Proceedings of Joint Conference of Korean Data And Information Science Society and The Korean Data Analysis Society,
April 28 - 29, 2006, pp. 7-15

Bootstrap 방법을 이용한 중앙값 관리도의 구축

박효일[1]

요 약

이번 연구의 목적은 평균 대신에 중앙값을 이용한 관리도를 제시하며 관리한계선을 결정하기 위하여 표본 중앙값에 대한 근사분포에서 bootstrap 방법을 이용한 분산을 추정하는 연구를 한다.

1. 서론

생산 현장에서 제품의 품질 향상과 품질 유지를 위하여 사용되는 위치모수에 대한 계량형 관리도 중의 하나인 $\overline{X}$-관리도의 특징은 다음과 같다.

(i) production process의 분포가 정규분포라는 가정 하에서 작성됨
(ii) 특이치 (outlier)에 대하여 민감한 반응을 보임

그러나 실제로는 정규분포보다 꼬리 부분이 두꺼운(heavy-tailed) 속성을 지닌 분포를 갖는 자료가 생산현장에서 자주 발생한다는 것이 적지 않은 통계학자와 공학자들의 생각이다(cf. Gunter, 1989). 특이치가 발생하는 이유도 process의 분포가 정규분포로부터 벗어나기 때문이라는 것이 가장 큰 원인이라고 여긴다. 이와 같이 정규분포라는 가정이 용이하지 않은 경우에 통계적 추론 방법으로서 robust적인 여러 가지 통계적 절차가 제시되었지만 비모수적 방법이 가장 성공적이며 또한 널리 이용되고 있다. 평균을 대신할 수 있는 위치모수로는 중앙값 (median)이 있으며 중앙값을 이용하는 비모수적 검정절차와 추정이 많이 연구 제시되어 있다. 이번 연구에서도 중앙값을 이용하여 중앙값 관리도를 제시하는 것이 주된 목표이다. 물론 중앙값을 이용하는 관리도에 관한 연구는 이번이 처음이 아니며 기존에 여러 연구가 있으나 다음에 review하는 바와 같이 여러 문제점이 존재하기 때문에 실제 생산 현장에서 거의 이용되지 못하고 있는 것이다. 먼저 기존의 중앙값 관리도에 관하여 review하기 전에 중앙값과 관련된 몇 가지 용어를 먼저 기술하고자 한다. process 분포함수 F에 대하여 quantile function, F^{-1}의 정의는 다음과 같다: $0<p<1$에 대하여

$$F^{-1}(p) = \inf\{x: F(x) \geq p\}. \tag{1}$$

따라서 중앙값은 $p=1/2$인 경우의 quantile이다. 중앙값에 대한 비모수적 추정치는 소위

[1]360-764 충북 청주시 상당구 내덕동 36, 청주대학교 응용통계학과 교수. E-mail : hipark@cju.ac.kr

Hodges-Lehmann 추정치라고 하며 비모수 검정통계량을 이용하여 구한다. 예를 들어 부호검정통계량을 이용하면 통상적인 표본 중앙값이 얻어지며 Wilcoxon 부호순위통계량을 이용하면 연구 내용편에서 다루게 되는 형태의 Hodges-Lehmann 추정치를 얻는다. 다음은 지금까지 연구 제시된 중앙값 관리도에 대한 review이다.

a. 표본 중앙값 (sample median)을 이용한 관리도

Nelson (1982), Janacek and Meikle(1997) 그리고 Khoo(2005)는 중심선 (CL)으로서 평균 대신에 중앙값을 이용하는 관리도를 제시하였으나 Nelson은 여전히 process의 분포가 정규분포인 경우에 한정하였으므로 단지 CL에 대하여 중앙값을 이용한다는 점이며 정규분포를 이용하여 $3-\sigma$ 관리한계선을 제시하고 있다. Janacek and Meikle은 관리한계선을 구하기 위하여 이항분포 및 beta-분포를 이용하여 소위 double expectation theorem (cf. Bickel and Doksum, 1977)을 이용하여 beta-binomial 분포를 도출하였다. 그러나 본인들이 지적하였듯이 beta-binomial 분포함수 자체가 다루기 곤란한 (awkward)형태이며 특히 이산형이기 때문에 $3-\sigma$에 대한 정확한 관리한계선을 얻기가 어렵다. Khoo의 연구 결과는 표본 중앙값의 분포를 이용하여 관리한계선을 결정하는 연구이지만 Khoo의 표본 중앙값의 분포에 대한 연구는 잘못된 결론을 도출하였다. 다음은 표본 중앙값의 분포를 직접 이용하여 관리한계선을 결정하는 경우의 문제점에 관하여 설명하기로 한다.

표본 중앙값의 분포를 이용하여 관리한계선을 결정하는 경우의 문제점

표본 중앙값의 분포는 미지(unknown)의 분포함수 F뿐만이 아니라 확률밀도함수 (pdf)까지 포함하고 있으며 대표본 근사이론에 의한 정규근사식 조차도 극한 분산에 pdf를 포함하고 있어 미지의 pdf를 추정하여야 하므로 정확한 관리한계선은 물론 근사적 관리한계선도 구하기가 어렵다. 미지의 pdf를 추정하여 분산의 추정치를 구한다고 하더라도 pdf의 추정으로 인한 분산의 추정치가 unstable 하므로 이러한 방법으로 제시된 중앙값 관리도는 이용에 제한이 있는 것은 자명하다. 따라서 이러한 문제점을 해결하기 위하여 Nelson은 여전히 정규분포 가정 하에서 CL만이 중앙값으로 하는 관리도를 제시하였다.

b. Hodges-Lehmann 추정치와 이와 관련된 구간추정을 이용한 관리도

Alloway와 Raghavachari (1991)은 Wilcoxon 부호순위통계량을 이용하여 중앙값 추정치로서 소위 Hodges-Lehmann 추정치를 사용하였으며 관리한계선을 구하기 위하여 역시 Wilcoxon 부호순위통계량을 이용하여 중앙값에 대한 비모수적 구간추정을 이용하였다. 그러나 Wilcoxon 부호순위통계량의 분포가 이산형이므로 $3-\sigma$에 대한 관리한계선을 정확하게 정할 수 없으며 표본의 크기가 최소한 10 이상이 되어야 한다.

c. Quantile function을 이용한 관리도

Grimshaw와 Alt (1997)는 quantile function을 이용하여 전통적인 관리도라는 개념 대신에 가설검정의 접근방법에 의한 공정의 변화를 확인하는 관리도로서 visualize할 수 없다는 점과 공정 평균 또는 중앙값의 변화에 대한 판단을 위한 것이라기보다는 분포함수의 변화를 감지하는 용도로서 적합도 검정 (goodness-of-fit test)에 사용하는 것이 더욱 적절한 것으로 판단된다. 더욱이 보다 더 불편한 사항은 검정절차를 이용하는 관리도 형태이므로 작업현장에 직접 적용하기가 용이하지 않다는 점이다.

위에서 review한 중앙값 또는 quantile function을 이용하는 관리도에서 공통적인 애로 사항은 표본중앙값의 분포함수와 극한 분산에는 미지의 pdf를 포함하고 있기 때문에 pdf를 추정을 해야 하며 비모수적 구간추정을 이용하는 경우에는 이산형 분포함수를 이용하기 때문에 $3-\sigma$ 관리한계선을 사용하는 경우에 일정 수 이상의 표본의 크기가 요구되며 더욱이 주어진 확률에 대한 정확한 관리한계선을 결정하기가 어렵다는 데 있다.

분포함수 F에 대한 아무런 정보 없이 분산과 신뢰구간 등을 추정하는 데 주어진 자료만을 이용하여 추정할 수 있는 bootstrap 방법이 통계학에 도입된 지 거의 30년이 되었지만 관리도 분야에서는 거의 응용이 이루어 지지 않았다. 따라서 중앙값을 이용한 관리도에서 나타난 여러 가지 단점을 해결하고 극복하기 위하여 통계학의 추론분야에서 널리 사용되고 있는 bootstrap 방법을 이용하여 관리한계선을 결정하는 중앙값 관리도를 제시하는 연구를 하는 것이 금번 연구의 목적이다.

2. 중앙값 관리도

분포함수 F를 갖는 모집단에서 추출된 표본 $X_1, \cdots, X_n$을 이용한 표본중앙값 $\widetilde{X}$은 다음과 같이 정의된다.

$$\widetilde{X} = \mathrm{med}\{X_1, \cdots, X_n\}.$$

만약 표본의 크기 n이 홀수이면 즉 $n=2k+1$이면 $\widetilde{X} = X_{(k+1)}$이며, n이 짝수이면 즉 $n=2k$이면 $\widetilde{X} = \dfrac{X_{(k)} + X_{(k+1)}}{2}$ 라고 정의한다. 여기서 $X_{(i)}$는 표본 $X_1, \cdots, X_n$에서 i번째 순서통계량을 의미한다. 위치모수에 대한 계량형 관리도로서 $\overline{X}$-관리도 대신에 표본중앙값 $\widetilde{X}$을 이용하는 관리도를 제시하기 위하여 표본의 크기가 n인 r개의 표본이 있으면 r개의 표본중앙값 $(\widetilde{X}_1, \cdots, \widetilde{X}_r)$을 이용하여 다음과 같이 중심선($CL$)을 얻는다.

$$\widetilde{\widetilde{X}} = \mathrm{med}\{\widetilde{X}_1, \cdots, \widetilde{X}_r\}.$$

관리상한(UCL)과 관리하한(LCL)을 구하기 위한 $\widetilde{X}$의 분포는 다음과 같다. i번째 순서통계량

$X_{(i)}$의 확률밀도함수(pdf) $g(x_{(i)})$는 다음과 같다 (cf. Bickel and Doksum, 1977).

$$g(x_{(i)})=\frac{n!}{(i-1)!(n-i)!}[F(x_{(i)})]^{i-1}[1-F(x_{(i)})]^{n-i}f(x_{(i)}).$$

위에서 f는 분포함수 F에 대응하는 pdf이다. 따라서 $n=2k+1$인 경우에는 $\widetilde{X}=X_{(k+1)}$의 pdf $g(x_{(k+1)})$는 다음과 같다.

$$g(x_{(k+1)})=\frac{n!}{k!(n-k-1)!}[F(x_{(k+1)})]^{k}[1-F(x_{(k+1)})]^{n-k-1}f(x_{(k+1)}). \qquad (*)$$

UCL과 LCL은 적절한 α와 β를 선택하여 다음의 방정식의 해를 구한다.

$$\alpha=\int_{UCL}^{\infty}g(x_{(k+1)})dx \text{ 그리고 } \beta=\int_{-\infty}^{LCL}g(x_{(k+1)})dx.$$

예를 들어 F가 정규분포이며 $3-\sigma$ 관리한계선이라면 $\alpha=\beta=0.005$를 선택할 수 있으며 지수분포같이 대칭형이 아닌 경우에는 반드시 $\alpha=\beta$일 필요는 없다. $n=2k$이면 $\widetilde{X}$에 대한 pdf를 정확하게 도출할 수 없어 근사식을 이용하여 구한다. 그러나 $n=2k+1$인 경우에도 pdf에는 미지(unknown)의 f가 포함되어 있으므로 UCL과 LCL에는 f의 간섭을 피할 수 없어 분포함수에 대한 정보를 보유해야한다. 만약 대표본 근사이론을 이용한다면 quantile 함수 또는 순서통계량에 대한 Bahadur representation theorem (cf. Serfling, 1980)을 이용하여 표본중앙값에 대한 정규근사의 결과는 다음과 같다.

$$\sqrt{n}(\widetilde{X}-\theta)\simeq N\left(0,\frac{1}{4f^2(\theta)}\right) \text{ 또는 } \widetilde{X}-\theta\simeq N\left(0,\frac{1}{4nf^2(\theta)}\right).$$

따라서 UCL과 LCL은 다음과 같다.

$$UCL=\widetilde{\widetilde{X}}+\frac{z_{\alpha}}{2\sqrt{n}f(\theta)} \text{ 그리고 } LCL=\widetilde{\widetilde{X}}-\frac{z_{\beta}}{2\sqrt{n}f(\theta)}.$$

여기서 $\Phi(z_{\alpha})=1-\alpha$이며 Φ는 누적표준정규분포이다. 이와 같이 근사식을 이용하여도 UCL과 LCL의 표현식에는 f의 간섭을 피할 수는 없다. 따라서 "연구목적"에서 언급된 중앙값을 이용하는 관리도에 대한 연구결과는 모두 f의 간섭을 피하기 위한 것이나 별로 성공적이지 않으며 심지어 잘못된 결론에 도달한 논문도 있다.

Maritz와 Jarrett (1978)은 $\widetilde{X}$의 극한분산 $1/(4nf^2(\theta))$의 추정치 $\widehat{\sigma}^2$를 순서통계량과 식 (*)를 이용하여 $n=2k+1$인 경우에 다음과 같이 제시하였다.

$A_{dn}=\sum_{i=1}^{n}X_{(i)}^{d}W_i,\ \ W_i=\frac{(2k+1)!}{(k!)^2}\int_{(i-1)/n}^{i/n}y^k(1-y)^k dy$이라고 하면

$$\widehat{\sigma^2}=A_{2n}-A_{1n}^2.$$

$n=2k$인 경우에는 $\widetilde{X}=(X_{(k)}+X_{(k+1)})/2$이므로

$$V(\widetilde{X})=\frac{1}{4}E(X_{(k)}^2+X_{(k+1)}^2)+\frac{1}{2}E(X_{(k)}X_{(k+1)})-\frac{1}{4}\{E(X_{(k)}+X_{(k+1)})\}^2.$$

이러한 관계를 이용하여 다음의 quantities를 구하여 $\widehat{\sigma^2}$를 얻었다.

$$B_{dn}=\sum_{i=1}^{n}X_{(i)}^{d}V_i,\ \ V_i=\frac{(2k)!}{2[k!(k-1)!]}\int_{(i-1)/n}^{i/n}y^{k-1}(1-y)^{k-1}dy$$

$$C_n=\sum_{i=1}^{n}\sum_{j=i}^{n}X_{(i)}X_{(j)}V_{ij},\ \ V_{ij}=\frac{2k!}{[(k-1)!]^2}\int_{(i-1)/n}^{i/n}y^{k-1}\int_{\max(y,(i-1)/n)}^{j/n}(1-x)^{k-1}dxdy$$

$$\widehat{\sigma^2}=\frac{1}{2}(B_{2n}+C_n)-B_{1n}^2.$$

따라서 Maritz와 Jarrett가 제시한 극한분산의 추정치를 이용하면 pdf에 대한 추정절차 없이 정규근사식을 이용하여 UCL과 LCL을 구할 수가 있다. 여기서 주목할 점은 Maritz와 Jarrett가 제시한 분산의 추정치는 bootstrap 추정치와 동일하다는 것이며 (cf. Shao and Tu, 1995) 아직은 이 추정치가 널리 사용되지 않고 있다. 따라서 이번 연구에서는 Maritz와 Jarrett가 제시한 $\widehat{\sigma^2}$을 이용한 중앙값 관리도에 대한 연구도 수행하게 된다. 이에 따라 $\widehat{\sigma^2}$에 대한 bias와 효율성에 대한 연구도 진행하게 된다.

이번 연구의 주요 목표인 bootstrap 방법을 이용하여 UCL과 LCL을 결정하는 연구를 위하여 먼저 bootstrap 분산 추정에 대한 연구를 하게 될 것이다. 그러나 놀랍게도 bootstrap 방법을 관리도 분야에 적용하는 연구는 중앙값 관리도에 대한 연구 결과보다도 더욱 연구결과가 드물었다. Seppala 등 (1995)은 소표본 그리고 크게 왜곡(skewed)된 분포인 경우 $\overline{X}$-관리도에 대한 UCL과 LCL을 결정하는 데 bootstrap percentile 방법을 이용하였으며 Liu and Tang (1996)은 dependent data에 대한 관리도를 제시하면서 역시 관리한계선을 결정하는 데 있어 bootstrap 방법을 사용하였다. Jones and Woodall (1998)은 $\overline{X}$-관리도에 대한 UCL과 LCL을 결정하는 데 사용되는 bootstrap 방법에 대한 효용성을 average run length를 이용하여 simulation을 통하여 연구하였다. Shao and Tu(1995)가 밝혔듯이 소표본인 경우에는 $\widehat{\sigma^2}$가 bootstrap 추정치와 일치한다는 것은 밝혀졌으나 표본의 크기가 큰 경우에는 Monte Carlo 방법을 이용하기도 한다. 따라서 simulation 연구에 이용하여 두 종류의 bootstrap 추정치에 대한 비교 연구를 할 필요가 있다.

UCL과 LCL을 정하기 위하여 bootstrap 방법을 이용하는 또 다른 접근 방법으로서 bootstrap

분포함수를 이용하는 방법이 있다. 구체적으로 $\widetilde{X}$의 분포에 대한 bootstrap 방법의 이용에 관한 본격적인 논의에 앞서 먼저 다음의 사실을 먼저 기술한다. 중앙값은 앞서 지적한 바와 같이 quantile values 중의 하나임을 주목한다. 만약 F_n을 표본 $X_1, \cdots, X_n$에 의한 empirical distribution function 이라고 하고 F_n^{-1}를 F_n에 대응하는 quantile function이라고 한다면 표본 quantile function은 다음과 같다. $0<p<1$에 대하여

$$F_n^{-1}(p) = \inf\{x: F_n(x) \geq p\}.$$

표본중앙값 $\widetilde{X}$은 $p=1/2$인 경우이므로 다음과 같다.

$$\widetilde{X} = \inf\{x: F_n(x) \geq 1/2\}.$$

따라서 표본중앙값 $\widetilde{X}$은 L-통계량이다. $X_1^*, \cdots, X_n^*$를 $X_1, \cdots, X_n$ (또는 F_n)으로부터 추출된 bootstrap 표본이라고 하고 F_n^*와 F_n^{*-1}를 bootstrap empirical function 그리고 bootstrap quantile function이라고 하자. 그러면 우리의 목표는

$$\sqrt{n}(F_n^{*-1}(1/2) - F_n^{-1}(1/2)) = \sqrt{n}(\widetilde{X}^* - \widetilde{X})$$

의 분포, P_{F_n}를 통하여

$$\sqrt{n}(F_n^{-1}(1/2) - F^{-1}(1/2)) = \sqrt{n}(\widetilde{X} - \theta)$$

의 분포, P_F를 추정하고자 하는 것이다. P_{F_n}와 P_F의 관계를 기술하기 전에 bootstrap에서 사용하는 consistency에 대한 정의를 기술한다. 사실 consistent라는 용어는 점 추정치에서 사용되는 용어이나 bootstrap에서는 다음과 같이 정의된다. 여기서 ρ는 실수 위에서 정의되는 모든 분포들의 공간상의 metric(거리함수)의 하나이다.

정의: $n\to\infty$함에 따라 $\rho(P_{F_n}, P_F) \xrightarrow{a.s.} 0$이면 P_{F_n}는 P_F에 대하여 strongly ρ-consistent하다라고 한며 $\rho(P_{F_n}, P_F) \xrightarrow{p} 0$이면 P_{F_n}는 P_F에 대하여 weakly ρ-consistent하다라고 한다.

다음의 P_{F_n}와 P_F의 관계는 잘 알려진 결과이다.

정리. bootstrap 추정치 P_{F_n}는 P_F에 대한 strongly consistent하다.

증명. $\widetilde{X} = F_n^{-1}(1/2)$이 L-통계량이므로 Bickel and Freedman (1981), Singh (1981) 그리고 Csörgö and Mason (1989) 등을 이용하면 쉽게 도출할 수 있다.

정리를 이용하여 UCL과 LCL를 다음과 같이 결정할 수 있다. 먼저 q_α를 bootstrap 분포 P_{F_n}의 upper αth percentile point라고하자. 즉 다시 말하면

$$P_{F_n}\{\sqrt{n}(\widetilde{X}^* - \widetilde{X}) > q_\alpha\} = \alpha.$$

정리를 이용하면 다음의 결과를 얻는다.

$$P_F\{\sqrt{n}(\widetilde{X} - \theta) > q_\alpha\} \approx \alpha.$$

따라서 UCL은 다음과 같이 결정된다.

$$UCL = \widetilde{\widetilde{X}} + q_\alpha/\sqrt{n}.$$

LCL을 구하기 위하여 q_α와 같은 방법으로 다음과 같이 lower βth percentile point $q_{1-\beta}$를 먼저 구한다.

$$P_{F_n}\{\sqrt{n}(\widetilde{X}^* - \widetilde{X}) < q_{1-\beta}\} = \beta.$$

정리를 다시 한 번 더 사용하면 다음과 같다.

$$P_F\{\sqrt{n}(\widetilde{X} - \theta) < q_{1-\beta}\} \approx \beta.$$

따라서 LCL은 다음과 같이 결정된다.

$$LCL = \widetilde{\widetilde{X}} + q_{1-\beta}/\sqrt{n}.$$

n의 값이 작은 경우에는 정확한(exact) P_{F_n}, 즉 정확한 bootstrap 분포를 얻을 수는 있어 직접 UCL과 LCL을 결정할 수 있으나 상당히 지루한 계산이 요구되므로 Monte Carlo 방법에 의한 근사분포를 얻는 것이 일반적인 접근방법이다. 따라서 반복적인 표본 재 추출 (re-sampling) 과정이 필요하게 되며 이러한 이유로 bootstrap 방법이 표본 재 추출 방법이라고 인식되고 있는 이유이다. 그런데 표본 재 추출 방법은 주어진 q에 대하여 다음과 같은 확률은 구하기 쉽다.

$$P_{F_n}\{\sqrt{n}(\widetilde{X}^* - \widetilde{X}) > q\}.$$

그러나 다음과 같이 주어진 확률 α에 대하여 quantile point q_α를 구하기는 어렵다.

$$P_{F_n}\{\sqrt{n}(\widetilde{X}^* - \widetilde{X}) > q_\alpha\} = \alpha.$$

왜냐하면 Monte Carlo 방법은 표본 재 추출 때 마다 $\widetilde{X}^* - \widetilde{X}$만을 계산하므로 주어진 q와는 비

교할 수 있지만 확률 α와는 비교할 수 있는 길이 없기 때문이다. 그런데 우리가 추구하는 UCL과 LCL은 바로 quantile points이기 때문에 Monte Carlo 방법에 의한 bootstrap 분포 P_{F_n}를 직접 이용할 수는 없어 bootstrap 신뢰구간을 이용하여 UCL과 LCL을 결정하는 방법을 취하기로 한다. bootstrap 신뢰구간을 구하는 방법으로 몇 가지가 사용되고 있으나 가장 널리 쉽게 사용되는 방법은 소위 bootstrap percentile 방법으로서 다음과 같은 절차로 사용되고 있다.

다음은 UCL과 LCL을 정하기 위하여 $\widetilde{X}$의 bootstrap 근사분포에서 $\beta\times100\%$th와 $(1-\alpha)\times100\%$th percentile points를 구하기 위한 bootstrap percentile 방법을 이용하는 절차를 다음과 같이 예시한다.

(i) 표본 $X_1,\cdots,X_n$로부터 bootstrap 표본 $X_1^*,\cdots,X_n^*$를 복원 추출한다.

(ii) $X_1^*,\cdots,X_n^*$로부터 empirical distribution F_n^*를 이용하여 quantile function F_n^{*-1}을 구하여 bootstrap median $\widetilde{X}^*$를 구한다.

(iii) (i)과 (ii)의 과정을 B번 반복한다. 여기서 B의 크기는 1000이상이어야 한다.

(iv) (iii)까지 구한 B개의 $\widetilde{X}^*$를 다음과 같이 크기 순으로 나열한다. $\widetilde{X}^*_{(1)}\leq\widetilde{X}^*_{(2)}\leq\cdots\leq\widetilde{X}^*_{(B)}$

(v) $LCL=\widetilde{X}^*_{([\beta B])}$, $UCL=\widetilde{X}^*_{([(1-\alpha)B])}$로 결정한다.

위에서 $[a]$는 실수 a를 초과하지 않는 정수이다. 또 다른 bootstrap 신뢰구간을 구하는 방법으로는 bias를 줄이기 위한 bootstrap bias-corrected percentile 방법, bootstrap accelerated bias-corrected percentile 방법 그리고 hybrid bootstrap 방법 등이 있으나 모두 복잡한 functional form을 가진 통계량에 대한 bootstrap 방법으로서 비교적 단순한 형태인 표본중앙값에 대한 bootstrap 방법으로는 bootstrap percentile 방법과 bootstrap-t 방법으로 충분하다고 할 수 있다.

결론적으로 말하자면 bootstrap 방법을 이용하여 직접 극한 분산의 추정치를 이용할 수도 있고 또는 bootstrap 분포를 이용할 수도 있다. 이상 언급된 bootstrap 이용방법 중에서 어떤 것이 보다 효율적인 가는 아직 연구된 바가 없다. 따라서 이상 언급된 방법에 대하여 효율을 알아보고 지금까지 널리 이용되고 있는 자료를 이용하여 우리가 제시한 중앙값 관리도를 예시할 것이며 마지막으로 simulation을 통하여 다양한 분포에 대한 적응도와 $\overline{X}$-관리도에 대한 performance를 비교하게 될 것이다.

참고문헌

[1] Alloway, Jr. J. A. and Raghavachari, M. (1991). Control chart based on the Hodges-Lehmann estimator, *Journal of Quality Technology*, 23, 336-347.

[2] Bickel, P. J. and Doksum, K. A. (1977). *Mathematical Statistics-Basic Ideas and Selected Topics*, Holden-Day, Inc. San Francisco.

[3] Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap, *Annals of Statistics*, 9, 1196-1217.

[4] Grimshaw, S. D. and Alt, F. B. (1997). Control charts for quantile function values, *Journal of Quality Technology*, 29, 1-7.

[5] Gunter, B. H. (1989). The use and abuse of C_{pk}, Part 2, *Quality Progress*, 22, 108-109.

[6] Khoo, M. B. C. (2005). A control chart based on sample median for the detection of a permanent shift in the process mean, *Quality Engineering*, 17, 243-257.

[7] Janacek, G. J. and Meikle, S. E. (1997). Control charts based on medians, *Journal of the Royal Statistical Society, Ser. D-The Statistician*, 46, 19-31.

[8] Jones, L. A. and Woodall, W. H. (1998). The performance of bootstrap control charts, *Journal of Quality Technology*, 30, 362-375.

[9] Liu, R. Y. and Tang, J. (1996). Control charts for dependent and independent measurements based on bootstrap methods, *Journal of the American Statistical Association*, 91, 1694-1700.

[10] Maritz, J. S. and Jarrett, R. G. (1978). A note on estimating the variance of the sample median, *Journal of the American Statistical Association*, 73, 194-196.

[11] Nelson, L. S. (1982). Control chart for medians, *Journal of Quality Technology*, 14, 226-227.

[12] Seppala, T., Moskowitz, H., Plante, R. and Tang, J. (1995). Statistical process control via the subgroup bootstrap, *Journal of Quality Technology*, 27,139-153.

[13] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

[14] Shao, J. and Tu, D. (1995). *The jackknife and bootstrap*, Springer, New York.

[15] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap, *Annals of Statistics*, 9, 1187-1195.