위해 다변량 공정관리를 이용한 재수정 절차를 제안한 Cho와 Park (2011),

¹ (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 석사.

² 교신저자: (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 응용통계학과, 교수.
E-mail: jtlee@dankook.ac.kr

다변량 통합공정관리에서 재수정 절차 및 모수의 추정을 연구한 Cho와 Park (2013), 우도원리를 사용한 평균과 분산에 대한 결합관리도에 대한 연구한 Park (2013) 등이 있다.

Shewhart 관리도와 이를 기반으로 한 단일 관리도는 품질특성치가 정규분포를 따르거나 중심극한정리를 통한 정규분포의 근사가 가능할 수 있을 정도로 부분군 크기가 커야 함을 가정한다. 하지만 Levinson (2000)에 의하면 제조현장의 품질특성치가 정규분포를 따르지 않거나 시간 및 재무적 비용증가, 공정상의 문제로 인해 부분군의 크기를 충분히 크게 할 수 없는 경우가 존재하기 때문에 본 연구에서는 단일 관리도의 로버스트성 (robustness)을 비정규분포에서 공정이 안정 상태 (in-control)와 불안정 상태 (out-of-control)일 때 각각 비교하고자 한다. 본 논문은 다음과 같이 구성되어 있다. 2절에서는 단일 관리도의 종류에 대하여 기술하였으며, 3절에서는 모의실험을 위한 평가지표, 자료구성, 분석과정 및 4가지 방법들의 비교를 통해서 분석결과를 설명하며 끝으로 4절에서는 본 연구의 결론에 대해 언급하였다.

2. 단일 관리도의 종류

이번 절에서는 본 연구에서 사용할 단일 관리도에 대해 간단히 소개한다. 구체적인 논의에 앞서 사용할 표기는 X 는 품질특성치, X 의 모평균과 모분산은 각각 μ 와 σ^2 , i 번째 부분군의 j 번째 측정치는 X_{ij} (단, $i = 1, 2, \dots, k, j = 1, 2, \dots, n$), i 번째 부분군의 평균 \bar{X}_i 는 $\bar{X}_i = \sum_{j=1}^n X_{ij}/n$, i 번째 부분군의 표본분산 S_i^2 은 $S_i^2 = \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 / (n - 1)$ 으로 정의한다.

2.1. 반원관리도

Chao와 Cheng (1996)에 의해 제안된 반원 관리도 (semicircle control chart)는 식 (2.1)과 같이 정의되는 2차원 그래프를 사용한다.

$$T_i = (\bar{X}_i - \mu)^2 + S_i^{*2} \quad (2.1)$$

식 (2.1)에서 S_i^* 는 $S_i^* = S_i(\sqrt{(n-1)/n})$ 으로 정의되며 통계량 T_i 는 x 축이 \bar{X}_i , y 축이 S_i^* 인 평면에서의 원을 나타내며, S_i^* 는 양수이기 때문에 각 부분군 (\bar{X}_i, S_i^*) 를 그래프로 그리면 반원이 된다. 한편 $(n/\sigma^2) T_i$ 는 자유도가 부분군의 크기 n 인 카이제곱분포를 따른다. 따라서 통계량 T_i 가 반지름 r 의 제곱보다 작을 확률이 $1 - \alpha$ 가 되도록 식 (2.2)를 이용하여 r 을 계산할 수 있다.

$$P(T_i < r^2) = P\left(\frac{n}{\sigma^2} T_i < \frac{n}{\sigma^2} r^2\right) = 1 - \alpha \quad (2.2)$$

따라서 r 은 $r = \sigma \sqrt{\chi_{n, (1-\alpha)}^2 / n}$ 로 구할 수 있으며, 여기서 $\chi_{n, (1-\alpha)}^2$ 은 자유도가 n 인 카이제곱분포의 $100(1 - \alpha)$ 백분위수이다. 만일 μ 와 σ 를 모른다고 하면, 부분군의 개수 k 에 대하여 μ 는 $\bar{\bar{X}} = \sum_{i=1}^k \bar{X}_i / k$, σ 는 $\bar{S}^* = \sum_{i=1}^k S_i^* / k$ 를 이용하여 추정한다.

2.2. 최대관리도

Chen과 Cheng (1998)의 최대 관리도 (max chart)는 반원관리도처럼 중심과 산포를 나타낼 수 있는 통계량을 제시한다. 품질특성치 X 의 수집된 데이터가 부분군의 개수 k 와 부분군의 크기 n 에 대하여 평균이 $\mu + a\sigma$, 표준편차가 $b\sigma$ 인 정규분포를 따른다고 했을 때, 다음 식 (2.3)과 식 (2.4)를 도입한다. 식 (2.4)에서 $\Phi(z)$ 는 누적표준정규분포함수, Φ^{-1} 은 Φ 의 역함수, H 는 W 가 자유도 ν 인 카이제곱분포를

따를 때 $H(w; \nu) = P(W \leq w | \nu)$ 으로 정의된다.

$$U_i = \frac{(\bar{X}_i - \mu)}{\sigma/\sqrt{n}} \quad (2.3)$$

$$V_i = \Phi^{-1} \left\{ H \left(\frac{(n-1)S_i^2}{\sigma^2}; n-1 \right) \right\} \quad (2.4)$$

한편 공정평균과 산포를 동시에 모니터링 하기 위해 새로운 통계량 $M(n_i)$ 를 식 (2.5)와 같이 정의하면, $M(n_i)$ 을 이용하여 최대관리도의 중심선과 관리상한선을 제1종 오류 α 를 이용하여 결정할 수 있다 (Chen과 Cheng, 1998).

$$M(n_i) = \max \{|U_i|, |V_i|\} \quad (2.5)$$

2.3. 평균제곱오차 관리도

Spiring과 Cheng (1998)이 제안한 평균제곱오차 관리도 (MSE chart)는 목표값 (target value) t 에 대한 근접성 (proximity)과 그 근접성에 대한 변동을 관리하는 단일 관리도이다. 부분군의 개수 k 와 부분군의 크기 n 에 대하여 품질특성치 X 의 수집된 데이터가 정규분포를 따른다고 가정할 때, 관리하는 지표는 중심을 나타내는 식 (2.6)과 산포를 나타내는 식 (2.7)이다.

$$C_i = (\bar{X}_i - t)^2 \quad (2.6)$$

$$MSE_i = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - t)^2 \quad (2.7)$$

식 (2.6)과 식 (2.7)을 하나의 관리도에 그리는 것이 평균제곱오차 관리도이며 품질특성치 X 가 $X \sim N(\mu, \sigma^2)$ 일 때, C 는 $C \sim n^{-1}\sigma^2\chi_{1,\lambda}^2$ 이며, 또한 MSE 는 자유도가 n 이고 비중심모수가 $\lambda = n[(\mu - t)/\sigma]^2$ 인 $(n-1)^{-1}\sigma^2\chi_{n,\lambda}^2$ 분포를 따르는데, 이 경우 $\chi_{n,\lambda}^2$ 는 자유도가 n 이고 비중심모수가 λ 인 비중심 카이제곱분포를 나타낸다. 이를 통해 $100(1 - \alpha)$ 백분위수가 되는 C 와 MSE 의 관리상한선을 식 (2.8)과 같이 구할 수 있다.

$$UCL_C = n^{-1}\sigma^2\chi_{1,\lambda,(1-\alpha)}^2, \quad UCL_{MSE} = (n-1)^{-1}\sigma^2\chi_{n,\lambda,(1-\alpha)}^2 \quad (2.8)$$

3. 모의 분석

3.1. 평가지표

평균 런길이 (average run length; ARL)는 품질특성치의 관리도를 그렸을 때, 관리한계선 (control limit)을 이탈하기 전까지의 점들의 평균 개수를 의미하고, 가짜 알람률 (false alarm rate; FAR) p 는 평균 런길이의 역수이며 공정이 안정 상태지만 관리한계선을 이탈하는 비율을 나타낸다. 공정이 안정 상태일 때는 ARL이 길수록, 반대로 공정이 불안정한 상태일 때는 ARL이 짧은 것이 좋다고 평가하며 FAR은 그 반대이다. 일반적으로 ARL은 FAR p 의 역수이며, 3σ 의 \bar{X} 관리도의 경우 공정이 안정 상태일 때, 점이 관리한계선 밖으로 나갈 확률은 $p = 0.0027$ 이며 R 관리도의 경우는 $p = 0.0046$ 이다. 하지만 본 연구에서는 중심과 산포를 동시에 확인하는 것이기 때문에 관리한계선을 이탈할 확률 p 를 재계산해야 한다. 예를 들어 \bar{X} 관리도에서 관리한계선을 이탈할 사건을 A , R 관리도에서 관리한계선을 이탈할 사건을 B 라고 했을 때, ARL은 2개의 관리도 중에 먼저 관리한계선을 이탈할 확률의 역수이기 때문에 $1/P(A \cup B)$ 으로 나타낼 수 있다. 따라서 사건 A 와 B 가 서로 독립이라는 사실을 이용하면 FAR p 는 $P(A \cup B) \cong 0.0073$ 이며, 평균 런길이는 137이 된다.

3.2. 자료구성 및 분석과정

일반적으로 관리도는 공정이탈을 감지하기 위해 관리한계선의 하한 (lower control limit; LCL)과 관리한계선의 상한 (upper control limit; UCL)을 계산하는 단계 I (Phase I)과 계산된 관리한계선을 이용하여 공정의 이상 유무를 탐지하는 단계 II (Phase II)로 나뉜다. 2개의 단계를 다시 3개의 절차로 나누어 본 연구의 분석과정을 설명하면 [절차 1]은 단계 I의 해석용 관리도를 작성하는 부분으로서 μ 와 σ 는 미지라고 가정하고 부분군의 개수 $k = 100,000$, 부분군의 크기 $n = 5$ 의 표본을 확률분포에서 추출하여 추정치 $\hat{\mu}$, $\hat{\sigma}$, \widehat{UCL} , \widehat{LCL} 을 각각 구하고, [절차 2]는 단계 II에 속하는 부분으로 [절차 1]에서 작성된 관리한계선을 이용하여 공정이 안정 상태일 때를 가정하고 부분군의 개수 $k = 10,000$, 부분군의 크기 $n = 5$ 인 표본을 각 분포로부터 100개를 추출하여 FAR의 평균을 계산한다. 그 후 n 을 7, 9, 25로 늘려가면서 각 분포에서 부분군의 크기가 관리도의 성능에 미치는 영향을 비교한다. 마지막 절차인 [절차 3]도 단계 II의 부분으로 공정이 불안정 상태일 때 품질특성치의 중심 변화 차이를 달리하여 관리도의 성능을 확인하기 위하여 [절차 2]의 표본에서 중심을 $\mu' = \mu + \delta\sigma$ 으로 설정하여 δ 만큼 이동해가면서 관리도의 성능을 비교 평가한다. 여기서 δ 는 $0 < \delta \leq 0.3$ 에서 0.075씩 이동시키며 4개의 δ 를 선정하고 $0.3 < \delta \leq 1.1$ 에서 0.2씩 이동하여 4개의 δ 를 선정하였다.

모의실험에서는 자유도가 3, 5, 10인 카이제곱분포와 자유도가 5, 10, 30인 t -분포를 사용하였는데, 그 이유는 일반적으로 실제 현장에서 수집된 품질특성치의 분포는 왜도가 양수이거나 좌우대칭이면서 양쪽 꼬리가 두꺼운 분포가 많기 때문이다 (Minitab Inc, 2014). 따라서 왜도와 첨도의 영향력을 살펴 보기 위해 카이제곱분포와 t -분포를 각각 사용하였다.

3.3. 분석 결과

3.3.1. 카이제곱분포

모의실험은 왜도가 1.6인 자유도가 3, 왜도가 1.2인 자유도가 5, 왜도가 0.8인 자유도가 10인 카이제곱분포를 각각 이용하여 살펴보았으며, Table 3.1은 Shewhart의 $\bar{X} - R$ 관리도 (XBR), 반원 관리도 (SC), 최대 관리도 (MAX) 및 평균제곱오차 관리도 (MSE)의 FAR 평균을 측정된 결과이다. 공정이 안정 상태인 $\delta = 0$ 에서는 4개의 관리도 모두 정규분포 FAR의 값 0.0073 보다 큰 것을 알 수 있으며 자유도가 3이고 부분군 크기 n 이 9 또는 25, 자유도가 5이고 부분군 크기 n 이 25인 경우를 제외하고는 MAX의 FAR이 상대적으로 낮다. 따라서 공정이 안정 상태이고 품질특성치 X 의 왜도가 양수일 때 MAX가 더 로버스트하며 SC와 MSE는 편이가 크기 때문에 직접 비교하기가 어렵다. 또한 왜도와 첨도의 감소는 XBR과 MAX의 FAR을 낮춘다. 불안정 공정상태인 $\delta > 0$ 의 경우는 중심을 $\mu' = \mu_0 + \delta\sigma_0$ 만큼 이동하면서 FAR을 계산하였으며, δ 는 중심의 변화가 0.75 간격과 0.2 간격으로 증가할 때를 비교하였다. 전반적으로 자유도와 부분군 크기가 증가할수록 MAX의 공정이탈 감지 성능이 좋아지며 특히 자유도가 3이고 n 이 25인 경우, 자유도가 5이고 $\delta \geq 0.9$ 인 경우, 자유도가 10이고 $\delta \geq 0.5$ 인 경우에 MAX의 공정이탈 감지 성능이 좋다.

Figure 3.1은 자유도가 10이고 부분군 크기가 7인 카이제곱분포일 때 FAR을 나타낸 것인데 공정이 불안정 상태일 때 XBR보다 MAX가 성능이 다소 좋으며 앞서 언급한 것처럼 MSE와 SC는 로버스트하지 않다. 또한 Figure 3.2는 자유도가 10이고 부분군 크기가 25인 카이제곱분포일 때 FAR을 나타낸 것인데, 공정이 불안정 상태일 때 MAX의 FAR이 높아 공정 이탈을 빠르게 감지하고 있으며 부분군 크기 n 이 커짐에 따라 관리도 성능의 우열에 크게 변동은 없지만 δ 가 같을 때 전반적으로 공정의 이탈 감지 속도가 빨라진다.

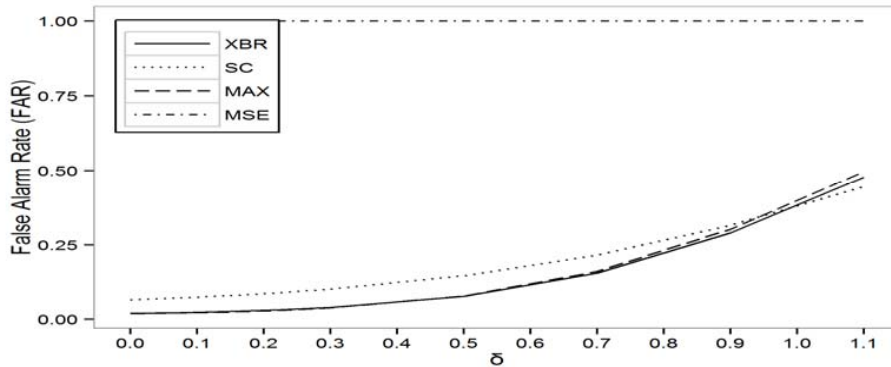


Figure 3.1 The FAR of four control charts under $\chi^2(10)$, $n = 7$

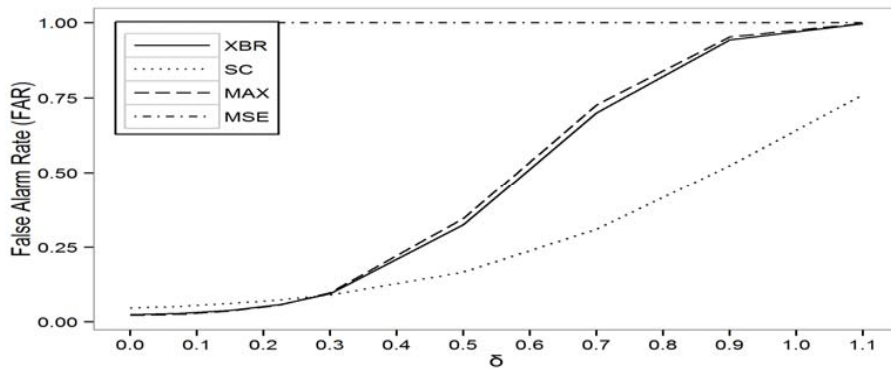


Figure 3.2 The FAR of four control charts under $\chi^2(10)$, $n = 25$

3.3.2. t -분포

모의실험은 첨도가 5.2인 자유도가 5, 첨도가 1.0인 자유도가 10, 첨도가 0.2인 자유도가 30인 t -분포를 각각 이용하여 살펴보았는데, Table 3.2는 5, 10, 30의 자유도를 가진 t -분포에서 부분군의 개수 $k = 10, 000$, 부분군의 크기 $n = 5, 7, 9, 25$ 가 되도록 각각 표본을 100개 추출하고 중심을 $\mu' = \mu_0 + \delta\sigma_0$ 로 이동하면서 XBR, SC, MAX, MSE의 FAR 평균을 측정된 결과이다.

공정이 안정 상태인 $\delta = 0$ 상황에서는 관리도 4개 모두 정규분포를 가정한 FAR의 측정기준 0.0073보다 크며 자유도가 5인 경우에는 MSE가 대체적으로 타 단일 관리도보다 더 로버스트한 것을 알 수 있으며 자유도가 증가함에 따라 MAX의 FAR이 점점 낮아져서 더욱 로버스트해진다. 공정이 불안정 상태인 경우는 동일한 자유도일 때 부분군 크기 n 이 증가할수록 각 관리도의 공정이탈 탐지능력이 우수해진다. 자유도가 5인 경우에는 MAX가 공정이탈 감지속도가 대체적으로 빠르며 자유도가 10과 30인 경우에는 MSE의 FAR이 대부분 높다.

Figure 3.3은 자유도가 5이고 부분군의 크기가 5인 t -분포일 때 FAR을 나타낸 것이며 중심의 변화가 클수록 MAX의 공정이탈 탐지속도가 빨라지는 것을 알 수 있다. Figure 3.4는 자유도가 5이고 부분군의 크기가 25인 t -분포일 때 FAR을 나타낸 것이다. 부분군의 크기가 증가하면서 MAX가 공정이 불안정 상태일 때 전반적으로 높은 성능을 보이며 δ 가 동일할 때 부분군의 크기가 증가하면 모든 관리도의 성능이 전반적으로 공정 이탈을 빠르게 탐지한다.

Table 3.2 The FAR of four control charts under t -distribution

type	df	n	δ								
			0.00	0.08	0.15	0.23	0.30	0.50	0.70	0.90	1.10
XBR	5	5	0.0659	0.0664	0.0678	0.0706	0.0748	0.0974	0.1502	0.2583	0.4336
		7	0.0725	0.0731	0.0750	0.0786	0.0843	0.1173	0.2016	0.3688	0.5980
		9	0.0779	0.0786	0.0807	0.0850	0.0918	0.1352	0.2520	0.4701	0.7195
		25	0.1181	0.1188	0.1215	0.1279	0.1420	0.2754	0.6175	0.9031	0.9850
	10	5	0.0183	0.0187	0.0199	0.0222	0.0260	0.0461	0.0913	0.1771	0.3113
		7	0.0192	0.0198	0.0215	0.0249	0.0304	0.0620	0.1379	0.2793	0.4753
		9	0.0205	0.0212	0.0234	0.0277	0.0350	0.0803	0.1914	0.3851	0.6196
		25	0.0273	0.0287	0.0342	0.0473	0.0749	0.2797	0.6525	0.9162	0.9910
	30	5	0.0104	0.0107	0.0119	0.0141	0.0176	0.0373	0.0832	0.1690	0.3021
		7	0.0101	0.0107	0.0124	0.0157	0.0212	0.0535	0.1311	0.2727	0.4682
		9	0.0103	0.0110	0.0132	0.0174	0.0249	0.0719	0.1855	0.3807	0.6153
		25	0.0117	0.0134	0.0200	0.0360	0.0688	0.2957	0.6753	0.9268	0.9929
SC	5	5	0.1125	0.1134	0.1159	0.1202	0.1268	0.1579	0.2188	0.3274	0.4958
		7	0.1097	0.1105	0.1132	0.1176	0.1240	0.1553	0.2182	0.3358	0.5236
		9	0.1103	0.1112	0.1138	0.1184	0.1250	0.1574	0.2256	0.3557	0.5649
		25	0.1210	0.1222	0.1259	0.1322	0.1417	0.1913	0.3048	0.5369	0.8348
	10	5	0.0710	0.0719	0.0752	0.0808	0.0889	0.1257	0.1913	0.2935	0.4314
		7	0.0576	0.0587	0.0617	0.0668	0.0747	0.1122	0.1833	0.3007	0.4643
		9	0.0513	0.0524	0.0555	0.0608	0.0689	0.1084	0.1877	0.3227	0.5099
		25	0.0390	0.0402	0.0441	0.0511	0.0626	0.1280	0.2837	0.5523	0.8252
	30	5	0.0549	0.0560	0.0593	0.0652	0.0739	0.1135	0.1834	0.2898	0.4293
		7	0.0404	0.0414	0.0445	0.0501	0.0583	0.0975	0.1731	0.2957	0.4610
		9	0.0339	0.0349	0.0379	0.0432	0.0514	0.0923	0.1755	0.3168	0.5068
		25	0.0207	0.0217	0.0248	0.0308	0.0409	0.1060	0.2714	0.5530	0.8263
MAX	5	5	0.0525	0.0531	0.0548	0.0577	0.0627	0.0898	0.1542	0.2822	0.4764
		7	0.0633	0.0640	0.0661	0.0703	0.0774	0.1206	0.2286	0.4268	0.6665
		9	0.0729	0.0737	0.0765	0.0820	0.0916	0.1543	0.3116	0.5636	0.8011
		25	0.1303	0.1323	0.1403	0.1599	0.2007	0.4733	0.8310	0.9784	0.9989
	10	5	0.0156	0.0161	0.0177	0.0206	0.0252	0.0497	0.1031	0.2006	0.3466
		7	0.0163	0.0170	0.0195	0.0238	0.0308	0.0699	0.1596	0.3168	0.5217
		9	0.0171	0.0181	0.0211	0.0269	0.0366	0.0942	0.2254	0.4358	0.6706
		25	0.0195	0.0222	0.0318	0.0535	0.0956	0.3571	0.7347	0.9479	0.9956
	30	5	0.0094	0.0098	0.0114	0.0142	0.0187	0.0429	0.0968	0.1934	0.3363
		7	0.0095	0.0102	0.0125	0.0168	0.0235	0.0627	0.1519	0.3062	0.5078
		9	0.0095	0.0104	0.0132	0.0188	0.0282	0.0851	0.2143	0.4219	0.6558
		25	0.0096	0.0120	0.0214	0.0429	0.0849	0.3460	0.7258	0.9452	0.9953
MSE	5	5	0.0459	0.0463	0.0475	0.0496	0.0527	0.0690	0.1059	0.1843	0.3314
		7	0.0521	0.0527	0.0545	0.0576	0.0625	0.0895	0.1567	0.3034	0.5384
		9	0.0568	0.0575	0.0599	0.0641	0.0712	0.1114	0.2191	0.4368	0.7092
		25	0.0734	0.0754	0.0822	0.0973	0.1276	0.3573	0.7533	0.9591	0.9962
	10	5	0.0188	0.0193	0.0212	0.0246	0.0300	0.0580	0.1182	0.2291	0.3973
		7	0.0191	0.0200	0.0225	0.0274	0.0351	0.0775	0.1745	0.3478	0.5759
		9	0.0197	0.0208	0.0241	0.0303	0.0405	0.1009	0.2399	0.4690	0.7229
		25	0.0217	0.0243	0.0337	0.0550	0.0955	0.3543	0.7452	0.9598	0.9981
	30	5	0.0103	0.0109	0.0127	0.0159	0.0212	0.0489	0.1118	0.2272	0.4001
		7	0.0102	0.0111	0.0135	0.0182	0.0258	0.0690	0.1694	0.3467	0.5784
		9	0.0103	0.0114	0.0145	0.0205	0.0305	0.0917	0.2339	0.4672	0.7237
		25	0.0101	0.0124	0.0214	0.0421	0.0823	0.3431	0.7410	0.9600	0.9982

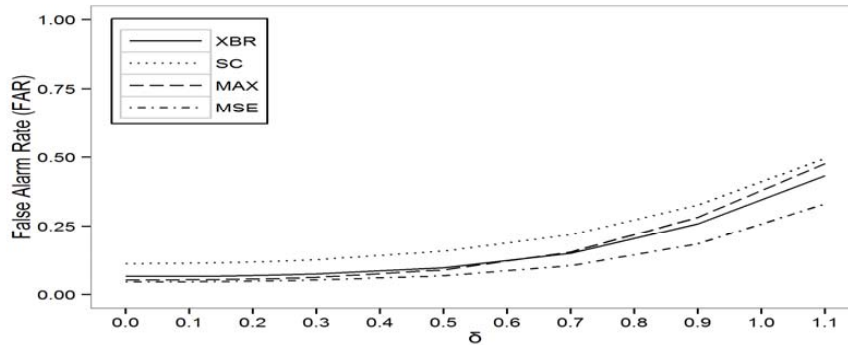


Figure 3.3 The FAR of four control charts under $t(5)$, $n = 5$

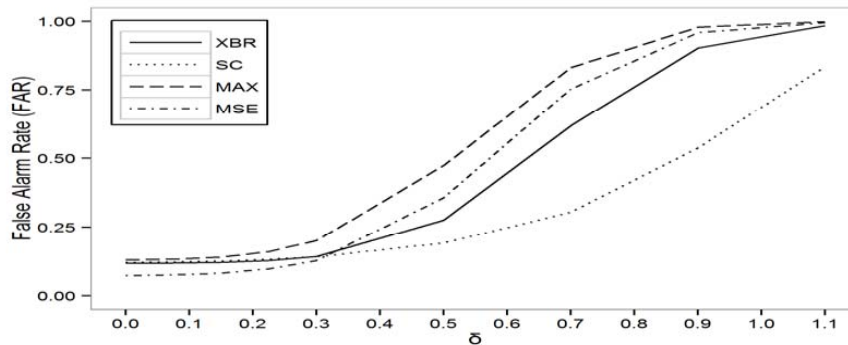


Figure 3.4 The FAR of four control charts under $t(5)$, $n = 25$

4. 결론

본 연구에서는 중심과 산포를 동시에 관리하는 단일 관리도들이 비정규성 데이터에 대해서 얼마나 로버스트성을 유지하는 지를 살펴보았다. 사용한 모의실험은 카이제곱분포와 t -분포를 이용하여 여러 가지 경우에 데이터를 생성한 후, 공정이 안정 상태와 불안정 상태를 각각 가정하고 측정지표인 FAR을 판단기준으로 사용하여 비교 평가하였다. 비교대상으로 사용한 관리도는 $\bar{X} - R$ 관리도와 단일 관리도인 반원관리도, 최대관리도, 평균제곱오차 관리도를 사용하였으며, 분석결과 공정이 안정 상태이면 카이제곱분포에서는 자유도와 상관없이 최대관리도의 성능이 좋고 t -분포에서는 자유도가 클수록 최대관리도가 더 로버스트하였다. 반대로 공정이 불안정 상태이면 카이제곱분포에서는 자유도와 부분군 크기 n 이 증가할수록 최대관리도가 우수했으며 t -분포에서는 자유도가 증가할수록 평균제곱오차 관리도의 공정 이탈 감지가 빨랐다. 본 연구에서는 실제 제조 및 품질 현장에서 수집되는 품질특성치의 데이터가 정규 분포를 따르지 않는 경우가 많기 때문에 단일 관리도의 비정규성 데이터에 대한 성능을 비교하였다. 향후 연구과제로는 단일 관리도에 대한 제1종 오류에 대해 편의를 줄이고 비정규성 데이터에 로버스트한 새로운 통계량을 더 깊이 연구한다면 단일 관리도의 활용도가 더욱 높아질 것으로 간주된다.

References

- Chao, M. T. and Cheng, S. W. (1996). Semicircle control chart for variable data. *Quality Engineering*, **8**, 441-446.
- Chen, G. and Cheng, S. W. (1998). Max chart: Combining X-bar chart and S chart. *Statistica Sinica*, **8**, 263-271.
- Cho, G. Y. (2010). Multivariate shewhart control charts with variable sampling intervals. *Journal of the Korean Data & Information Science Society*, **21**, 999-1008.
- Cho, G. and Park, J. (2011). A readjustment procedure in the multivariate integrated process control. *Journal of the Korean Data & Information Science Society*, **22**, 1123-1135.
- Cho, G. and Park, J. (2013). Parameter estimation in a readjustment procedure in the multivariate integrated process control. *Journal of the Korean Data & Information Science Society*, **24**, 1275-1283.
- Lee, H. Y. and Lee, J. (2009). Change point estimators in monitoring the parameters of an IMA(1,1) model. *Journal of the Korean Data & Information Science Society*, **20**, 435-443.
- Lee, J. (2011). Comparison of monitoring the output variable and the input variable in the integrated process control. *Journal of the Korean Data & Information Science Society*, **22**, 679-690.
- Levinson, W. A. (2000). SPC for real-world processes : What to do when the normality assumption doesn't work. *ASQ's Annual Quality Conference*, **54**, 82-89.
- Minitab Inc. Variables control charts in the assistant <http://support.minitab.com/en-us/minitab/17/technical-papers>.
- Park, H. I. (2013). Combined control chart for the mean and variance using likelihood method. *Journal of the Korean Data Analysis Society*, **15**, 2889-2898.
- Shewhart, W. A. (1926). Quality control charts. *System Technical Journal*, **5**, 593-603.
- Spiring, F. A. and Cheng, S. W. (1998). An alternate variables control chart : The univariate and multivariate case. *Statistica Sinica*, **8**, 273-287.

A comparison of single charts for non-normal data

Myunggoo Kang¹ · Jangtaek Lee²

^{1,2}Department of Statistics, Dankook University

Received 23 March 2015, revised 11 May 2015, accepted 24 May 2015

Abstract

In this paper, we compare the robustness to the assumption of normality of the single control charts to control the mean and variance simultaneously. The charts examined were semicircle control chart, max chart and MSE chart with Shewhart individuals control charts. Their in-control and out-of-control performance were studied by simulation combined with computation. We calculated false alarm rate to compare among single charts by changing subgroup size and shifting mean of quality characteristics. It turns out that max chart is more robust than any of the others if the process is in-control. In some cases max chart and MSE chart are more robust than others if the process is out-of-control.

Keywords: False alarm rate, max chart, semicircle control chart, single chart, MSE chart.

¹ Graduate student, Department of Statistics, Dankook University, Yongin 448-701, Korea.

² Corresponding author: Professor, Department of Statistics, Dankook University, Yongin 448-701, Korea. E-mail: jtlee@dankook.ac.kr