응용통계연구
제 4 권 제 2 호, 1991년

위치 및 척도 모수에 대한 L_1 불편 추정†

성내경* · 임용빈*

요 약

위치와 척도 모수에 대하여 간단히 L_1 불편 추정량을 구하는 방법을 요약 소개하고, L_1 불편 추정량에 대한 자연적인 퍼짐의 측도로 제안된 diffusivity를 이용하는 예를 든다.

1. L_1 불편성

$X=(X_1, \cdots, X_n)$을 미지의 모수 $\theta \in \Theta$로 규정된 분포함수 F를 갖는 모집단에서 추출된 랜덤 표본이라 하자. 관심대상이 되는 모수의 함수 $\tau(\theta)$에 대한 추정량 $\delta(X)$를 결정하려면 추정량의 성능을 사정할 수 있는 잘 정의된 판정 기준이 필요하다. 이러한 판정 기준을 선택하는 일반적인 접근 방법은 모수에의 근접도를 재는 측도를 마련하고, 이 근접도의 측도를 최소화하는 추정량을 선택하는 것이다. 그러나, 이를 만족하는 추정량은 일반적으로 구할 수 없으므로, 찾고자 하는 추정량이 가져야 할 성질을 제한한 후, 제한된 범위의 추정량 중 근접도의 측도를 최소화하는 최적 추정량을 선택하는 방법이 흔히 쓰인다. 대표적인 제한된 추정량의 부류가 불편 추정량이라 할 수 있다. 불편성에 대한 일반적인 개념은 Lehmann(1951)에 의하여 주어졌다.

모수의 참값이 θ일 때 $\tau(\theta)$에 대한 추정량 $\delta(X)$로 기인되는 손실이 $L(\theta, \delta(X))$인 추정문제를 생각하자. Lehmann(1983)은 다음과 같은 조건이 모든 $\theta \in \Theta$에 대하여 만족될 때 추정량 $\delta(X)$는 $\tau(\theta)$에 대하여 위험 불편이라 정의하였다.

$$E_\theta L(\theta, \delta(X)) \leq E_\theta L(\theta', \delta(X)), \quad \theta' \neq \theta \tag{1}$$

랜덤 손실 함수 $L(\theta, \delta(X))$는 $\tau(\theta)$에 대하여 $\delta(X)$에 주어진 선호도를 표현한다. 위험 불편성의 관점에서 우리의 주된 목표는 $\tau(\theta)$에 평균적으로 충분히 근접하는 올바른 추정량을 선택하는 것이다.

---

† 본 연구는 1990년도 문교부 기초과학연구비 지원(과제번호 BSRI-90-108)에 의한 연구결과임.
* (120-750) 서울특별시 서대문구 대현동 이화여자대학교 통계학과

그러나, 기대 손실, 즉, 위험의 함수 형태를 규정하는 자연스러운 방법은 없다. 위험 불편 추정과 관련된 상대적으로 폭넓은 추정량의 부류는 민코우스키 거리, 혹은 L_p 거리를 최소화하여 얻어지는 것들로서, 이 추정량들을 L_p 불편 추정량이라 부른다. L_p 불편 추정에 관하여는 Barankin(1949)의 연구가 있다. 만일 $p=2$이면, (1) 식의 손실은 제곱 오차 손실로 귀착되며, 위험은 따라서 유클리디안 거리가 된다. 추정 문제에 대한 고전적인 방법은 전부 유클리디안 거리를 기준으로 한다. 이 경우, 위험은 $E_\theta\delta(X)=\tau(\theta)$일 때 최소화된다. 즉, 위험 불편성의 조건은 통상의 평균 불편성 혹은 L_2 불편성으로 축소된다.

흡사하게 만일 $p=1$로 놓으면 절대거리를 얻게 되며, 따라서 (1)식은 다음과 같이 된다.

$$E_\theta\,|\,\delta(X)-\tau(\theta)\,|\leq E_\theta\,|\,\delta(X)-\tau(\theta')\,|,\quad \theta'\neq\theta \tag{2}$$

잘 알려진 사실이지만 (2) 식은 임의의 $\delta(X)$에 대한 중앙값으로 최소화된다. 결국 (2)식은 L_1 불편성, 혹은 중앙값 불편성의 조건으로 축소되며, 수식으로 표현하면

$$\mathrm{median}_\theta\delta(X)=\tau(\theta).$$

L_2 불편 추정에 관한 문제는 지금까지 수많은 문헌에서 다루어졌으나 L_1 불편 추정은 상대적으로 드물게 다루어졌다. 이 이유는 많은 경우에 L_1 불편 추정량을 닫힌 형태로 나타내기가 쉽지 않고, 또, 이산 분포에 대하여 L_1 불편 추정량을 정의하기가 어려운데 기인한다. 게다가 결합 추정의 경우에는 아직까지 일반적으로 받아들여지는 L_1 불편성의 정의가 마련되지 않은 탓도 있다(다변량 L_1 불편성의 한가지 정의에 관하여는 Sung(1990a)을 들 수 있다).

이상과 같은 문제점도 있으나 L_1 불편 추정이 L_2 불편 추정에 비하여 강점을 갖고 있는 것도 사실이다. 무엇보다도 L_1 불편성은 협의의 단조 변환에 대하여 불변이다. 이 성질은 L_2 불편 추정량에서는 일반적으로 성립하지 않는다. 게다가 Birnbaum(1963)이 지적하였듯이 오차의 분포로 정규분포가 가정되는 통상적인 선형 모형론에서 최상의 L_2 불편 추정량들은 전부 정규 분포를 하고, 따라서 이 추정량들은 L_2 불편 추정량인 동시에 L_1 불편 추정량이 된다. 이 경우, 최상 L_2 불편 판정 기준을 최상 L_1 불편 판정 기준으로 바꾸어도 통계분석에는 아무런 문제가 없다. 또한, Huber(1987)에 의하여 잘 요약된 바, L_1 불편 추정량은 정의 자체가 중앙값을 기준으로 하고 있는 만큼 대체로 L_2 불편 추정량보다 더 로버스트하다.

이 소고에서는 위치 및 척도 모수에 대하여 간단히 L_1 불편 추정량을 구하는 방법들을 요약하고, Sung(1990a, c)에 의하여 제안된 L_1 불편 추정량에 대한 일종의 퍼짐의 측도인 diffusivity를 이용하여 코쉬 분포와 같은 경우에는 L_1 불편 추정량을 사용하는 편이 우월함을 정량화 할 수 있음을 보인다.

2. 위치와 척도 모수의 추정

$X=(X_1, \cdots X_n)$은 결합 분포 함수가 $F(x\,;\theta)$, 결합 밀도함수가 $f(x\,;\theta)$인 랜덤 표본으로 X_i은 연속 확률변수이다. M을 $X_1, \cdots X_n$의 표본 중앙값이라 하고, m을 F의 모집단 중앙값이라 하자. 이 표기는 앞으로도 계속 쓰인다. 다음 결과는 잘 알려진 것으로 van der Vaart(1961) 등에 나와

있다.

정리 1. 만일 n이 홀수이면 M은 m에 대하여 L_1 불편이다.

표본 크기에 제한이 가해지기는 하나 정리 1을 위치 모수족이나 척도 모수족에 속하는 분포들에 직접 적용하면 위치 모수나 척도 모수에 대한 L_1 불편 추정량을 표본 중앙값의 함수로 나타낼 수 있음은 자명하다. 이 결과는 다음의 정리로서 간단히 요약할 수 있다.

정리 2. 위치 모수족에 속하는 분포 $f(x;\theta)=h(x-\theta)$에 대하여, 만일 n이 홀수이면 $M-(m-\theta)$는 위치 모수 θ에 대하여 L_1 불편이다. 또한, 척도 모수족에 속하는 분포 $f(x;\sigma)=g(x/\sigma)/\sigma$, $\sigma>0$, 에 대하여 만일 n이 홀수이고 $m\neq 0$이면, $M\sigma/m$은 척도 모수 σ에 대하여 L_1 불편이다.

예 1. 만일 $X_1, \cdots X_n$이 iid 확률 변수들로 다음과 같은 극값 밀도를 갖는다면,

$$f(x;\theta)=\exp[-(x-\theta)-\exp\{-(x-\theta)\}]$$

n이 홀수일 때, $M+\log\log 2$는 θ에 대하여 L_1 불편이다.

예 2. 만일 $X_1, \cdots X_n$이 iid 확률 변수들로 다음과 같은 지수 밀도를 갖는다면,

$$f(x;\sigma)=\exp(-x/\sigma)/\sigma, \quad x>0,$$

n이 홀수일 때, $M/\log 2$는 σ에 대하여 L_1 불편이다.

정리 2에서 m=0인 경우, 즉 일례로 분포가 대칭인 경우에는 표본 중앙값의 함수로 척도 모수에 대한 L_1 불편 추정량을 구할 수 없다. 그러나, 대칭 분포인 경우에는 잘 알려진 분포를 갖는 표본 통계량을 선택하고 L_1 불편 추정량이 협의의 단조 변환에 대하여 불변임을 이용하면 의외로 쉽게 L_1 불편 추정량을 구할 수 있다.

예 3. $X_1, \cdots X_n$이 iid $N(0,\sigma^2)$ 확률 변수들이라면,

$$\sum_{i=1}^{n} X_i^2/\sigma^2 \sim \chi_n^2$$

이므로, $(\Sigma X_i^2/\chi^2_{n,0.5})^{1/2}$은 σ에 대하여 L_1 불편이다. 여기서 $\chi^2_{n,0.5}$는 자유도의 n의 χ^2 분포의 중앙값이다.

예 4. $X_1, \cdots, X_n$이 다음과 같은 이중 지수 분포를 따르는 iid 확률 변수들이다.

$$f(x;\sigma)=\exp(-|x|/\sigma)/2\sigma$$

$2\sum_{i=1}^{n}|X_i|/\sigma \sim \chi^2_{2n}$이므로, $2\Sigma|X_i|/\chi^2_{2n,0.5}$는 σ에 대하여 L_1 불편이다.

위치 모수에 대한 L_1 불편 추정에서 잘 알려진 통계량의 분포가 위치 모수를 중심으로 대칭형일 때는, 바로 그 통계량을 L_1 불편 추정량으로 선택하는 편이 유리하다. 예를 들어, 정규 분포 $N(\mu,1)$을 따르는 랜덤 표본에서 표본 평균은 모평균에 대하여 L_1 불편이다.

위치 모수의 추정 문제는 언제나 척도 모수의 추정 문제로 변환될 수 있고, 그 역도 가능하다.

다음 정리는 이 사실을 요약한 것이다.

정리 3. X가 위치 모수가 θ인 분포 $f(x;\theta)=h(x-\theta)$를 따르는 확률 변수라면, exp(X)는 척도 모수가 $\exp(\theta)$인 척도 모수족에 속하는 분포를 갖는다. 역으로, 확률변수 X의 분포가 $f(x;\sigma)=g(x/\sigma)/\sigma$, $x>0$, $\sigma>0$, 이라면 log(X)는 위치 모수가 $\log(\sigma)$인 위치 모수족에 속하는 분포를 갖는다.

정리 2와 연관되어 정리 3은 위치와 척도 모수에 관한 L_1 불편 추정에서 자주 언급되는 바, 지수 함수와 로그 함수 변환은 모두 협의의 단조 변환이므로 L_1 불편성이 이들 변환하에서는 불변이기 때문이다.

3. 상대 효율

앞 절에서 위치와 척도 모수에 대한 L_1 불편 추정량을 구하는 간단한 방법과 뒷받침되는 변환 등을 소개하였는데, L_1 불편 추정량은 표본 중앙값의 함수로뿐만 아니라 어떠한 합리적인 통계량으로도 표현할 수 있음은 자명하다. 하나의 모수에 대하여 둘 이상의 L_1 불편 추정량이 존재할 때 대두되는 한가지 문제는 최적 추정량을 결정할 때 사용되는 추정량에 대한 적절한 퍼짐의 측도의 선택이다. 통상적인 분산은 Lehmann의 위험 불편 판정 기준의 관점에서 L_2 불편 추정량에 대한 자연스러운 퍼짐의 측도라 하겠으나 정의된 공간 자체가 다르므로 L_1 불편 추정에도 분산을 도입하는 것은 부적절해 보인다. 이런 관점에서 Sung 등(1990c)에 의하여 새로 제안된 L_1 불편 추정량의 측도로 diffusivity를 고려할 수 있다.

$Y\equiv\delta(x)$가 $\tau(\theta)$에 대한 중앙값 불편의 추정량이고, $g_y(y;\theta)$를 Y의 연속 확률밀도라면 diffusivity는 $1/\{2g_y(\tau(\theta);\theta)\}$로 정의된다. 즉, diffusivity는 확률밀도의 높이에 반비례하는 함수인데 이와 흡사한 한가지 예로, 위치 모수족에 속하는 분포의 중앙값이 유일할 때, 표본 중앙값의 근사 표준편차가 위치 모수점에서 산출된 확률 밀도의 높이에 반비례 한다는 사실을 들 수 있다.

그러나, 때에 따라서 L_1 불편 추정량에 대하여 효율성을 측정할 때, 퍼짐의 측도로서 표준편차를 사용하든 diffusivity를 사용하든 동일한 값을 얻는 경우도 있다. 예를 들어, Sung 등(1990c)이 밝힌 바에 따르면 정규분포 $N(\theta,1)$과 같은 경우에는 표본 평균과 표본 중앙값이 둘다 θ에 대하여 L_1 불편이나, 상대 효율도로 두 통계량의 표준편차의 비를 사용하든 diffusivity의 비를 사용하든 그 값은 서로 동일하고, 표본 평균이 표본 중앙값보다 더 낫다는 결과를 얻고 있다. 다음 예는 이와 유사한 결과를 보이는 예이다.

예 5. 만일 $X_1,\cdots,X_n$이 iid 확률 변수들로 다음과 같은 로그 정규 밀도를 갖는다면,

$$f(x;\sigma)=\exp[-(\log(x/\sigma)^2/2]/x\sqrt{2\pi},\quad x>0$$

표본 중앙값 M과 $(\prod_{i=1}^{n}X_i)^{1/n}$이 둘다 σ에 대하여 L_1 불편이나 M이 덜 효율적이다.

L_2 불편 추정의 관점에서 코쉬 분포의 경우에 표본 평균은, 표본 평균의 분포가 단일 관측의

분포와 일치하므로 위치 모수에 대한 아주 열등한 추정량이다. 따라서, 일반적으로 표본 중앙값이 표본 평균보다 더 낫다고 생각되어지나 L_2 불편 추정 형식의 처리로는 표본 평균이나 표본 중앙값 중 어느 것도 일차 이상의 적률을 갖지 못하므로 이러한 짐작을 확인할 수 없다. 다음 예의 요점은 diffusivity를 이용하여 이러한 가설을 정량화 할 수 있음을 보이는 데 있다.

예 6. $X_1, \cdots, X_n$이 다음과 같은 코쉬 분포를 따르는 iid 확률 변수들이라 하자.

$$f(x;\theta)=1/\pi[1+(x-\theta)^2].$$

코쉬에서는 표본 평균 $\bar{X}$와 표본 중앙값 M이 둘다 θ에 대하여 L_1 불편이다. 아래 표에서는 몇가지 n 값, $n=2k+1$, $k=0,1,\cdots$, 에 대하여 표본 평균과 표본 중앙값에 대하여 산출된 diffusivity를 보였고, 가장 오른쪽의 하한은 Sung(1990a)에 의하여 주어진 L_1 불편 추정량에 대한 Cramer−Rao 형식의 diffusivity에 대한 하한값으로 diffusivity들이 하한값보다 더 작을 수는 없다. 주어진 diffusivity들은 예를 들어 Yakowitz(1977)에서 제시된 것과 같은 표본 평균 몬테칼로 앨고리듬을 사용하여 얻은 근사값들이다.

k	$\text{diffusivity}_\theta(\bar{X})$	$\text{diffusivity}_\theta(M)$	하한
0	1.56	1.56	1.56
1	1.56	1.05	0.99
2	1.56	0.84	0.78

결과를 보면, 두 값이 같은 k=0일 때를 제외하면 표본 중앙값이 표본 평균보다 더 나은 L_1 불편 추정량이라고 할 수 있다. 그러나, 표본 중앙값도 k=0일 때를 제외하면 하한을 구현하지 못한다.

참 고 문 헌

[1] Barankin, E.W.(1949). "Locally Best Unbiased Estimates," *Annals of Mathematical Statistics*, 20, 477-501.

[2] Birnbaum, A.(1964). "Median−unbiased Estimators," *Bulletin of Mathematical Statistics*, 1, 25-34,

[3] Huber, P.J.(1987). The Place of the L_1−norm in Robust Estimation, in *Statistical Data Analysis Based on the L_1−norm and Related Methods* Edited by Dodge, Y., 23-34, North-Holland.

[4] Lehmann, E.L.(1951). "A General Concept of Unbiasedness," *Annals of Mathematical Statistics*, 22, 587-592.

[5] Lehmann, E.L.(1983). *Theory of Point Estimation*. John Wiley & Sons, Inc., New York.

[6] Sung, N.K.(1990a). "A Generalized Cramer−Rao Analogue for Median−unbiased Estimators," *Journal of Multivariate Analysis*, 32, 204-212.

[7] Sung, N.K. (1990b). "An Optimality Criterion for Median−unbiased Estimators," *Journal of the Korean Statistical Association*, **19**, 176-181.

[8] Sung, N.K., Stangenhaus, G., and David, H.T. (1990c). "A Cramer−Rao Analogue for Median−Unviased Estimators," *Trabajos de Estadistica*, **5**, 83-94.

[9] van der Vaart, H.R. (1961). "Some Extensions of the Idea of Bias," *Annals of Mathematical Statistics*, **32**, 436-447.

[10] Yakowitz, S.J. (1977). *Computational Probability and Simulation*. Addison−Wesley, Reading.

A Note on L_1 – unbiased Estimation for Location and Scale Parameters†

Nae-Kyung Sung* and Yong-Bin Lim*

ABSTRACT

A few simple methods to find L_1 – unbiased estimators for location and scale parameters are summarized and examples utilizing diffusivity, a natural measure of dispersion for L_1 – unbiased estimators, are given.

† Research supported in part by the Basic Science Research Institute Program, Ministry of Education, 1990, under Project No. BSRI – 90 – 108.

* Statistics Department, Ewha Women's University, Daehyundong, Seodaemungu Seoul 120 – 750, Korea.