응용통계연구 제 9권 2호
1996년 9월 pp. 1-11

붓스트랩방법을 활용한 절사추정량의 이론 및 응용연구[1)]

이 재 창[2)], 전 명 식[3)], 강 창 완[4)]

요 약

주어진 일변량자료의 위치모수에 대한 추정방법으로 α-절사평균 (α-trimmed mean)은 로버스트 성질 등의 장점에 근거하여 오랜기간을 두고 연구되어왔다. 본 연구에서는 자료에 근거한 α의 선택을 붓스트랩(bootstrap) 방법에 의거하여 제시하고 그에 근거한 새로운 절사추정량의 응용을 다루었다. 나아가 추정의 정도를 높이기 위해 이중붓스트랩(double bootstrap)을 사용하였다. 한편, 이러한 붓스트랩방법의 타당성 연구를 위해 모의실험(simulation)과 실제 자료 적용에 기존방법과 새로운 방법의 비교연구를 제시하였다.

1. 서 론

랜덤표본 $X_1, X_2, \cdots, X_n$ 의 모집단의 위치모수(location parameter)에 대한 전통적 추정방법인 표본평균 $\overline{X_n}$ 는 계산의 단순성과 정규성하에서의 수리적 최적성등의 이유로 널리 이용되어왔다. 그러나, 잘 알려진바와 같이, 표본평균은 너무 크거나 작은 관찰값에 대하여 로버스트(robust)하지 못한 단점이 있으며 이에 로버스트한 추정량으로 α-절사평균(α-trimmed mean)을 고려할 수 있다. α-절사평균은 표본을 크기순으로

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}$$

와 같이 나열하여 크기가 작거나 큰 부분을 제거하고 가운데 $(1-2\alpha)\cdot 100\%$의 표본만을 이용하여

$$m(\alpha) = \frac{1}{n-2[n\alpha]} \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{(i)}$$

로 위치모수 θ를 추정하는 것이다. 이러한 α-절사평균의 이용은 기저분포가 정규성을 만족하지 않는 경우 효율성이 매우 높은 것으로 알려져있다 (Bickel & Doksum, 1977).

---

1) 이 연구는 1994년도 한국과학재단의 핵심전문연구비에 의하여 연구되었음.
2) (136-701) 서울시 성북구 안암동 고려대학교 통계학과 교수.
3) (136-701) 서울시 성북구 안암동 고려대학교 통계학과 교수.
4) (136-701) 서울시 성북구 안암동 고려대학교 통계연구소 연구원.

그런데, 개념적으로 쉽고 합리적인 방법임에도 불구하고, α-절사평균 $m(\alpha)$는 α가 고정되어 있다는 점에서 실제 사용에 많은 문제점이 있다. Jaeckel(1971)은 $m(\alpha)$의 극한분산의 추정량을 최소화하는 α의 선택법을 제시하였으나 이는 불안정하여 실용성에 제약을 주고 있다. 한편, Efron(1979)에 의해 제안된 붓스트랩(bootstrap) 방법은 기저분포에 대한 모수적 가정없이도 표본으로부터 재표본(resampling)을 취하는 방법에 근거하여 추정량의 편의와 분산은 물론 표본분포의 추정에도 활용될 수 있다. 이러한 붓스트랩에 관한 연구는 Beran(1984), Hall(1992) 등에 의해 많은 발전이 있었으며 컴퓨터의 발전과 더불어 그 실용성이 날로 높아지고있다.

본 논문에서는 붓스트랩방법을 이용하여 추정량 $m(\alpha)$의 분산을 추정하여 이를 최소화하는 절사량 α^*의 선택을 제안하고, 그에 따른 최적절사추정량 $m(\alpha^*)$의 성질과 응용을 다루고자 한다. 2장에서는 붓스트랩방법의 사용을 설명하고 그에 따른 문제점을 보완하는 방법으로 이중 붓스트랩(double bootstrap)을 제안하고자 한다. 3장에서는 모의실험을 통해 기존의 방법과 본 논문에서 제안한 방법들을 비교하겠으며, 4장에서는 응용사례들을 살펴보았다.

2. 붓스트랩방법

2.1 이론적 정당성

위치모수 θ에 대해 대칭인 모분포 $F(x) = F_o(x-\theta)$ 로부터 구한 랜덤표본의 순위통계량을 $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ 이라 하자. 그러면, 적절한 조건하에서, α-절사평균 $m(\alpha)$는

$$\sqrt{n}\,(\,m(\alpha) - \theta\,) \;\Rightarrow\; N(0, \sigma^2(\alpha))$$

를 만족한다. 이 때, 극한분산 $\sigma^2(\alpha)$는 모분포 F와 절사량 α에 의존한다. Jaeckel (1971)은 $\sigma^2(\alpha)$의 추정량을 최소화하는 α의 사용을 제안하였으나 이는 매우 불안정하여 이에 대한 개선책이 요구되어왔다.

한편, Efron(1979)은 모분포 F와 표본 $\mathbf{X}=(X_1,X_2,\cdots,X_n)$의 함수로서 관심대상인 확률변량 $R_n(\mathbf{X}, F)$ 의 표본분포를 재표본과정을 통해 근사하는 붓스트랩방법을 제안하였다. 이 방법의 근본생각은 $R_n(\mathbf{X}, F)$의 표본분포 $J_n(x, F) = P_F[R_n(\mathbf{X}, F) \le x]$ 를 추정함에 있어 미지의 분포 F 를 경험적 분포 F_n 으로 대체하여 $J_n(x, F_n)$ 를 사용하는 것으로, 추정량 $J_n(x, F_n)$의 폐쇄형은 매우 복잡하나 몬테칼로 근사법이 가능하다는 것이다. 이러한 붓스트랩방법에는 미지의 분포 F를 어떤 분포로 대체하느냐에 따라 여러 가지 버전이 있으며 이들의 이론적 타당성과 응용성이 연구대상이 되고 있다.

붓스트랩방법의 일치성

이제, α-절사평균의 붓스트랩분포의 수리적 타당성을 알아보자. 우선, $\sqrt{n}\,(\,m(\alpha) - \theta)$의 표본분포를

$$J_n(x,F) = P_F [\sqrt{n} (m(\alpha) - \theta) \le x]$$

라고 표기하자. 다음으로, $X_1^*, X_2^*, \cdots, X_n^*$ 를 경험적 분포 F_n으로부터의 랜덤표본(이하 붓스트랩표본이라 부름)이라 하고 그의 α-절사평균을 $m^*(\alpha)$라고 하면, $J_n(x, F)$의 붓스트랩분포는

$$J_n(x, F_n) = P_{Fn} [\sqrt{n} (m^*(\alpha) - m(\alpha)) \le x]$$

이 된다. 여기서 $\sqrt{n} (m(\alpha) - \theta)$의 실제 표본분포 $J_n(x,F)$와 그의 붓스트랩분포 $J_n(x,F_n)$가 같은 극한분포를 가지면 두 분포는 표본의 크기가 커짐에 따라 가까와지며 붓스트랩방법은 일치성(consistency)을 갖는다고 한다. 붓스트랩방법의 일치성에 대한 일반적인 필요충분조건은 아직 완전한 상태로 구해지지 않았으며 주어진 문제에 따라 여러 버전이 활용되고 있다. 또한, 이러한 일치성이 만족되면 붓스트랩방법에 의한 신뢰영역구축이 가능하며 이는 근사적으로 정확한 신뢰계수를 갖는다.

한편, α-절사평균의 경우 분위수과정(quantile process) $Q_n(t) = \sqrt{n} [F_n^{-1}(t) - F^{-1}(t)]$ 를 사용하면

$$\sqrt{n} (m(\alpha) - \theta) = \frac{1}{1-2\alpha} \int_{\alpha}^{1-\alpha} Q_n(t)\, dt$$

로 표현되며 이를 붓스트랩하면

$$\sqrt{n} (m^*(\alpha) - m(\alpha)) = \frac{1}{1-2\alpha} \int_{\alpha}^{1-\alpha} Q_n^*(t)\, dt$$

가 된다. (단, $Q_n^*(t) = \sqrt{n} [F_n^{*-1}(t) - F_n^{-1}(t)]$). 그런데, Q_n 과 Q_n^* 는 같은 극한과정을 가지므로 $J_n(x, F)$와 $J_n(x, F_n)$이 같은 극한 분포를 가지게 된다.

붓스트랩방법의 최적성

이와 같은 붓스트랩의 일치성에 근거하여 $m^*(\alpha)$의 분산을 최소화하는 절사량 $\alpha = \alpha^*$를 얻을 수 있다. 따라서, 절사량 α^*를 사용한 $m(\alpha^*)$를 자료에 근거한 절사추정량으로 제안한다. 이러한 절사량 α^*는, a를 미지의 기저분포 F에 대한 최적절사량이라 할 때, 붓스트랩방법의 일치성에 의하여,

$$\lim_{n \to \infty} \frac{E [m(\alpha^*) - \theta]^2}{E [m(a) - \theta]^2} = 1$$

를 만족한다(Kang,1993). 즉, 붓스트랩에 의해 선택된 절사량 α^*를 사용한 절사추정량 $m(\alpha^*)$는 미지의 분포 F를 안다고 가정할 때 얻을 수 있는 최소평균제곱오차와 같은 평균제곱오차를 근사적으로 가진다는 점에서 '근사 최적성(asymptotic optimality)'이 있다.

나아가, 가설검정이나 신뢰구간구축등을 위해, 우리가 원하는 것은 $m(\alpha^*)$의 표본분포인데, α^*가 자료에 근거한(data dependent) 선택이기 때문에 단순히 $m^*(\alpha^*)$의 분포를 활용하는 데에는 문제점이 따른다. 따라서, 첫 번째 붓스트랩에서 구한 각 붓스트랩표본에 관한 최적절사량 α^{**}을 다시 붓스트랩을 통해 얻어내서 $m^*(\alpha^{**})$의 분포를 사용하는 이중붓스트랩방법을 제안하고자 한다. 이렇게 구한 $m^*(\alpha^{**})$의 분포는 위치모수 θ에 관한 신뢰구간을 비롯한 통계적 추론에 활용될 수 있다.

2.2 몬테칼로 근사법

이제 붓스트랩방법의 실제적인 사용에 필요한 몬테칼로(Monte Carlo) 근사방법은 다음과 같이 구해진다.

(단계 1) 주어진 표본 $X_1, X_2, \cdots, X_n$ 으로부터 모집단의 경험적 분포 F_n을 만든다.

(단계 2) 경험적 분포 F_n으로부터 붓스트랩표본 $X_1^*, X_2^*, \cdots, X_n^*$ 를 얻고 그에 근거한 $m^*(\alpha)$를 구한다.

(단계 3) 단계 2를 독립적으로 반복시행하여 B_1개의 $m^*(\alpha)$값을 가능한 α의 값에 대해 구하고 그로부터 $m^*(\alpha)$의 분산을 최소화하는 절사량 $\alpha = \alpha^*$를 얻는다.

이러한 과정을 통해 구한 붓스트랩분산 $Var[m^*(\alpha)]$을 최소화하는 α^*를 사용하여 최적절사평균 $m(\alpha^*)$를 얻는다. 이제, 우리가 원하는 바는 $m(\alpha^*)$의 분산이며 나아가 표본분포를 통한 위치모수에 대한 신뢰구간의 구축이다. 단순하게 생각하면, $Var[m^*(\alpha^*)]$로 $m(\alpha^*)$의 분산을 추정할 수도 있겠으나 이는 $Var[m(\alpha^*)]$를 과소추정하는 경향이 있다. 왜냐하면 최적절사량 α^*는 $Var[m^*(\alpha)]$를 최소화하는 α이기 때문이다. 이러한 문제는 다음과 같은 이중붓스트랩을 통하여 해결될 수 있다.

(단계 4) 단계 2에서 구한 각 붓스트랩표본의 경험적 분포 F_n^*로부터 이중붓스트랩표본 $X_1^{**}, X_2^{**}, \cdots, X_n^{**}$를 얻고 그에 근거한 $m^{**}(\alpha)$를 구한다.

(단계 5) 단계 4를 독립적으로 반복시행하여 B_2개의 $m^{**}(\alpha)$값을 가능한 α의 값에 대해 구하고 그로부터 $m^{**}(\alpha)$의 분산을 최소화하는 $\alpha = \alpha^{**}$ 를 단계 2의 각 붓스트랩표본에 대하여 모두 B_1개 구한다.

(단계 6) 단계 5에서 구한 α^{**}를 해당하는 붓스트랩표본의 절사량으로 사용한 B_1개의 $m^*(\alpha^{**})$의 값을 $m(\alpha^*)$의 표본분포 추정에 그리고 그의 분산을 $Var[m(\alpha^*)]$의 추정량으로 사용한다.

물론 이와 같은 근사방법에 계산능력이 좋은 컴퓨터의 이용은 필수적이며 지수적으로 좋아지는 컴퓨터환경은 이러한 근사방법을 현실화시키고있다.

3. 모의실험

이 장에서의 모의실험은 세 가지로 구분되어 시행되었다. 첫번째는 기존의 Jaeckel이 제시한 방법과 붓스트랩방법을 절사량 α에 대한 선택의 측면에서 살펴보았으며, 두번째는 두 방법에 의한 α-절사평균의 평균제곱오차를 비교하였다. 마지막으로 붓스트랩방법에 의해 구해진 최적 절사추정량 $m(\alpha^*)$의 분산추정에 대하여 '단순' 붓스트랩과 이중붓스트랩의 차이를 알아보았다.

다음의 <표 1>은 N(0,1) 분포로부터 크기가 20인 랜덤표본에 근거하여 절사량 α의 값에 해당하는 절사평균의 분산을 붓스트랩방법과 Jaeckel의 방법에 대해 1000회의 독립시행을 통해 구한 것이다. 붓스트랩방법에 필요한 반복횟수는 B_1=200을 사용하였으며, 표에 나타난 숫자는 α-절사평균의 분산에 20을 곱한 것이다. $\sigma^2(\alpha) = n\ Var[m(\alpha)]$이다.

<표 1> α-절사평균의 분산추정 비교

α	붓스트랩방법		Jaeckel방법		$\sigma^2(\alpha)$
	평균	표준편차	평균	표준편차	
0.05	1.022	0.344	0.955	0.333	1.026
0.10	1.066	0.377	0.976	0.369	1.061
0.15	1.105	0.417	1.003	0.423	1.099
0.20	1.150	0.467	1.034	0.487	1.145
0.25	1.204	0.530	1.043	0.556	1.193
0.30	1.266	0.608	1.050	0.650	1.250
0.35	1.345	0.709	1.064	0.804	1.314

<표 1>에서 보면 붓스트랩방법에 의한 분산의 변화가 참분산 $\sigma^2(\alpha)$의 그것과 매우 유사하다는 것을 알 수 있다. 반면, Jaeckel의 방법은 절사량의 변화에 따라 불안정한 형태를 나타내고 있다. 특히, 분산을 최소화하는 α의 값이 매우 변화가 많을 것으로 예측된다.

다음의 <표 2>는 4가지의 기저분포에 대하여 붓스트랩방법과 Jaeckel방법의 평균제곱오차(Mean Square Error)를 비교한 결과이다. 모의실험에서 붓스트랩반복은 300회이고 총 반복횟수는 1000회를 시행하였다. <표 2>의 결과로부터 붓스트랩방법이 Jaeckel 방법에 의한 추정보다 모든 기저분포에서 더 작은 평균제곱오차를 가지는 것을 볼 수 있다. 즉, 위의 결과들은 최적절사량 선택에 있어서 붓스트랩방법이 기존의 Jaeckel방법보다 더 우월하다는 것을 나타내고 있다.

<표 2> α-절사평균의 평균제곱오차 비교

기저분포	붓스트랩방법	Jaeckel방법
	평균제곱오차	평균제곱오차
N(0,1)	0.053	0.061
t 분포 (df=5)	0.068	0.072
0.9N(0,1)+ 0.1N(0,4)	0.067	0.078
0.8N(0,1)+ 0.2N(0,4)	0.075	0.082

다음으로 2장에서 제안한 최적절사추정량 m(α^*)의 분산추정방법으로 단순붓스트랩과 더욱 합리적이라고 할 수 있는 이중붓스트랩을 비교하고자 한다. 또한, 위치모수 θ에 대한 신뢰영역의 구축에서도 두 방법을 비교하고자 한다. 먼저, 이 두 방법의 비교는 Tukey(1960)가 제안한 gross error model에서 표본크기 n=20, 단순붓스트랩 반복횟수 B_1=200, 이중붓스트랩 반복횟수 B_2=200을 모두 500회 독립반복시행하여 분산의 평균을 비교하였다. 이 때, 사용된 Tukey의 모형에서 오차밀도함수(error density)는

$$f(x) = \frac{1-\varepsilon}{\sigma}\varphi(x/\sigma) + \frac{\varepsilon}{\tau}\varphi(x/\tau)$$

(단, $\varphi(\cdot)$는 표준정규확률밀도함수)

이다. 모의실험에서는 σ = 1, τ = 4를 사용하였고 오염율이 ε = 0.0, 0.05, 0.1, 0.2 에 대하여 구한 결과가 <표 3>에 제시되어있다.

<표 3> 최적절사평균의 분산 추정 비교

ε	$\widehat{Var}(m(\alpha^*))$	단순붓스트랩		이중붓스트랩	
		평균	표준편차	평균	표준편차
0.00	0.054	0.048	0.019	0.059	0.023
0.05	0.065	0.051	0.022	0.063	0.027
0.10	0.069	0.056	0.025	0.068	0.030
0.20	0.078	0.065	0.029	0.081	0.037

<표 3>에서 보면 이중붓스트랩 분산추정이 단순붓스트랩분산의 과소추정 경향을 개선시키고 있다는 점을 보여주고 있다. 이러한 사실은 다음의 신뢰구간 구축에서도 확인된다.

한편 위의 Tukey모형하에서 θ(=0)에 대한 신뢰구간의 포함확률(coverage probability)을 500번의 독립반복시행을 통하여 구한 결과가 <표 4>에 제시되어 있다. 이 때 붓스트랩신뢰구간은 더 발전된 붓스트랩방법들을 사용하여 구할 수도 있으나, 여러가지 여건상 가장 간단한 분위수(percentile)방법을 사용하여 구하였다. <표 4>는 정규분포하에서도 통상적인 표본평균을 이용한 신뢰구간과 비교했을 때 붓스트랩방법 특히 이중붓스트랩방법은 충분히 잘 작동하고 있음을 보여주고 있다. 앞의 <표 3>에서 단순붓스트랩의 분산이 이중붓스트랩의 그것보다 작은 사실은 포함확률이 과소추정되는 것과 잘 부합된다. 이러한 결과들은 표본의 크기가 20인 경우이지만, 붓스트랩방법의 실제적 타당성은 표본의 크기가 커지면서 더 높아질 것으로 기대된다.

<표 4> 포함확률의 비교

오염율		ε = 0	ε = 0.1	ε = 0.2
표본평균		0.930	0.946	0.942
절사평균 m(α*)	단순붓스트랩	0.900	0.886	0.892
	이중붓스트랩	0.916	0.906	0.910

(단, 명목포함확률은 0.95임)

4. 사례연구

4.1 Newcomb 자료 분석

<표 5>는 Newcomb이 1882년 7월부터 1882년 9월사이에 빛의 통과속도를 측정한 자료이다. 원 자료는 표에 주어진 값에 10^{-3}을 곱하고 24.8을 더한 값이다.

<표 5> Newcomb의 자료

28	-44	29	30
26	27	22	23
33	16	24	29
24	40	21	31
34	-2	25	19

<표 5>로부터 α-절사평균을 구하면 다음과 같다.

표본평균 m(0) = 21.75
m(0.05) = 24.39, m(0.10) = 25.44,
m(0.15) = 25.57, m(0.20) = 25.67
m(0.25) = 25.70, m(0.30) = 25.75
m(0.35) = 25.75, m(0.40) = 25.50

이로부터 표본평균은 α-절사평균에 비해 불안정하며 이것은 표본평균이 특이값에 민감한 영향을 반영한 결과로 추측되어진다. 한편 요즘 과학계에서 가정하고 있는 빛의 통과속도의 참값은 33.02이고 20%-30% 절사평균이 참값에 가깝지만 당시의 측정실험에 있어 체계적 편의가 존재했다는 지적에 따라 참값과의 비교는 합리적이라 할 수 없다.

한편 붓스트랩방법으로 최적절사량을 구하면 α^*=0.35 이고 자료에 근거한 최적절사량은 m(0.35)가 되어 주어진 자료 하에서 최선의 추정이 되고 있음을 알 수 있다.

4.2 Cavendish의 지구밀도 자료 분석

다음의 <표 6>은 1798년 Cavendish가 평균 지구밀도를 측정한 자료이다. 이 자료로부터 붓스트랩방법을 사용하여 구한 최적절사량은 α^* = 0.15이고 이에 따른 최적절사평균은 5.458이다. 한편, 현재 지구밀도의 참값은 5.517로 여기고 있다. 이제 2장에서 제안한 붓스트랩방법을 이용하여 최적절사평균의 붓스트랩분포를 <그림 1>에, 그리고 참값에 대한 붓스트랩신뢰구간을 표본평균을 이용한 통상적인 신뢰구간과 비교하여 <표 7>에 제시하였다. 여기서 붓스트랩 크기는 B_1=300, B_2=300회를 실시하였다.

<표 6> Cavendish의 자료

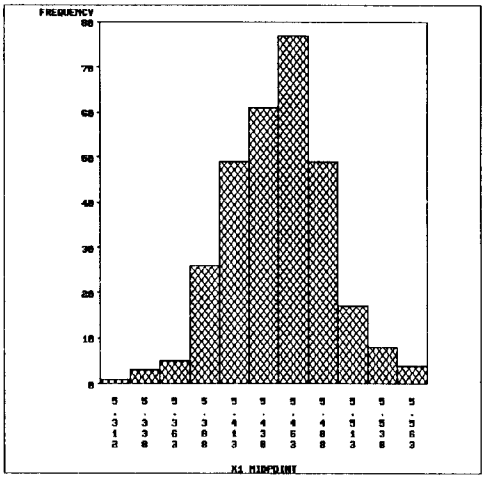
5.50	5.55	5.57	5.34	5.42	5.30
5.61	5.36	5.53	5.79	5.47	5.75
4.88*	5.29	5.62	5.10	5.63	5.68
5.07	5.58	5.29	5.27	5.34	5.85
5.26	5.65	5.44	5.39	5.46	

* 표시의 측정치는 이상치로 간주되고 있음.

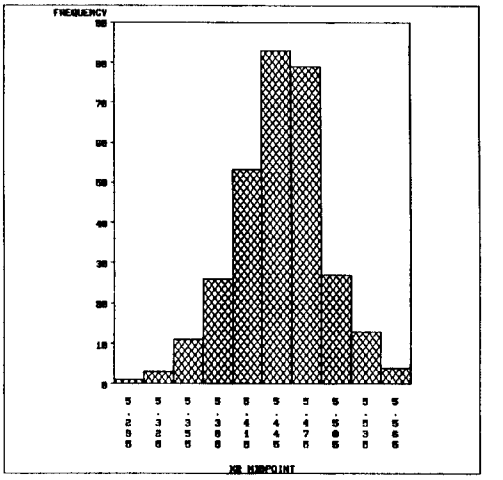
<표 7>에서 참값이라고 여겨지는 5.517이 90%, 95% 신뢰구간에 모두 포함되어 있지만 통상적인 신뢰구간보다 붓스트랩방법을 이용한 신뢰구간이 더욱 바람직하다는 것을 알 수 있으며 특히 이중붓스트랩방법을 이용한 신뢰구간이 단순붓스트랩보다 더 넓으나 포함확률이 정확하다는 점을 유념할 필요가 있다. 또한, <그림 1>을 살펴보면 단순붓스트랩과 이중붓스트랩의 차이는 단순붓스트랩분포의 꼬리부분이 얇은 것으로 보이며 이는 앞에서 언급한대로 단순붓스트랩이 분산을 과소추정하는 것을 잘 나타내고 있다. 붓스트랩신뢰구간에는 앞의 <표 4>를 구할 때와 마찬가지로 분위수방법을 활용하였다.

<표 7> 신뢰구간의 비교

신뢰구간		90% C.I	95% C.I
표본평균		(5.378, 5.518)	(5.364, 5.532)
절사 평균 m(α*)	단순붓스트랩	(5.399, 5.530)	(5.370, 5.551)
	이중붓스트랩	(5.392, 5.547)	(5.380, 5.560)



(단순 붓스트랩분포)



(이중 붓스트랩분포)

<그림 1> 붓스트랩분포의 비교

감사의 글

본 논문을 심사하여 주신 익명의 두 분 심사위원께 깊은 감사를 드립니다.

참고문헌

[1] Beran,R. (1984). Bootstrap Method in Statistics, *Jahresbericht des Deutschen Mathematischen Verein*, 86, 14-30.

[2] Bickel,P.J. and Doksum,K.A. (1977). *Mathematical Statistics,* Holden-Day, Inc.

[3] Efron,B. (1979). Bootstrap Methods: Another Look at the Jackknife, *Annals of Statistics*, 7, 1-26.

[4] Hall,P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York, Inc.

[5] Jaeckel,L.A. (1971). Some Flexible Estimates of Location, *Annals of Mathematical Statistics*, 42, 1540-1552.

[6] Kang,C. (1993). Trimmed Estimators in Statistical Inference, *Ph.D. thesis, Korea University.*

[7] Tukey,J.W. (1960). A Survey of Sampling from Contaminated Distribution, in *Contributions to Probability and Statistics*, Stanford Universty Press, Olkin,I. et al. editors, 448-485.

[8] Stigler,S.M. (1977). Do Robust Estimators Work with Real Data ?, *Annals of Statistics*, 5, 1055-1098.

Bootstrapping Trimmed Estimator in Statistical Inference

Jae Chang Lee[1], Myoungshic Jhun[2], Changwan Kang[3]

Abstract

As an estimate of a location parameter for a given data set, α-trimmed mean has been studied for a long time by many statisticians because of its nice properties including robustness. However, its perfomance depends on the proportion of trimming say α. In this paper, we suggest a data-driven choice of α and study its validity. Also, we suggest a new estimator and consider double-bootstrap to improve its performance. By using simulation study, the proposed method is compared with the exiting one in various cases. Real data sets are also analyzed by using the proposed method.

1) Professor, Department of Statistics, Korea University, 136-701, Seoul, Korea.
2) Professor, Department of Statistics, Korea University, Seoul, 136-701, Korea.
3) Researcher, Institute of Statistic, Korea University, 136-701, Seoul, Korea.